



(12)发明专利申请

(10)申请公布号 CN 107614689 A

(43)申请公布日 2018.01.19

(21)申请号 201680030538.4

(22)申请日 2016.03.29

(30) 优先权数据

2015901121 2015.03.27 AU

(85)PCT国际申请进入国家阶段日

2017.11.27

(86)PCT国际申请的申请数据

PCT/AU2016/050239 2016.03.29

(87)PCT国际申请的公布数据

W02016/154675 EN 2016.10.06

(71) 申请人 昆士兰大学

地址 澳大利亚昆士兰省

(72)发明人 谢尔盖·穆尔伊夫 正林·崔

基里尔·亚历山德罗夫

(74)专利代理机构 广州市天河区倪律专利代理
事务所(普通合伙) 44348

事务所(普通合伙) 44348

代理人 倪小敏

(51) Int.Cl.

C12N 15/67(2006.01)

C12P 21/02(2006.01)

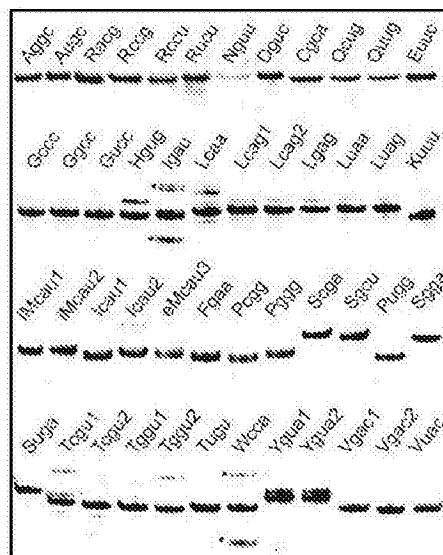
权利要求书2页 说明书49页 附图21页

(54)发明名称

用于将非天然氨基酸并入蛋白质中的平台

(57)摘要

本发明提供了一套tRNA和蛋白质翻译系统，其使得能够并入非天然部分(例如非天然氨基酸)而不损害将所有二十种天然氨基酸并入蛋白质的能力。这通过将通常由至少两种不同的tRNA反密码子解码的氨基酸的tRNA反密码子之一重新分配给非天然部分来实现，其中至少一个密码子可以被重新分配的反密码子独特地识别并且至少另外一个来自同一个密码子框的密码子不能被重新分配的反密码子识别。因此，用于翻译的mRNA被工程化以包含对应于重新分配的tRNA反密码子的一个或多个特定密码子，使得非天然部分在选定位置被并入翻译的蛋白质中。



1. 一种生产适于包含至少一个非天然部分的蛋白质的翻译的一套tRNA的方法,所述方法包括将至少一个包含天然氨基酸的反密码子的tRNA替换为至少一个包含相同但被重新分配为非天然部分的反密码子的tRNA,其中所述一套tRNA能够被操作而便于RNA的翻译,所述RNA包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,从而使得所翻译的蛋白质可包含任何一种或全部二十种(20)天然氨基酸。

2. 根据权利要求1所述的方法,所述方法包括一个或多个以下步骤: (i) 从适于包含天然氨基酸的蛋白质的翻译的一套tRNA中耗尽一种或多种tRNA; 和 (ii) 以一个或多个分别被重新分配至非天然部分并且分别被连接至所述非天然部分的tRNA来重构该被耗尽的一套tRNA。

3. 根据权利要求2所述的方法,其中在步骤(i)中,基本上所有天然氨基酸的tRNA都被从该套tRNA中耗尽。

4. 根据权利要求2所述的方法,其中在步骤(i)中,一种或多种的天然氨基酸tRNA被选择性地从该套tRNA中耗尽。

5. 根据权利要求2、3或4所述的方法,其中在步骤(ii)中,重构tRNA包含合成tRNA、天然tRNA或其混合物。

6. 一种组合物,其包含适于包含至少一个非天然部分的蛋白质的翻译的一套tRNA,所述一套tRNA包含至少一个含有重新分配为非天然部分的天然氨基酸反密码子的tRNA,其中该一套tRNA能够被操作而便于RNA的翻译,所述RNA包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,从而使得所翻译的蛋白质可包含任何一种或全部二十种(20)天然氨基酸。

7. 一种生产适于包含至少一个非天然部分的蛋白质的翻译的翻译系统的方法,所述方法包括生产包含至少一个tRNA的一套tRNA,该至少一个tRNA包含已被重新分配为非天然部分的天然氨基酸反密码子;以及生产可转录RNA,其包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,其中该mRNA可被转录而生产出可包含任何一种或全部二十种(20)天然氨基酸的已翻译蛋白质。

8. 一种适于包含至少一个非天然部分的蛋白质的翻译的翻译系统,所述系统包括:包含至少一个tRNA的一套tRNA,该至少一个tRNA包含已被重新分配为非天然部分的天然氨基酸反密码子;以及可翻译mRNA,其包含与已被重新分配为非天然部分的所述反密码子相对应的密码子,其中该mRNA可被转录而生产出可包含任何一种或全部二十种(20)天然氨基酸的已翻译蛋白质。

9. 一种生产包含至少一个非天然部分的重组蛋白质的方法,所述方法包括翻译mRNA的步骤,该mRNA包含与一套tRNA中的一个tRNA中已被重新分配为非天然部分的反密码子相对应的密码子,该一套tRNA包含至少一个含有已被重新分配为非天然部分的天然氨基酸反密码子的tRNA,其中所翻译出的蛋白质可包含任何一种或全部二十种(20)天然氨基酸。

10. 前述权利要求任一项所述的组合物、方法或系统,其中该被重新分配的反密码子是多个通常是天然氨基酸的不同反密码子中的一个。

11. 前述权利要求任一项所述的组合物、方法或系统,其中氨酰化tRNA是体外合成的。

12. 根据权利要求11所述的组合物、方法或系统,其中tRNA利用RNA合成酶来生产。

13. 根据权利要求12所述的组合物、方法或系统,其中tRNA利用RNA合成酶来氨酰化。

14. 根据权利要求13所述的组合物、方法或系统,其中所述RNA合成酶选自Py1RS或其适于吡咯赖氨酸tRNA合成酶介导的氨酰化的变体;詹氏甲烷球菌酪氨酰转移RNA合成酶(*Mj* TyrRS)或其变体);Flexizyme;以及半胱氨酰tRNA合成酶。

15. 根据权利要求14所述的组合物、方法或系统,其中包含被重新分配的密码子的tRNA是突变的半胱氨酸tRNA。

16. 根据权利要求15所述的组合物、方法或系统,其中氨酰化tRNA利用半胱氨酰tRNA合成酶在体外合成。

17. 根据权利要求16所述的组合物、方法或系统,其中突变的半胱氨酸tRNA不能通过内源性半胱氨酰tRNA合成酶而被再负载。

18. 前述权利要求任一项所述的方法、组合物或系统,其中重新分配的反密码子是四倍或六倍简并的反密码子。

19. 前述权利要求任一项所述的方法、组合物或系统,其中重新分配的反密码子是Ile、Ala、Gly、Pro、Thr、Val、Arg、Leu或Ser的反密码子。

20. 前述权利要求任一项所述的方法、组合物或系统,其中该套tRNA包含两种或更多种不同的、通常编码相应天然氨基酸、被重新分配至相应非天然部分的反密码子。

21. 前述权利要求任一项所述的方法、组合物或系统,其中该套tRNA包含至少一个含有不是用于天然氨基酸的反密码子的tRNA,其中所述反密码子已被重新分配至非天然部分。

22. 根据权利要求21所述的方法、组合物或系统,其中已被重新分配至所述非天然部分的所述反密码子是琥珀抑制剂反密码子。

23. 通过权利要求9至22的任一项的方法生产的重组蛋白质。

24. 根据权利要求23所述的重组蛋白质,其包含一个或多个相同或不同的非天然部分,所述非天然部分便于PEG化、结合至小分子、标记、固定、分子间和/或分子内交联或其他相互作用、形成更高级别的结构和/或一个或多个催化活性。

25. 根据权利要求23所述的重组蛋白质,其包含两个或更多个相同或不同的能够实现分子内共价键合的非天然部分。

26. 根据权利要求23、24或25所述的重组蛋白质,其中所述重组蛋白质是大环状蛋白质。

27. 一种蛋白质文库,其包含多种根据权利要求23至26的任一项所述的重组蛋白质。

28. 一种mRNA,其编码权利要求23至26的任一项的重组蛋白质,或其编码权利要求27所述的蛋白质文库中的多种重组蛋白质中的至少一种。

用于将非天然氨基酸并入蛋白质中的平台

技术领域

[0001] 本发明涉及蛋白质工程。更具体地,本发明涉及生产包含一个或多个非天然部分(如非天然氨基酸)以及所有二十种(20)天然氨基酸中的任何一种的重组蛋白。

背景技术

[0002] 蛋白质是所有有生命的生物体的核心功能成分。蛋白质由天然生物合成途径形成,该途径使用二十种(20)遗传编码的“天然”氨基酸的遗传学上确定的序列以形成这些蛋白质的结构单元。鉴于生产包含天然氨基酸的蛋白相对简单,因而迄今为止,蛋白质结构、功能和种类方面的研究进展大部分集中在天然结构、功能和种类。虽然人类基因组编码约30000种不同的蛋白质,但如果能够制造出包括天然氨基酸之外的结构部分(如非天然氨基酸、天然氨基酸的化学衍生物、具有化学反应性的部分等)的蛋白质,那么其种类可能会更加庞大。这些“非天然”蛋白质可被用来产生合成蛋白库、基于蛋白质的生物活性物(如药物)、生物治疗剂(如抗体、肽类激素、免疫调节剂)、检测或诊断试剂,它们具有在天然蛋白质中未发现的结构和活性。但是,这种“非天然”蛋白质的重组生产已被证明是一个困难的命题。

[0003] 一般而言,生产重组“非天然”蛋白质的方法依赖于对蛋白翻译过程进行操作(特别是在无细胞表达系统中进行),这种操作破坏正交tRNA为蛋白质链延长的每个步骤提供特定的天然氨基酸的正常功能。通过酶解法或化学方法可用非天然氨基酸使正交tRNA氨酰化。酶法氨酰化需要使用正交氨酰tRNA合成酶(aaRS)。通常需要多轮定向进化才能生成可接收非天然氨基酸作为其底物的正交tRNA/aaRS对。虽然使用这些方法已将数百种非天然氨基酸并入到蛋白质中,但是它们大多数属于一个窄的结构类别(通过它们的“亲本”aaRS的特异性来界定)。

[0004] 与此概念明显不同的是使用能够在体外将非天然氨基酸负载到tRNA上的一种进化的tRNA酰化核酶(Flexizyme)。该方法是通过与tRNA的3'-端形成暂时的复合物来实现的,其中末端核糖的3'-OH与非天然氨基酸的活化的羧基碳通过亲核攻击而结合。

[0005] 此外,还开发了化学酰化tRNA方案。该方案可并入与天然氨基酸非常不同的化学官能团,因而难以通过aaRS途径来负载。虽然已开发出几种合成途径,但通常它们的产率低并且未广泛应用。

[0006] 其他方案集中于在蛋白质翻译过程中通过指定具有正交反密码子的tRNA来编码非天然氨基酸而并入非天然氨基酸。这还要求翻译的mRNA在非天然氨基酸要被并入的位置包括相应的“重新分配”的密码子。生成正交密码子的最常用方法是“无义抑制”,其使用遗传密码中的三个终止密码子:琥珀密码子、蛋白石密码子和赭石密码子。在一些生物体中,终止密码子可被通读并编码非天然氨基酸,例如UGA密码子编码硒代半胱氨酸,UAG密码子编码吡咯赖氨酸。该系统的主要缺陷在于释放因子与正交抑制tRNA的竞争导致蛋白质表达产率内在降低。

[0007] 另一种方案是“移码抑制”。包含长的反密码子环的天然的移码抑制tRNA可读四碱

基密码子,如UAGN和ACCN,并抑制+1阅读框移位。但是,通过四碱基密码子编码非天然氨基酸的方案效率远不如使用无义抑制,目标蛋白质的产率因为三联密码子被误读为三联密码子而显著降低。近期的一个氨基酸并入策略是从遗传密码中排除某些氨基酸(如Phe),从而生成可被指定至非天然氨基酸的“游离”密码子。但是,这导致氨基酸种类的减少。

发明内容

[0008] 本发明广义上来说提供一种方法,通过该方法tRNA反密码子可从天然氨基酸被“重新分配”至非天然部分(例如非天然氨基酸),从而能够生产包括一个或多个非天然部分的重组蛋白质而不损害将天然氨基酸并入该重组蛋白质的能力。

[0009] 在第一个方面,本发明提供一种生产适合于包含至少一个非天然部分的蛋白质的翻译的一套tRNA的方法,所述方法包括将至少一个包含天然氨基酸的反密码子的tRNA替换为至少一个包含相同但被重新分配为非天然部分的反密码子的tRNA,其中所述一套tRNA能够被操作而便于RNA的翻译,所述RNA包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,从而使得所翻译的蛋白质可包含任何一种或全部二十种(20)天然氨基酸。

[0010] 适当地,该被重新分配的反密码子是该相同的天然氨基酸的多个不同反密码子中的一个。

[0011] 在第二个方面,本发明提供一种组合物,其包含适合于包含至少一个非天然部分的蛋白质的翻译的一套tRNA,所述一套tRNA包含至少一个含有非天然部分的反密码子的tRNA,所述非天然部分的反密码子由天然氨基酸的反密码子重新分配,其中该一套tRNA能够被操作而便于RNA的翻译,所述RNA包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,从而使得所翻译的蛋白质可包含任何一种或全部二十种(20)天然氨基酸。

[0012] 适当地,该被重新分配的反密码子是该相同的天然氨基酸的多个不同反密码子中的一个。

[0013] 在第三个方面,本发明提供一种生产适于包含至少一个非天然部分的蛋白质的翻译的翻译系统的方法,所述方法包括:生产包含至少一个tRNA的一套tRNA,该至少一个tRNA包含已被重新分配为非天然部分的天然氨基酸反密码子;以及生产可转录RNA,其包含与已被重新分配为所述非天然部分的所述反密码子相对应的密码子,其中该mRNA可被转录而生产出可包含任何一种或全部二十种(20)天然氨基酸的已翻译蛋白质。

[0014] 在第四个方面,本发明提供一种适于包含至少一个非天然部分的蛋白质的翻译的翻译系统,所述系统包括:包含至少一个tRNA的一套tRNA,该至少一个tRNA包含已被重新分配为非天然部分的天然氨基酸反密码子;以及可翻译mRNA,其包含与已被重新分配为非天然部分的所述反密码子相对应的密码子,其中该mRNA可被转录而生产出可包含任何一种或全部二十种(20)天然氨基酸的已翻译蛋白质。

[0015] 在第五个方面,本发明提供一种生产包含至少一个非天然部分的重组蛋白质的方法,所述方法包括翻译mRNA的步骤,该mRNA包含与一套tRNA中的一个tRNA中已被重新分配为非天然部分的反密码子相对应的密码子,该tRNA包含至少一个含有已被重新分配为非天然部分的天然氨基酸反密码子的tRNA,其中所翻译出的蛋白质可包含任何一种或全部二十

种(20)天然氨基酸。

[0016] 适当地,根据前述各个方面的任何一个的方法、组合物或系统,所述重新分配的反密码子是四倍或六倍简并的。

[0017] 适当地,根据前述各个方面的任何一个的方法、组合物或系统,所述重新分配的反密码子是Ile、Ala、Gly、Pro、Thr、Val、Arg、Leu或Ser的反密码子。

[0018] 在前述各个方面的任何一个的方法、组合物或系统的具体实施方式中,所翻译出的蛋白质可包含一个或多个相同或不同的非天然部分。

[0019] 在第六个方面,本发明提供一种由第五个方面的方法生产的重组蛋白质。

[0020] 将意识到的是,分离的蛋白质可包含一个或多个相同或不同的非天然部分,而便于PEG化、结合至小分子、标记、固定、分子间和/或分子内交联或其他相互作用、形成更高级别的结构和/或一个或多个催化活性,但不限于此。

[0021] 在一个实施方式中,所述重组蛋白质包含两个或更多个相同或不同的能够实现分子内共价键合的非天然部分。在一个特定实施方式中,所述重组蛋白质是大环状蛋白质(macrocyclic protein)。

[0022] 本发明还提供编码第六个方面的所述重组蛋白质的mRNA分子。

[0023] 将意识到的是,不定冠词“一个(a)”和“一个(an)”不能被解读为单数不定冠词或以其他方式排除多于该不定冠词所指的单个对象的情况。例如,“一个(a)”蛋白包括一个蛋白、一个或多个蛋白或多个蛋白。

[0024] 在本文中,除非另有说明,否则术语“包含”和“包括”将被理解为包括所描述的一个整数或一组整数,但不排除其他整数或整数组。

附图说明

[0025] 图1. 48种体外合成的t7tRNA的变性PAGE分析。这些t7tRNA以各自单字母氨基酸代码(大写)和5' -3' 反密码子三联体(小写)来表示。tRNA多态性变体以各数字表示。Ile和Trp的tRNA(Igau和Wcca)通过含HHRz的RNA前体的自动剪接而产生。相应tRNA条带之上和之下的星号分别表示该前体和剪接后的HHRz。起始tRNA和延伸tRNA(Met)分别加前缀“i”和“e”。

[0026] 图2.亲和钳肽生物传感器(affinity clamp peptide biosensor)的原理和校准。(A)实验方案示意图。顶部RNA序列代表RGS肽及其衍生物RGS1和RGS2的编码框架,后者包含“绝缘”密码子(绿色)及其后的测试密码子三联体(XXX,红色)。所产生的包含恒定的八氨基酸C末端和可变N末端的肽结合至由自身抑制的(autoinhibited)的TVMV蛋白酶构成的生物传感器,PDZ和FN3结构域形成所述亲和钳。肽的结合导致亲和钳的构象变化,进而将该抑制性肽驱离TVMV的活性位点,导致蛋白酶活化以及随后的淬灭的底物的酶切。Q和F分别代表荧光淬灭剂和荧光团。(B)通过在热变性的E.coli提取物中分析不同浓度的合成肽而获得的RGS肽的校准曲线。以初始速率对肽浓度作图,用回归系数(R²)拟合三次实验的均值。

[0027] 图3.eGFP和RGS肽在tRNA耗尽的E.coli裂解物中的无细胞翻译。(A)天然tRNA依赖性eGFP表达。(B)通过阻止t7tRNA混合物中的单个t7tRNA并调节该tRNA耗尽的E.coli裂解物中的RGS肽翻译来分析对于每个密码子天然tRNA的耗尽效率。与被单个排除的t7tRNA相

对应的密码子/反密码子对位于相应柱子下方。补充延伸t7tRNA至终浓度为0.8μM,起始t7tRNAiMet终浓度为1.6μM,天然tRNA混合物终浓度为1μg/μl。

[0028] 图4.t7tRNA解码表。Ser、Arg和Leu(蓝色底纹)由混合的密码子家族盒(mixed codon family boxes)编码,其中两个密码子(N1N2N3)属于分割(split)密码子家族盒,其余四个属于未分割(unsplit)密码子家族盒¹⁹。4重和2倍简并氨基酸分别为绿色和灰色底纹。天然tRNA和t7tRNA以它们各自的反密码子表示(N34N35N36)。天然tRNA反密码子中的A/U/G/C之外的字母表示修饰的核苷⁵⁵。天然和t7tRNA解码模式分别在密码子栏的左右两侧以箭头表示。对于Lys、Glu和Ile特异性tRNA的反密码子,其中需要在密码子或其他部分进行修饰以实现氨酰化的反密码子以粉色高亮表示。具有U34到C34反密码子替换(基于tRNA^{Lys}(UUU))的Lys t7tRNA以红色高亮表示。连接t7tRNA和各密码子的箭头表示在肽生物传感器实验中测试的tRNA/密码子组合。灰色和黑色连续虚线箭头分别对应于<10%或≥10%密码子解码效率。黑色箭头旁的关联数字表示图11中所述的t7tRNA对于所分析的密码子的计算出的解码效率。N34修饰包括“V”-尿苷5-羟乙酸,“[”-5-甲基-氨基甲基尿苷(mnm⁵U)、“\$”-5-羧甲基氨基甲基-2-硫代尿苷(cmm⁵s²U)、“S”-5-甲基氨基甲基-2-硫代尿苷(mnm⁵s²U)、“)”-5-羧甲基氨基甲基-2’-O-甲基尿苷(cmm⁵Um)、“B”-2’-O-甲基胞苷(Cm)、“M”-N4-乙酰基胞苷(ac4C)、“J”-2-赖胞苷(k2C)、“I”-肌苷、“Q”-癸苷(queuosine)和“Q*”-谷酰基癸苷⁵⁵。

[0029] 图5.纯化的Ile、Glu和Asn的天然tRNA的分析。(A)在SYBR绿染色的变性PAGE上分析纯化的tRNA。(B-D)肽表达实验分析的纯化的天然tRNA^{Glu}(B)、tRNA^{Asn}(C)和tRNA^{Ile}(GAU)(D)的活性。试验中天然tRNA的终浓度为1.6μM。

[0030] 图6.sGFP在半合成体外翻译系统中的表达。(A)在tRNA耗尽的裂解物中表达具有不同密码子组成的三个sGFP ORF的DNA模板,该裂解物具有不同浓度的半合成tRNA套件或天然tRNA混合物。(B)sGFP_T2模板在含有半合成tRNA混合物而缺乏所示t7tRNA的tRNA耗尽裂解物中的表达。相应密码子示于tRNA之下,其中y表示U或C,r表示A或G。

[0031] 图7.sGFP_T1在PURE体外翻译系统中以及在tRNA耗尽的S30裂解物中的表达。体外翻译实验在无tRNA(圆圈)、有天然tRNA(方形)或有半合成tRNA(三角形)的情况下进行。

[0032] 图8.用于体外合成48个E.coli tRNA的DNA模板的构建。(A)以G、U或C起始的tRNA的DNA模板;(B)以A或U起始的tRNA的DNA模板。这些DNA模板含有HHRz编码序列,其在转录后从前体上自动剪接出来而释放tRNA。方框表示HHRz编码序列,方框中的箭头表示核酶剪切位点;(C)48个E.coli tRNA的汇总,在表中定义为a1至d12。表中显示了根据图1的tRNA名称以及它们各自的基因拷贝数、第一个碱基及尺寸。

[0033] 图9.RGS肽衍生物GGRGS和DDRGS在亲和钳试验中的校准曲线。此处,TVMV底物淬灭的初始速率被对肽的浓度作图并被拟合为线性方程。

[0034] 图10.以不同密码子开始的RGS模板的肽表达水平。密码子位于(A)紧接在AUG起始密码子的下游或(B)AUG后两个连续GAU密码子下游。前序密码子显示于各柱子之下。

[0035] 图11.天然tRNA耗尽效率和t7tRNA解码活性的计算。使用具有一个或两个测试密码子的两个编码框架。对于对应RGS肽的8个氨基酸的密码子,在肽编码框架(灰色)内测试它们的同义变体(synonymous variants)。如果是双密码子模板,除了RGS编码框架内的相同密码子变体外,还使用GAU对后的2个连续测试密码子。对于其他未在RGS肽序列中出现的

12个氨基酸,紧接在连续GAU密码子对的下游设置一个或两个测试密码子。红色的“xxx”表示测试密码子。以不同tRNA/密码子组合的体外翻译反应在tRNA耗尽的裂解物中进行。以E.coli天然tRNA (1 μ g/ μ l终浓度)进行的反应作为阳性对照(Σ 天然tRNA)。以缺少单个tRNA的不完整t7tRNA混合物进行的反应(Σ t7 (-1) tRNA混合物)被用来评价给定测试密码子的天然tRNA的耗尽效率(使用方框中显示的公式)。以完整(Σ t7tRNA混合物)和不完整t7tRNA混合物进行的反应之间的差异被用来评估测试密码子特异性t7tRNA的相对活性以及用来揭示未分割或分割密码子家族盒中所有可能的同源或非同源交叉识别事件。对于单密码子和双密码子模板,测试密码子特异性t7tRNA在无细胞翻译反应中的浓度均为1.6或分别为6.4 μ M。所有20个典型氨基酸的密码子/tRNA组合都被分析,并通过亲和钳试验来测定肽产率。除非另有说明,否则每个反应都重复一次。

[0036] 图12.t7tRNA对Ser、Arg和Leu的活性评价。A:使用单密码子模板在1.6 μ M tRNA浓度(上)或双密码子模板在1.6 μ M(中)和6.4 μ M(下) tRNA浓度时测试四种t7tRNA Ser同工受体对六种Ser密码子的解码效率。B:对大多数Ser、Arg和Leu密码子的平均t7tRNA解码效率被计算为单密码子和双密码子模板在1.6 μ M或6.4 μ M的t7tRNA时的相对活性的平均值。星号标记的密码子(CGU、CGC、CGA和CUG)是例外,这些密码子的天然同工受体不完全耗尽,因而掩盖了单密码子模板实验中的信号。因此,t7tRNA对这四种密码子的功能性用双密码子模板来测试,它们的相对解码活性被计算为在1.6和6.4 μ M的各自t7tRNA时的活性的平均值。

[0037] 图13.t7tRNA对四倍简并氨基酸(Val、Pro、Thr、Ala和Gly)解码效率的评价。Prot7tRNA同工受体的密码子阅读活性基于单密码子模板反应而计算,因为相应的双密码子模板无法翻译。以星号标记的其他密码子(天然同工受体仅部分耗尽)的评价解码活性基于仅用双密码子模板测量的活性来计算。

[0038] 图14.t7tRNA对于1重、2重和3倍简并氨基酸的密码子的分析。图中显示出对每个密码子来说表现最佳的t7tRNA同工受体(见图8的名称)以及它们相对活性。图13中计算出了t7tRNA相对活性。

[0039] 图15.分离在不同寡核苷酸基质上的天然tRNAGlu的PAGE分析。天然tRNAGlu同工受体通过分别用OligoDNA1-Glu-胺和OligoDNA1-Glu-生物素官能化的NHS-和链霉亲和素树脂纯化,使用两种不同的缓冲系统(含TMA⁺或Na⁺)¹¹。洗脱组分(E1和E2)在8%变性PAGE上分析纯度。未结合的tRNA通过用10倍树脂量的10mM Tris-HCl (pH7.5)缓冲液在室温下洗涤而除去,直至洗脱液在260nm的吸收值在0.01A260单位/ml以下。目标tRNA以3倍树脂体积的10mM Tris-HCl (pH7.5)在68 $^{\circ}$ C洗脱两次。

[0040] 图16.使用RNA适配体(吻合环(kissing loop))色谱法从总tRNA混合物中耗尽特定tRNA。(A)吻合环:tRNA复合物形成的碱基配对机制示意图,(B)特定tRNA耗尽的体外翻译系统的制备流程图,(C) tRNA重构(reconstitution)后GFP-分析的测定。

[0041] 图17.ARS反冲(recoiling)过程。

[0042] 图18.ARS反冲过程在tRNA耗尽上的结果。

[0043] 图19.A:包含接枝的GCU-或CCU-反密码子(Cys-tRNA (Cys) gcu/ccu)的半胱氨酰化或硒代半胱氨酰化tRNACys (Secys)。B:使用全部tRNA耗尽裂解物与缺乏t7tRNA (Arg) ccu的半合成tRNA套件重构翻译反应物,并以携带单个AGG密码子的编码eGFP的模板引发该反应物,Cys-tRNA (Cys) ccu能够有效抑制单个AGG密码子。C:在使用KL直接耗尽了内源性tRNA

(Ser)GCU的裂解物的翻译反应中有效抑制两个连续的AGC密码子。D:通过含有接枝的CUA反密码子的翻译前硒代半胱氨酰化的tRNACys有效抑制编码eGFP的ORF中的琥珀密码子(第153位)(反应条件如B)。E:用BodipyFL碘乙酰胺纯化Secys-tRNACys(CCU-反密码子)偶联产物。

[0044] 图20.通过并入非天然氨基酸来设计大环肽。(A)使用BODIPY-tRNA-Cys在缺乏tRNAArg(CCU)的重构无细胞系统中重新分配EGFP第108位的AGG密码子。(B)重新分配EGFP第4位的AGC密码子。左图是加有从大肠杆菌体外翻译反应纯化的EGFP的SDS-PAGE扫描。泳道1代表未修饰的无细胞系统,而泳道2和3代表其中_{gcut}tRNA被负载有环辛烯赖氨酸(COC)的Pyr_{gcut}tRNA替换的反应。在纯化之前COC用Atto488-四嗪标记。(C)具有一个或两个重新分配的Ser(AGC)或琥珀(TAG)密码子的EGFP的表达产量。(D)四嗪连接反应。(E)能够连接两个含环辛烯的氨基酸的双官能四嗪衍生物的结构。(F)三官能四嗪衍生物。(G)SDS-PAGE分析体外翻译混合物,该混合物表达含有与双四嗪(DT)交联的含COC的EGFP蛋白(B)。凝胶用未煮沸的样品上样并扫描EGFP荧光。(H)使用密码子重新分配和铜催化的点击反应形成大环多肽的实例。

[0045] 图21.在加入所示3种tRNA混合物的翻译反应中,全部六个(6)或仅一个精氨酸密码子变为AGG(图中分别为6AGG或1AGG)的两种GFP编码模板的翻译效率。总tRNA和Dep1tRNA分别表示tRNA_{accu/ucu}耗尽前后总自制分离的tRNA混合物。第三种混合物含有合成的t7tRNA_{accu},其终浓度为5μM。

[0046] 图22.具有不同nnAA/o-tRNA/aaRS组合的大肠杆菌无细胞翻译系统中琥珀密码子抑制效率的比较分析。A:此处所使用的nnAA的结构。B-D:A_{151X}模板的琥珀抑制。抑制效率被定义为与ORF内不含终止密码子的野生型eGFP荧光相比的比率。B:通过三种Py1RS变体并入四种nnAA。Py1T_{cua}、Py1RS变体和nnAA的最终浓度分别为20μM、20μM和1mM。C:通过四种MjY_{tRNA}变体进行AzF并入。MjY1、MjY2、MjY3和MjY4对应于原始文献^{35,37}中的tRNA_{pAzPhe1}、_{pAzPhe2}、_{pAzPhe3}和tRNA_{opt CUA}。MjY_{tRNA}变体、AzFRS和AzF的终浓度分别为10μM、10μM和1mM。D:o-tRNA和o-aaRS浓度对MjY2/AzFRS/AzF和Py1T/Py1RSAF/PrK的琥珀密码子抑制效率的影响。E:密码子背景对两个正交系统中nnAA并入效率的影响。MjY2和AzFRS的浓度均为10μM,而Py1T和Py1RS的浓度分别为20μM和30μM。在分析中使用了四种模板,即A_{1X}、A_{151X}、B_{1X}和B_{151X}(序列显示在SI中)。A和B表示具有不同密码子偏好的模板骨架,其中A中的密码子表针对eGFP进行了优化,而在B中主要是利用独特的密码子来解码每种氨基酸类型而实现简化。数字表示ORF序列中琥珀密码子的位置。单个反应中的相对活性以使用EctRNA_{Tyr(cua)}作为抑制剂时获得的荧光强度的百分比表示。

[0047] 图23.AGG在四种GFP ORF中重新分配至AzF。GFP蛋白在耗尽了解码AGG密码子的tRNA的无细胞系统中的表达,分添加和未加MjY抑制系统(MjY2/AzFRS)两种情况。MjY2、AzFRS、AzF的最终浓度分别为10μM、10μM和1mM。将MjY2中的反密码子改为CCU以将AGG密码子重新分配给AzF。

[0048] 图24.用于AGG和UAG抑制的BPFL-tRNA的合成。(A):结合BPFL的tRNA的HPLC纯化。带有CUA或CCU反密码子的tRNACys通过CysRS被负载有Cys,并与BP-FL上的碘乙酰胺基团反应。在254nm处的吸光度测量检测到核糖核酸,而490nm通道用于BP-FL检测。(B-C):通过荧光扫描分析标记的蛋白质产量,通过12%PAGE凝胶的Western印迹分析总蛋白质产量。通过

煮沸样品淬灭GFP荧光,随后在凝胶上检测BP-FL并入。用抗GFP抗体通过Western印迹检测总蛋白质。每个条带的强度通过ImagJ按照设置为1的最微弱条带的强度来计算。(B) 带有一个AGG密码子的sGFP_T2模板在正常裂解物或补充有所示tRNA混合物的tRNA耗尽裂解物中表达,分添加或不加负载有Biodipy的tRNA_{Accu}两种情况。(C) 在用tRNA耗尽裂解物和缺乏AGG抑制剂的总tRNA混合物重构的翻译反应物中,加入分别在GFP ORF第1位和第151位具有单个AGG的模板,测试AGG抑制效果,分添加或未添加BPFL-tRNA_{Accu}抑制剂两种情况。对于琥珀抑制也进行类似的反应,但使用总tRNA混合物、含单个琥珀密码子的模板和BPFL-tRNA_{Acua}抑制剂。

[0049] 图25. CaM位点特异性双重标记用于基于smFRET的构象变化分析。(A) 单标记和双重标记的CaM的荧光扫描。双重标记的CaM蛋白质是在用tRNA耗尽裂解物和缺乏AGG抑制剂的总tRNA混合物重构的无细胞翻译系统中制备的。两个抑制系统, MjY2 (Cua) /AzFRS/AzF和BPFL-cys-tRNA_{Accu}, 分别添加用于重新编码在第1位和第149位的UAG和AGG密码子。然后通过无铜点击化学使并入到蛋白质中的AzF与DIBO-TAMRA反应。单标记的CaM蛋白以相同的方式制备,但仅使用一种抑制系统,而另一个o-密码子通过tRNA_{Tyr_{cua}}或tRNA_{Arg_{ccu}}解码。使用所示的两个通道来扫描用单标记和双标记的CaM上样的SDS-PAGE凝胶。(B) 在488nm激发的单标记和双重标记CaM的荧光发射光谱。调整TAMRA标记的蛋白质的浓度以确保其在543nm激发时其发射荧光与双重标记蛋白质的发射荧光相同。(C) :无Ca²⁺形式 (PDB:1CFC) 和Ca²⁺结合形式 (PDB:4CLN) 的CaM结构示意图。绿色和红色的圆点分别表示安装在CaM ORF第1位和第149位的BP-FL和TAMRA。(D-E) smFRET直方图记录在不同条件下的双重标记CaM。(D) :在50mM Tris-HCl、150mM NaCl缓冲液中,不含Ca²⁺, (E) :2mM Ca²⁺, (F) :在10mM EDTA中。实线代表使用Origin软件对数据进行高斯拟合。每个峰表示一个单独的构象。

[0050] 图26. pOPINE_CaM模板的载体图谱。

[0051] 图27. 在补充或不补充t7tRNA_{Accu}的情况下耗尽的商业tRNA混合物的翻译表现。在用3种tRNA混合物编程的翻译反应中测试了两种GFP-编码模板,所述模板具有全部或仅一个AGG密码子,即图中所示的6AGG或1AGG。总tRNA和Depl tRNA表示tRNA_{Accu/ucu}耗尽前后的商业tRNA混合物。第三种混合物含有t7tRNA_{Accu}以使终浓度为5μM。

[0052] 图28. 四种先前报道的MjYtRNA的比对。MjY1、MjY2、MjY3和MjY4对应于原始文献³中的pAzPhe1、pAzPhe2、pAzPhe33tRNA和原始文献⁴中的tRNA_{opt CUA}。与原文报道的序列不同, MjY2_tRNA中的N63到N67是“CATCG”而不是“CATCGT”(最后的“T”在原始文件中似乎是错误的)。

[0053] 图29. 使用flexizyme预负载有各自活性基团的5nAA的并入效率。A: 通过Flexizyme进行tRNA酰化的5个活化酸底物的结构,包括L-炔丙基甘氨酸4-氯苄基硫酯 (Pra-CBT)、L-叠氮基赖氨酸4-氯苄基硫酯 (Lys (N3)-CBT)、L-叠氮基高丙氨酸4-氯苄基硫酯 (Aha-CBT)、N^c-生物素基-L-赖氨酸3,5-二硝基苄基酯 (Lys (生物素)-DBN) 和L-叠氮基苯丙氨酸-氰甲基酯 (AzF-CME)。nnAA以黑色显示,而其各自的活性基团显示为红色。B: 使用A_151X模板在无细胞翻译系统中测试的5种nAA预负载的tRNA的琥珀抑制效率。预负载的tRNA用于在10、20或30μM的最终tRNA浓度下引发翻译反应。图中显示了flexizyme反应时间和预负载的nnAA-tRNA的最终浓度。在此测试的五种nnAA中, AzF显示出最高的翻译活性。这表明,与其他nnAA相比, flexizyme可以很好地在o-tRNA上负载AzF,并且这种预负载的AzF-

tRNA被EF-Tu和核糖体良好地接受。

[0054] 图30.两种抑制系统MjY2/AzFRS/AzF和Py1T/Py1RSAF/PrK在大肠杆菌体外翻译系统中的正交性。使用缺乏nnAA-底物、o-tRNA或酶的反应来评估对模板A_151X中琥珀密码子的非特异性抑制水平。MjY2和AzFRS的浓度均为10 μ M,而Py1T和Py1RS分别为20 μ M和40 μ M。

[0055] 图31. (A) 模板A_151X中的TAG密码子被重新分配给AzF。图中显示了含有nnAA并入位点的鉴定的肽的MS/MS分析。(B). 模板A_151R中的AGG密码子被重新分配给AzF。图中显示了含有nnAA并入位点的鉴定的肽的MS/MS分析。(C) 模板A_1R中的AGG密码子被重新分配给AzF。图中显示了含有nnAA并入位点的鉴定的肽的MS/MS分析。

[0056] 图32. BPFL-tRNACys (CCU) 的结构。

[0057] 图33. tRNA、氨基酸和延伸因子的三级复合物。负载有不同氨基酸的tRNA^{Trp}的结构与延伸因子EF-Tu形成三级复合物。A: Trp-tRNA-EFTu, B: BPFL-cys-tRNA-EFTu。

[0058] 图34. tRNA^{Cysccu}支持AGG翻译的能力。无细胞翻译反应是在补充缺乏AGG同工受体的半合成tRNA混合物的tRNA耗尽裂解物的情况下进行的。带有一个AGG密码子的sGFP_T2模板在添加或不添加含CCU反密码子的tRNA的情况下表达。tRNA^{Argccu}是合成的wt AGG tRNA同工受体,其用作阳性对照。

[0059] 图35. 商业tRNA混合物在0.25和0.5 μ g/ml时的翻译表现,该商业tRNA混合物耗尽了天然AGC-和AGG-密码子tRNA抑制剂(Σ -a5/6, -c11),反应物中补充或未补充其t7tRNA对应物:t7tRNA^{Accu} (a05) 和t7tRNA^{Gcu} (c11)。在抑制试验中使用以下模板:只在第4位具有一个独特的 (eGFP x1AGC) AGC密码子或在第4,5位具有两个连续的 (x2AGC) AGC密码子的两种GFP-编码模板(都不含任何AGG),以及具有独特AGC密码子(第4位)和AGG(第151位)密码子 (eGFP x1AGC/x1AGG) 的一个模板。对于每个模板,翻译反应用缺乏所需的AGC-和AGG-密码子的天然同工受体的总tRNA混合物编程,或用补充了两者的t7tRNA类似物的相同混合物进行编程。所使用的ORF的示意图显示在每个图下方。

[0060] 图36. 用于双重密码子标记的实验工作流程。CaM天然ORF以黑色显示,precision蛋白酶切割位点以蓝色圆圈表示,而亲和钳结合肽标签显示为棕色六角形。X1和X2分别表示在CaM的第1位或第2和第3位的单个或两个连续AGC密码子。小写字母表明其余Ser和Arg密码子的偏好。

[0061] 图37. CaM的位点特异性双标记。单标记和双标记的CaM的荧光扫描。在无细胞翻译系统中制备双重标记的CaM蛋白质,所述系统由tRNA耗尽裂解物和缺乏AGG/AGA密码子和AGC密码子的抑制剂的总tRNA混合物重构。向反应物中补充两个抑制系统MjY2 (gcu) / AzFRS/AzF和BPFL-cys-tRNA^{Accu}分别用于解码第1位和第151位的AGC和AGG密码子。携带AzF的蛋白质直接上样在SDS-PAGE上,或者通过无铜菌株促进的环加成与DIBO-TAMRA偶联。使用了两个编码CaM的ORF,它们除了位于151位的独特的AGG之外,在第1位或第1,2位具有1个 (x1AGC) 或两个 (x2) 连续的AGC密码子。使用两个通道在两个指定的波长下扫描SDS-PAGE凝胶。

[0062] 详细描述

[0063] 本发明提供一种蛋白质翻译系统,其允许将非天然部分(如非天然氨基酸)并入到所翻译的蛋白质中而不损害将所有二十种(20)天然氨基酸并入到该蛋白质中的能力。这是通过以下方法实现的:将通常由至少两个(2)不同tRNA反密码子解码的氨基酸的其中一个

tRNA反密码子重新分配至非天然部分,其中至少一个密码子可被该重新分配的反密码子独特地识别,并且至少一个另一密码子(即来自相同密码子框)不能被该重新分配的反密码子识别。因此,所得到的用于翻译的mRNA包含一个或多个对应于重新分配的tRNA反密码子的特定密码子,因而该非天然部分在适当的或期望的位置被并入到所翻译的蛋白质中。

[0064] 在一个方面,本发明提供一种生产适于包含至少一个非天然部分的蛋白质的翻译的一套tRNA,所述方法包括将至少一个包含天然按实际上的反密码子的tRNA替换为至少一个包含被重新分配至非天然部分的相同反密码子的tRNA,其中该套tRNA能够被操作从而便于RNA的翻译,该RNA包含对应于已被重新分配至该非天然部分的反密码子的密码子,从而使所翻译的蛋白质可包含任何一种或所有二十种(20)天然氨基酸。

[0065] 在另一个方面,本发明提供一种组合物,其包含适于包含至少一个非天然部分的蛋白质的翻译的一套tRNA,所述一套tRNA包含至少一个含有非天然部分的反密码子的tRNA,该反密码子是从天然氨基酸的反密码子重新分配的,其中该一套tRNA能够被操作而便于RNA的翻译,该RNA包含对应于已被重新分配至该非天然部分的反密码子的密码子,从而使所翻译的蛋白质可包含任何一种或所有二十种(20)天然氨基酸。

[0066] 如本文通常所使用的,蛋白质是一种包含两个(2)或更多个共价链接部分的聚合物,这些部分可为天然L-氨基酸和/或非天然部分(如非天然氨基酸)。通常,肽是一种包含不多于五十个(50)连续部分的蛋白质。通常,多肽是包含多于五十个(50)连续部分的蛋白质。

[0067] 如本文所使用的,“天然氨基酸”是可由生物体的基因组遗传编码的L-氨基酸。通常,天然氨基酸在肽基转移酶中心以生理范围的动力学常数反应。优选地,所述天然氨基酸选自:丙氨酸、天冬酰胺、天冬氨酸、半胱氨酸、甘氨酸、赖氨酸、谷氨酸、谷氨酰胺、精氨酸、组氨酸、蛋氨酸、丝氨酸、苏氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、酪氨酸、色氨酸、苯丙氨酸或它们的衍生物,这些衍生物通常可通过具有任何这些天然氨基酸的反密码子的tRNA而被并入到所翻译的蛋白质中。衍生物可为天然存在的衍生物,例如翻译后修饰的氨基酸(如硒代半胱氨酸)。

[0068] 如本文所使用的,“非天然部分”可为能够被并入到由RNA模板经核酶介导的链延伸而翻译得到的蛋白质中的任何分子,条件是它不是前文所定义的天然氨基酸。一般而言,这种非天然部分可包括非天然氨基酸、天然氨基酸的天然的或合成的化学衍生物、和/或具有化学反应性的部分(如能够形成分子内共价键的部分)。在一些实施方式中,受上述条件约束,上述非天然氨基酸可为能够形成肽键的具有氨基($-NH_2$)和羧基($-COOH$)的任何有机化合物。非天然部分的非限制性例子包括D-氨基酸、硒代半胱氨酸、吡咯赖氨酸、N-甲酰蛋氨酸、 α -氨基-n-丁酸、戊氨酸、正亮氨酸、别异亮氨酸、t-亮氨酸、 α -氨基-n-庚酸、哌啶酸、 α, β -二氨基丙酸、 α, γ -二氨基丁酸、鸟氨酸、别苏氨酸、同型半胱氨酸、同丝氨酸、 β -丙氨酸、 β -氨基-n-丁酸、 β -氨基异丁酸、 γ -氨基丁酸、 α -氨基异丁酸、异缬氨酸、肌氨酸、N-乙基甘氨酸、N-丙基甘氨酸、N-异丙基甘氨酸、N-甲基丙氨酸、N-乙基丙氨酸、N-甲基 β -丙氨酸, N-乙基 β -丙氨酸。

[0069] 可能适合形成分子内共价键的化学部分可为任何有利于以下作用的物质:加入同质缩合(homo-condensation)的两个相同反应基;加入经双官能反应基而缩合的两个相同反应基;或加入异质缩合(hetero-condensation)的两个不同反应基。非限制性的例子包括

由例如NHS-酯、马来酰亚胺、卤代乙酰基的部分修饰的氨基酸侧链或末端氨基或羧基,但不限于此。也可参考Current Protocols in Protein Science Eds.Coligan et al (2001-2012) 第14、15章有关氨基酸和蛋白质的化学修饰的更详细讨论。

[0070] 如本文通常所使用的,“tRNA”是包括天然的和合成的tRNA分子在内的转移RNA或同工受体RNA分子。通常,tRNA分子包含约60-93个核苷酸的核苷酸序列,具有内部碱基对区域,形成二级结构中的四或五个双链茎以及三个或四个单链环。5'末端和3'末端位于内部碱基配对的“接受茎”的末端。3'末端包含核苷酸序列CCA,末端“A”核苷酸是氨基酸特异性氨酰tRNA合成酶的氨酰化位点。氨基酸特异性反密码子位于环II内。

[0071] 如本文所使用的,“反密码子”是tRNA分子中对应于mRNA或其DNA前体中编码天然氨基酸或用作翻译终止子(UAA、UGA、UAG)的密码子的核苷酸序列。在许多情况下,反密码子便于通过合适的氨酰tRNA合成酶使用该反密码子所对应的可翻译mRNA密码子编码的正确氨基酸(Ser、Ala以及某些情况下Leu除外)对tRNA进行氨酰化。此时,反密码子可为至少在对应于相应5'-3' mRNA密码子序列的第一和第二位置处形成Watson-Crick碱基对的3'-5'三核苷酸序列。因为遗传密码是简并的,所以对于给定的氨基酸可能具有不止一种解码密码子框的tRNA。这些解码相同氨基酸并且包含不同反密码子的tRNA被称为“同工受体”。对于所有密码子(编码Met和Trp的密码子除外),反密码子和mRNA密码子之间的碱基配对特异性由头两个核苷酸(即反密码子3'-5'方向、mRNA密码子5'-3'方向)确定,因而相同的反密码子可能对应于两个或更多个简并的mRNA密码子。Asp、Asn、Cys、Glu、Gln、His、Lys、Phe和Tyr的每个单个tRNA同工受体对应两个不同的mRNA密码子(即“二倍简并”)。Ile具有两个同工受体,其中一个对应单个mRNA密码子(即“三倍简并”)。Ala、Gly、Pro、Thr和Val具有两个对应于RNA模板中的各个密码子的同工受体(即“四倍简并”)。对于Arg、Leu和Ser,有四个或五个对应各个密码子的同工受体(即“六倍简并”)。

[0072] 适当地,本发明公开的组合物包括至少一个含有由天然氨基酸重新分配至非天然部分的反密码子的tRNA。在本文中,“重新分配”是指该反密码子不再对应于编码其正交的(orthogonal)天然氨基酸的RNA或DNA密码子。如此,使用该组合物的任何可翻译RNA也会在预期将该非天然部分并入到所合成的蛋白质的位置包含相应的密码子。通过利用对应于冗余密码子的反密码子,该组合物使得能够并入非天然部分而不损害将所有二十种(20)天然氨基酸包括在所翻译的蛋白质中的能力。

[0073] 原则上,本发明的重新分配的反密码子可为对应于编码天然氨基酸的冗余RNA密码子的任何反密码子,其中所述重新分配的密码子是相同氨基酸的多个不同密码子中的一个。适当地,在给定氨基酸的冗余密码子中,至少一个RNA密码子独特地对应于该重新分配的反密码子,并且至少一个RNA密码子不对应于该重新分配的反密码子。如本领域技术人员将会理解的,遗传密码子具有简并性,因而除色氨酸和蛋氨酸之外,所有天然氨基酸都由不止一个密码子编码。本发明一个优选的目的在于提供一种蛋白质翻译系统,在该系统中,尽管向该蛋白质并入了非天然部分,但仍保留了并入所有二十种(20)非天然氨基酸的能力。因此,优选地,该重新分配的反密码子是如前所述的四倍简并的或六倍简并的。这使得“非重新分配的”反密码子在蛋白质翻译期间发挥其并入天然氨基酸的正常作用。在特别优选的形式中,所述反密码子是Leu、Arg或Ser的反密码子。

[0074] 如可从实施例中意识到的,在任何特定氨基酸的可重新分配反密码子的每个库

(pool) 或套件 (complement) 中, 可能存在用以筛选重新分配的合适反密码子的规则或条件。这些规则或条件可包括碱基配对的“摆动”效应、tRNA 反密码子和 mRNA 密码子之间可能的交叉识别、和/或制备包含该重新分配的反密码子的合成 tRNA 分子的适宜性。有关这些因素的更详细的考察在实施例中进行。

[0075] 在一个实施方式中, 生产所述组合物的方法包括以下步骤: (i) 从适合于包含天然氨基酸的蛋白质的翻译的一套 tRNA 中耗尽一种或多种 tRNA; 和 (ii) 以一个或多个分别被重新分配至非天然部分并且分别被连接至所述非天然部分的 tRNA 来重构该被耗尽的一套 tRNA。

[0076] 在一些实施方式中, (i) 中的所述一套 tRNA 存在于一个无细胞翻译系统中。

[0077] 在步骤 (i) 的一个实施方式中, 基本上所有天然氨基酸的 tRNA 都被从该套 tRNA 中耗尽。在一个特别的实施方式中, 耗尽是通过将 tRNA 结合至乙醇胺琼脂糖 (ethanolamine sepharose) 实现的。

[0078] 在步骤 (i) 的另一个实施方式中, 一种或多种天然氨基酸的 tRNA 被选择性地从该套 tRNA 中耗尽。在一个实施方式中, 选择性耗尽是通过将一种或多种 tRNA 选择性地结合至各自的特异性 tRNA 耗尽试剂实现的。这种试剂可为天然或合成的蛋白质、DNA、PNA 等。

[0079] 在一个实施方式中, tRNA 耗尽试剂是 RNA 适配体。适当地, 每个 RNA 适配体与特定的 tRNA 形成特异的、高亲和性复合物。在一个实施方式中, 所述 RNA 适配体包含核苷酸序列, 其通过结合至靶 tRNA 的含反密码子核苷酸序列而形成高亲和性复合物。在一些实施方式中, 所述 RNA 适配体可被称为“吻合环”RNA。

[0080] 在另一个实施方式中, 所述 tRNA 耗尽试剂包含一种或多种单链 DNA 寡核苷酸, 这些寡核苷酸序列具有对各自 tRNA 特异的核苷酸序列。该实施方式特别适用于耗尽特定 tRNA。

[0081] 在步骤 (ii) 的一个实施方式中, 重构 tRNA 可包含合成 tRNA、天然 tRNA 或它们的混合物 (本文称“半合成 tRNA 套件”)。合成 tRNA 可通过本领域已知的任何化学法或酶法制成, 包括 RNA 聚合酶介导的合成法。非限制性的例子包括 SP6、SP3 和 T7 介导的合成, 但不限于此。如下文将更详细描述, 天冬酰胺、谷氨酸和异亮氨酸的合成 tRNA 实质上是无功能的。因此, 为重构的目的, 适当地, Asn、Glu 和 Ile 的 tRNA 是天然 tRNA, 即使可替换为工程化的 tRNA 突变体。

[0082] 在一些实施方式中, 特异性 tRNA 可在选择性地结合至 RNA 适配体 (如前所述的“吻合环”RNA) 之后获得并用于重构。

[0083] 非天然部分可通过本领域任何已知的方法而被结合、负载或装载至重构 tRNA。这些方法可包括使用化学氨酰化或酶氨酰化。酶氨酰化的非限制性例子包括使用天然的、修饰的氨酰 tRNA 合成酶 (例如 Py1RS 或其在吡咯赖氨酸 tRNA 合成酶介导的氨酰化中使用的变体、詹氏甲烷球菌酪氨酰转移 RNA 合成酶 (Mj TyrRS) 或其变体)、Flexizyme 介导的氨酰化和/或通过半胱氨酰 tRNA 合成酶进行的氨酰化。

[0084] 在通过半胱氨酰 tRNA 合成酶以非天然部分使 tRNA 氨酰化的实施方式中, 优选的实施方式实现了具有被重新分配至非天然部分的反密码子的合成 tRNA 的体外负载。野生型半胱氨酸 tRNA 反密码子的突变降低了半胱氨酰 tRNA 合成酶对半胱氨酸 tRNA 的亲合力, 但不会实质性降低或抑制半胱氨酰 tRNA 合成酶在体外的氨酰化活性, 在体外, 高浓度的反应物补偿了亲和力的降低。半胱氨酸 tRNA 反密码子突变体在体内不能被半胱氨酸氨酰化 (如在无

细胞翻译系统中),因为半胱氨酰tRNA合成酶通常水平较低并且存在竞争性内源性半胱氨酸tRNA。因此,可通过在体外负载具有被重新分配至非天然部分的反密码子的合成tRNA并用该合成tRNA或每个包含不同非天然部分的多个不同tRNA来重构tRNA套件,从而生产一套tRNA。如前所述,本发明公开的组合物可适用于重组蛋白质生产方法或系统。

[0085] 因此,本发明的一个方面提供一种生产翻译系统的方法,该翻译系统适合于包含至少一个非天然部分的蛋白质的翻译,所述方法包括:生产一套tRNA,该一套tRNA包含至少一个具有已被重新分配至非天然部分的天然氨基酸反密码子的tRNA;以及生产可转录(transcribable) RNA,该可转录RNA包含对应于该已被重新分配至所述非天然部分的反密码子的密码子,其中该mRNA可被转录而生产可包含任何一种或所有二十种(20)天然氨基酸的已翻译蛋白质。

[0086] 本发明的另一个方面提供一种翻译系统,其适合于包含至少一个非天然部分的蛋白质的翻译,所述系统包含:一套tRNA,包含至少一个具有已被重新分配至非天然部分的天然氨基酸反密码子的tRNA;和可翻译mRNA,其包含对应于该已被重新分配至非天然部分的反密码子的密码子,其中所述mRNA可被转录以生产可包含任何一种或所有二十种(20)天然氨基酸的已翻译蛋白质。

[0087] 本发明的另一个方面提供一种生产包含至少一个非天然部分的重组蛋白质的方法,所述方法包括翻译包含密码子的mRNA,该密码子对应于已被重新分配至非天然部分的、在一套tRNA中的一个tRNA的反密码子,所述一套tRNA包含至少一个具有已被重新分配至非天然部分的天然氨基酸反密码子的tRNA,其中所翻译的蛋白质可包含任何一种或所有二十种(20)天然氨基酸。

[0088] 适当地,这些方面的系统和方法用于进行体外蛋白质翻译或用于其他无细胞翻译系统。非限制性的例子包括麦胚、昆虫、HeLa裂解物、兔网织红细胞裂解物、基于E.coli和利什曼虫的系统。

[0089] 如可从前文理解到的,本发明至少部分地基于重新分配遗传密码的一个或多个冗余密码子以编码非天然部分。因此,本发明提供一种可用来生产出包含一个或多个密码子的可翻译RNA的方法,该一个或多个密码子选择性地对应于已被从天然氨基酸重新分配至非天然部分的各个tRNA的反密码子。

[0090] 本发明的一个特别的优点在于该可翻译mRNA除包含所述一个或多个非天然部分的密码子之外还可包含任何一种或所有二十种(20)天然氨基酸的密码子,因而所翻译出的蛋白质除包含所述一个或多个天然部分之外还可包含任何一种或所有二十种(20)天然氨基酸。

[0091] 因此,本发明的一个方面提供一种由所公开的方法生产的重组蛋白质。

[0092] 本发明还提供一种编码所公开的重组蛋白质的mRNA分子。

[0093] 将意识到的是,分离的蛋白质可包含一个或多个相同或不同的非天然部分,而有利于进一步的衍生化,例如PEG化、结合小分子、标记、标签化、固定、分子间和/或分子内交联或其他相互作用、形成更高级别的结构和/或一种或多种催化活性,但不限于此。

[0094] 在一个具体实施方式中,所翻译的蛋白质可包含一个或多个非天然部分,而有利于形成一个或多个分子内共价键。在一个具体实施方式中,所述重组蛋白质是大环状(macrocyclic)蛋白质。

[0095] 作为例子,环状多肽可用于生物信息学方法中用来设计集中肽库(focused peptide libraries),其包含来自某些或所有天然结构折叠类别的代表。

[0096] 在包含可环化(cyclizable)氨基酸的蛋白质的生产中提供多种重新分配密码子可有利于构建大环状肽的文库。为形成分子内共价键,可使用以下方法:a)加入同质缩合(homo-condensation)的两个相同反应基;b)加入经双官能反应基而缩合的两个相同反应基;和c)加入异质缩合(hetero-condensation)的两个不同反应基。如本文所述,tRNA反密码子的重新分配提供了一种测试这些方法的系统。实施例中将更详细地描述构建编码合成的10-聚体肽的测试mRNA的非限制性例子,该测试mRNA携带Ser和Arg密码子和C末端亲和铈标签。

[0097] 将意识到的是,本文公开的发明可用于生产在靶位点或选定位点并入了天然氨基酸以外的部分的任何蛋白质或肽、或包含多种不同这样的蛋白质和肽的文库,同时允许按需并入天然氨基酸。

[0098] 为使本发明能够被完全理解并付诸实际应用,请参考以下非限制性的实施例。

实施例

[0099] 实施例1

[0100] 半合成蛋白质翻译系统的开发

[0101] 材料和方法

[0102] 肽传感器(亲和铈)试验

[0103] 不同序列的报道肽RGSIDTWV、GGRGSIDTWV、DDRGSIDTWV以及荧光淬灭的TVMV底物肽(5-氨基-2-硝基苯甲酸-ETVRFQSK-7-甲氧基香豆素-4-基)由Mimotopes合成。自身抑制蛋白酶与亲和铈(肽传感器)的融合蛋白通过Ni²⁺-NTA亲和色谱纯化并存储于50mM Tris-HCl,1M NaCl,5mM EDTA,2mM TCEP,10%甘油缓冲液中(pH 8.0)。

[0104] 通常,亲和铈试验在缓冲液A中进行,该缓冲液A组成为50mM Tris-HCl,1M NaCl,1mM DTT,0.5mM EDTA(pH8.0),并补充有1.3μM肽传感器和300μM TVMV底物肽。RGS肽在缓冲液中被用作溶液或在体外翻译中使用,并且反应进程使用Synergy酶标仪通过在330nm激发样品并在405nm读取荧光变化1h来监控。生成校准曲线以建立底物酶切的初始速率(V_{max})与对照肽的已知浓度之间的关系。实验重复三次。

[0105] 为定量无细胞翻译反应中RGS肽及其衍生物,向配制用于偶联转录-翻译并补充了x2蛋白酶抑制剂混合物(Roche)的S30大肠杆菌提取液中加入期望的肽编码DNA模板并在32℃孵育1h。翻译后,向反应混合物加入NaCl至终浓度为1M,再在65℃孵育10min以使与TVMV蛋白酶竞争底物肽的内源性蛋白酶失活。10000rpm离心5min,取10μl上清液用于上述亲和铈试验。在体外翻译反应中,通过将不同量的合成肽加入缺少DNA模板的无细胞翻译反应体系中获得校准曲线。

[0106] tRNA耗尽的裂解物的制备

[0107] 如文献27中所述从BL21(DE3)GOLD制备s30大肠杆菌提取物,在tRNA耗尽程序之前于-800冻存于10mM Tris-乙酸(pH8.2),14mM Mg(OAc)₂,0.6mM KOAc,0.5mM DTT缓冲液中。对于tRNA耗尽,将2.5ml s30提取物重新缓冲于用缓冲液B平衡的NAP-25柱上(GEhealthcare),缓冲液B组成为25mM KCl,10mM NaCl,1.1mM Mg(OAc)₂,0.1mM EDTA,10mM

Hepes-KOH (pH7.5), 1mM Mg (OAc) 2, 120mM KOAc。重新缓冲后,将裂解物与根据之前文献28制备的0.8-1.2ml的沉积(settled)乙醇胺-琼脂糖基质于40C在定轨摇床中一起孵育30min。孵育之后,收集上清,基质用1ml含180mM KOAc的缓冲液B洗涤。流过液与之前步骤的上清混合以产生tRNA耗尽的裂解物,速冻并存储于-800。

[0108] 在S30裂解物中的无细胞翻译反应根据标准程序27使用30nM DNA模板和10mM终浓度的Mg (OAc) 2在320进行。PURExpress® Δ (aa, tRNA) Kit (E6840S) 购自NEB并根据制造商说明书使用。

[0109] 构建用于肽和sGFP表达的tDNA和质粒

[0110] tRNA的编码序列获得自Genomic tRNA数据库 (GtRNAdb) 29。DNA模板 (tDNAs) 通过3-步法PCR合成(图8)。所有编码肽和sGFP的DNA模板都基于pOPINE-eGFP质粒 (GenBank: EF372397.1) 构建。为构建编码肽的DNA模板,两个具有NcoI和NotI限制位点突出端的互补寡核苷酸被用来组建这些期望的肽的ORF。寡核苷酸的浓度调节至100μM,在水中以1:1摩尔比混合,在95℃加热5min,随后缓慢冷却至室温以退火。pOPINE-eGFP质粒载体以NcoI和NotI消化,与退火的寡核苷酸混合,再用T4DNA连接酶连接。通过Sanger测序确认阳性克隆 (AGRF Brisbane)。

[0111] 具有各种密码子偏好性的编码sGFP ORF的片段 (sGFP_T1、sGFP_T2和sGFP_T3) 由IDT合成为G-block,并根据标准Gibson克隆程序克隆进基于pOPINE的质粒中。

[0112] T7tRNA合成及纯化

[0113] 在含有0.25μM DNA模板、10μg/ml T7聚合酶和0.25U/ml酵母无机焦磷酸酶的40mM Hepes-KOH (pH7.9)、18mM Mg (OAc) 2、2mM亚精胺、40mM DTT、5mM各rNTP中预先在320进行标准的run-off t7转录反应2h。合成tRNA^{His}时,转录反应液首先加入5mM各rATP/rCTP/rUTP和6.8mM rGMP反应5min,随后加入1.7mM rGTP并孵育2h。tRNA^{His}的DNA模板含有对应于tRNA的-1位置的额外的G。转录后,反应液被以5倍稀释于缓冲液C中 (125mM NaOAc pH 5.2, 0.25mM EDTA)。通过亲和色谱以乙醇胺-琼脂糖基质纯化tRNA转录本。对于1ml转录反应液,使用0.2ml的沉积基质。在40孵育匀浆1h后,结合了tRNA的基质用含200mM NaOAc的缓冲液C充分洗涤。t7tRNA被从基质洗脱到含2MNaOAc的缓冲液C中。用乙醇使tRNA析出,沉淀溶解于含1mM MgCl₂和0.5mM NaOAc (pH 5.0)的tRNA缓冲液中。

[0114] 通过半合成tRNA混合物来表达sGFP

[0115] 从商业上合成用于编码sGFP的具有不同同义密码子组成的三者ORF,并克隆进pOPINE质粒。模板1 (T1) 具有最大的密码子变化,包括编码Leu的五种不同的同义密码子、编码Val的四种同义密码子和编码Pro、Arg和Thr的3个三种同义密码子(表4)。模板2 (T2) 具有最高的密码子偏好性,Ser、Arg和Leu仅有两个同义密码子,Val、Pro、Thr、Ala和Gly仅有一个密码子(表5)。模板3 (T3) 具有中等的密码子变化性,Ser和Arg有两个密码子,Leu、Val、Pro、Thr、Ala和Gly像T1一样有几个密码子(表6)。除了出现10次以上的密码子以及对应于最少耗尽的天然tRNA的密码子之外,半合成tRNA混合物中的单个tRNA大致与sGFPORF序列中它们的密码子丰度成比例。半合成混合物中的这些t7tRNA比表4-6中所示的它们密码子利用率以更低的比例被使用。以半合成tRNA套件在翻译反应液中生产对应于T1-3的sGFP的过程在荧光读板仪上监控3h,激发波长为485nm,发射波长为528nm。

[0116] 结果

[0117] 为了构建能够支持蛋白质翻译的合成大肠杆菌tRNA套件,我们在体外进行了48个DNA模板的run-off转录^{29,30},这些模板具有t7启动子及随后的相应tRNA编码序列(图8)。

[0118] 除tRNAIle (GAU) 和tRNA^{Trp} (CCA) 外,所有其他tRNA都获得了大量的t7tRNA转录本(图1)。编码这些tRNA的序列以腺苷开始,这可能导致在早期转录阶段的高中止率。对于这两种tRNA,在tRNA编码序列上游引入了前端为强转录起始位点的锤头核酶(HHRz) 编码序列,以保证HHRz介导的自身剪切后的有效转录(图8)³¹。变性PAGE分析显示,90%以上的RNA前体被剪切而生成期望的tRNA(图1)。虽然某些tRNA(如tRNA^{Leu} (CAA) 和tRNA^{Tyr}) 包含额外的小条带,但我们在所有情况下都获得了具有预期尺寸的大条带。

[0119] 体外肽表达试验

[0120] 随后,我们尝试设计一种体外合成试验以评价t7tRNA的解码效率。我们推测对应于使用有限组的密码子的短开放阅读框(ORF) 的报道肽相比于需要完整组的tRNA以进行合成的典型报道蛋白质(如GFP或荧光素酶) 更有优势。

[0121] 为建立一个多路(multiplexed) 定量均质肽表达试验,我们利用了我们团队最近开发出的肽生物传感器³²。该肽生物传感器由人工改造的肽结合结构域(称“亲和钳”³³) 和自身抑制的烟草叶脉斑点病毒(TVMV) 蛋白酶构成。一个8氨基酸配体肽RGSIDTWV(RGS肽) 结合至该生物传感器会导致引起蛋白酶激活的构象改变,蛋白酶的激活可通过淬灭的荧光素TVMV底物肽的剪切而检测到(图2A)。我们证实使用该试验可检测低至50nM的肽。TVMV底物肽剪切的初始速率(V_{max}) 与RGS肽的绝对浓度之间的关系在50–400nM范围内是线性的(图2B)。而且,该试验可在大肠杆菌无细胞系统中进行,虽然该系统需要蛋白酶抑制剂混合物来减轻内源性蛋白水解活性。

[0122] 虽然RGS肽结合至亲和钳严重依赖于游离的C末端^{34,35},但我们预期在RGS肽的N末端引入额外的氨基酸不会影响这种钳-肽结合³³。我们通过在我们的试验中测试两个N末端延伸的合成肽GG RGSIDTWV和DD RGSIDTWV确认了这一点(图9),所有三个肽的校准曲线都显示了良好的线性, $R^2 > 0.99$ (图2B和图9)。我们的结论是,我们的表达试验很适合在无细胞系统中定量RGS相关肽的表达,因此可被用来测试各种同工受体/密码子对的解码效率。为降低紧接在起始密码子AUG下游的密码子对翻译起始的影响,在起始密码子和测试密码子之间插入了两个GAU密码子以生成RGS2模板(图2A)。对照实验证实,这样的模板介导了始终如一的肽表达水平,而与RGS肽编码序列上游的测试密码子无关(图10)。

[0123] tRNA耗尽的大肠杆菌无细胞翻译系统的表征

[0124] 为获得tRNA耗尽的大肠杆菌S30细胞提取物,我们修改了之前公开的色谱法tRNA耗尽方案²⁸。在该程序中,内源性tRNA结合至乙醇胺-琼脂糖基质,而其他蛋白质合成所需成分保留在流通液中。为获得最佳tRNA耗尽效果同时保持翻译效率,我们优化了钾/镁浓度以及基质与裂解物比率。tRNA耗尽程度通过比较eGFP在添加或未添加总天然tRNA混合物的被耗尽裂解物中的表达水平来评价。如可从图3A中看到的,在缺乏天然tRNA时,没有产生eGFP,而添加了总天然tRNA混合物使得翻译恢复至原裂解物水平的60%。tRNA耗尽裂解物中观察到了约10分钟的时间延迟,这可能反映出重新加入的tRNA的氨酰化所需的时间。

[0125] 与eGFP类似,短RGS1模板的翻译(图2A) 也是tRNA依赖性的。加入天然tRNA和一个9密码子特异性t7tRNA混合物(表1) 使得翻译恢复至类似水平(数据未示出)。

[0126] 考虑到tRNAIle需要转录后修饰以实现有效的氨酰化²¹,因此观察到合成tRNA能够

支持RGS1模板的翻译是有些令人惊讶的。因此,我们进行了一个对照实验,其中我们从混合物中省去了单个t7tRNA,并测定了所产生的反应混合物的翻译活性。如图3B所示,当AUG (iMet)、CGG (Arg) 或UCC (Ser) 密码子的单个t7tRNA被独立地从t7tRNA混合物中排除时,肽表达水平显著降低。但是,当排除GGC (Gly)、AUC (Ile)、GAC (Asp)、ACC (Thr) 或UGG (Trp) 的t7tRNA时,观察到了一些残余表达,这可能是由于裂解物残留的天然tRNA导致的。CGG (Arg) 和UCC (Ser) 的tRNA耗尽效率为75–90%,表明这些密码子更适合重新分配,因为被耗尽裂解物中剩余的相应天然tRNA同工受体的量可忽略不计(未加入所选tRNA而没有产生肽产时被定义为100% tRNA耗尽效率)³⁶。这两个密码子属于由两个密码子家族构成的混合密码子盒,使得它们有希望成为用于重新分配的候选物(见下文)。

[0127] t7tRNA功能和特异性的系统分析

[0128] 所开发出的试验提供了系统性测试整套t7tRNA的平台(图2A)。然而,初步实验表明在耗尽的裂解物中存在残余量的一些同工受体。这些代表了在大肠杆菌中含量丰富的tRNA,表明在单个tRNA的耗尽效率与它们的丰度之间存在关系(表2)^{10,36}。不完全耗尽内源性tRNA可能因为掩盖了它们t7tRNA对应物的信号而使功能性测试复杂化。但是,我们观察到,在模板内包括特定tRNA的两个连续密码子显著增强了其耗尽在肽翻译效率上的不利作用。这可能反映出之前报道的低丰度tRNA在降低的tRNA浓度下体内翻译动力学的变化³⁷。因此,我们利用含有两个连续的靶密码子的报道肽ORF以不完全耗尽的同工受体重新筛选了密码子。解码效率计算为单密码子模板和双密码子模板的平均值(图11)。t7tRNA解码61个密码子的能力汇总于图4和图12–14中。

[0129] 解释试验结果时要考虑的一个重要方面是t7tRNA在裂解物中能够进行修饰的程度。例如,文献显示大肠杆菌粗提液介导了合成tRNA中假尿苷的形成³⁸。这种修饰仅需要异构酶活性,而不需要可能存在于我们系统中的低分子量底物或辅因子³⁹。但是,N34和N37上的修饰涉及多达20种具有中继链(relay chain)的酶以及多种底物和辅因子¹⁹。因此,在未显著优化该系统的情况下,这种修饰不太可能出现在粗提液中的合成tRNA上。我们在本文稍后通过实验解决了这个问题(见下文)。

[0130] 分割密码子家族盒

[0131] 以U和C结尾的分割密码子家族(如Ser、Phe、Tyr、His、Asn、Asp和Cys)的解码通过第一反密码子位置为G或其修饰形式(Q)的tRNA进行。以A和G结尾的密码子的解码使用经修饰的uri尿苷(Lys和Glu)或通过加入在第一反密码子位置为C的同工受体(Arg、Leu和Gln)(图4蓝色和灰色底纹的氨基酸)进行。前者中尿苷通过不同的使识别位点仅限于A和G结尾的密码子的氨甲基衍生物来修饰⁴⁰,此外,当两个Leu同工受体的第一反密码子位置为U和C时,还需要核糖2'-O-甲基化。

[0132] 当与它们的天然对应物相比时,Phe、His、Asp和Cys的t7tRNA在解码它们的具有Watson-Crick几何学的同源(以C结尾)密码子(也称为“同源WC”) (以C结尾) 和摆动(以U结尾) 密码子方面是有效的(50–70%)。t7tRNATyr (GUA) 解码UAC的效率为~50%,解码UAU的效率低于30%。T7tRNAHis (GUG) 具有额外的G–1C73碱基对,它被证实能有效解码其同源WCCAC密码子和摆动CAU密码子。缺少G–1的t7tRNA无法恢复肽翻译(数据未显示),推测是由于无法完成氨酰化步骤导致的⁴¹。

[0133] t7tRNASer (GCU) 识别AGC和AUG密码子的效率分别是123%和75%。对于Leu和Arg,

两个tRNA同工受体负责解码每个分割密码子盒,在我们的试验中,带UAA和CAA反密码子的Leu t7tRNA只通过典型的WC碱基对识别它们的同源密码子。这之前报道的通过未修饰的尿苷的受限识别模式一致⁴²,虽然效率分别为36%和88%。带UCU和CCU反密码子的Arg t7tRNA都非常严格地识别它们的同源WC密码子,效率分别为340%和202%。这些t7tRNA较高的表观活性可能反映出这些同工受体在天然tRNA混合物中的低丰度³⁶。与之前从核糖体结合试验中观察到的一致⁴³,t7tRNAGln (UUG) 不能解码其同源WC CAA密码子。t7tRNAGln (CUG) 也不能解码CAA密码子,即使它能够以40%的效率解码其同源WC CUG密码子。

[0134] Glu (UUC)、Ile (GAU)、Asn (GUU) 和Lys (UUU) 的T7tRNA无法支持从包含它们的同源WC密码子和摆动密码子的模板翻译出肽。之前的报道显示,Glu和Ile t7tRNA的反密码子环内缺少修饰会阻止它们的氨酰化,使得它们在肽翻译中无活性^{44,45},并且无论有没有Hrz辅助,所制备的t7tRNAAsn (GUU) 在报道肽合成中都表现不佳。虽然具有来源于tRNAAsn的嫁接反密码子和识别位碱基 (discriminator base) 的嵌合t7tRNALys能够被AsnRS以Asn氨酰化⁴⁶,但它仍无法在我们的试验中支持肽翻译 (数据未显示)。关于这点,之前报道称具有未修饰的U34的tRNALys无法解码其任何密码子,可能是因为反密码子环中结构等级 (structural order) 丢失以及由三个连续的最不重叠的A-U碱基屏形成的密码子-反密码子复合物内的堆垛 (stacking) 不佳^{47,48}。在我们的系统中,将t7tRNALys的第一反密码子位置的U突变为C完全恢复了其对AAG密码子的解码活性。这种作用可能来源于反密码子环内以及密码子-反密码子螺旋内的胞苷提供的更强的堆垛作用,以及来源于对赖氨酰基aaRS更强的亲和力⁴⁹。

[0135] Trp (CCA) 和Met (CAU) 的T7tRNA分别以46%和61%的效率解码它们的同源WC密码子。

[0136] 未分割密码子家族盒

[0137] 仅除Arg之外,2或3个在第一反密码子位具有G、U和/或C的同工受体的标准子集 (standard subset) 被用在细菌中解码8个未分割密码子家族盒。此处,G34与C-和U-和C34配对以专门介导以G结尾的同义密码子的解码 (图4蓝色和绿色底纹)。在肽合成试验中,所有在第一反密码子位具有G或C的t7tRNA同工受体都以40%至160%的效率特异性识别它们的同源WC密码子 (图4)。在除tRNAGly (UCC) 外的所有天然tRNA同工受体中,U34携带5'-羟乙酸修饰,该修饰通过部分地改变核苷酸糖折叠几何结构^{51,52}而将其对同源WC以A结尾密码子的识别⁵⁰延伸至Val、Pro和Ala的同源WC以G-、U-和C结尾密码子。

[0138] 而且,t7tRNA^{Leu} (UAG) and t7tRNA^{Ser} (UGA) (推测无修饰) 显示出对以A结尾密码子强烈的偏好,对以U结尾密码子较低程度的偏好,但无法识别以G和C结尾的密码子²⁴。解码两个连续UCU密码子的tRNA^{Ser} (UGA) 在1.6μM时的相对效率为~10%,而该同工受体浓度增加至6.4μM时,解码效率升至~70% (图12A)。这个发现很好理解,因为tRNA (包括反密码子环) 中缺少一些转录后修饰会导致对aaRS的亲和力降低或密码子-反密码子相互作用的较高解离速率和/或tRNA入位 (accommodation)⁵³。这些效应至少部分地能够被tRNA浓度的增加抵消。

[0139] 惊人的是,Val、Pro、Thr和Ala的最可能具有未修饰U34的t7tRNA显示出与它们的天然对应物类似的密码子阅读模式,即这些t7tRNA不仅有效地解码它们的同源WC以A结尾密码子,而且能够以较低的程度解码以U和C结尾的密码子⁵⁴。以G结尾的密码子的解码表明

对于Val和Ala(缺少含C34的备份同工受体)具有强烈的U34-G3介导的识别。这与对Pro和Thr的低效的U34-G3介导的识别(图4)相反,它们可能通过同源WC同工受体介导。

[0140] 天然tRNA^{Gly} (UCC)与解码未分割密码子盒的在第一反密码子位为U的其他tRNA的区别在于它在U34携带氨酰化修饰(见上文),这是解码分割密码子盒的tRNA的特征⁵⁰。

[0141] T7tRNA^{Gly} (UCC)是上述关联性的一个例外,因为它有效地解码以C和G结尾的密码子,即使是存在用于后者的同源WC解码的C34同工受体。

[0142] 如上所述,解码细菌的未分割家族盒的四个Arg密码子是不寻常的,因为它依赖于两个同工受体,其中一个携带在真核细胞中更加常见的肌苷修饰。这种修饰能够通过分别与摆动和WC几何结构进行碱基配对来解码以A、U和C结尾的密码子。t7tRNA (ACG)的未修饰的反密码子茎环显示出对其同源WC密码子CGU的几乎相同的亲和力,但是在结合至其摆动CGC和CGA密码子方面低效⁵⁶。在我们的研究中,t7tRNA^{Arg} (ACG)不仅能够有效地解码以U结尾密码子,而且能够有效解码以C和A结尾的密码子,以及以较低效率解码以G结尾密码子。

[0143] 上述实验证实,合成tRNA能够在体外对于17个氨基酸在功能上替代它们的天然对应物。重要的是,对所有测试的t7tRNA,未观察到交叉识别,无论是对来自编码不同氨基酸的相同家族的密码子(图4,Ser和Arg、Leu和Phe、His和Gln、Asn和Lys)或对于属于具有相同混合密码子盒的不同家族的同义密码子(如Ser、Arg和Leu)而言都是如此。通过连接有未分割密码子盒的分割密码子家族来分析这三个氨基酸是特别令人感兴趣的,因为前者和后者具有非重叠的解码模式并且代表可能的可再指定的密码子。

[0144] 从大肠杆菌天然tRNA混合物中分离Glu、Asn和Ile的天然tRNA

[0145] 我们的研究结果表明,通过使用体外转录的tRNA可将大多数氨基酸纳入蛋白质中。为了重新构建能够支持含有所有20个标准氨基酸的蛋白质翻译的tRNA混合物,我们需要有效地解码Glu、Asn和Ile的密码子。为此,我们决定通过DNA/RNA杂交色谱从天然tRNA混合物中纯化特异于这些氨基酸的天然tRNA^{57,58}。我们测试了几种固定化策略,并通过将3'-胺化的寡核苷酸偶联到NHS-琼脂糖凝胶上获得了最佳的结果(图15)。通过与对靶标tRNA的D环和反密码子环互补的寡核苷酸选择性杂交,我们从天然tRNA混合物中成功获得具有良好纯度的所有三种Glu、Asn和Ile的特异tRNA(表3)。通过热变性我们将tRNA从基质中洗脱,通过变性聚丙烯酰胺凝胶电泳(PAGE)分析显示其纯度>90%(图5A)。如上所述,我们使用具有它们的同源性WC或简并同义密码子的模板测试了纯化的天然Glu(SUC)、Asn(QUU)和Ile(GAU)tRNA的功能性。我们使用两个具有连续的GAA或GAG密码子的模板测试纯化的tRNA^{Glu} (SUC)的功能,如图5B所示,其可以有效地解码两个密码子。类似地,纯化的天然天冬氨酸tRNA^{Asn} (QUC)恢复了具有AAU或AAC密码子的模板的翻译。纯化的天然的tRNA^{Ile} (GAU)解码AUU密码子的效率只有40%,这可能是由于修饰的同工受体变体变性或共分离后低效的重折叠导致的。

[0146] 半合成蛋白质翻译系统

[0147] 我们为20个标准氨基酸中的每一个获得至少一种纯化的功能性tRNA。为了测试这种简化的tRNA套件是否可以支持全长蛋白质的合成,我们合成了三种编码超折叠GFP(sGFP)59的DNA模板,这些模板具有可变的密码子组成。设计这些模板是为了排除由合成tRNA(如CCU(Pro)、UAU(Tyr)、CAA(Gln)和AAA(Lys))(表4-6)低效解码的密码子。模板在补充有不同半合成tRNA混合物(表4-6)的tRNA耗尽的裂解物中表达,其支持翻译所有三种模

板,其效率与天然tRNA混合物相当(图6A)。

[0148] 为确保sGFP表达是半合成tRNA的补充的直接结果,并再次证实各个tRNA的功能,我们制备了缺乏单种tRNA的tRNA混合物。然后,我们分析了这些混合物在tRNA耗尽的无细胞系统中支持sGFP模板合成的能力,并观察到了翻译降低了数倍至数个数量级(图6B)。与通过肽表达试验(包括编码UCG (Ser)、CGG和AGG (Arg)、UUG (Leu)、GGA (Gly)、CCA (Pro)、ACA (Thr)、GUG (Val)、AAG (Lys)和UUC (苯丙氨酸))所得到的结果类似,半合成tRNA混合物恢复了sGFP表达,从而再次确认相应的t7tRNA的功能。然而,与RGS-肽表达模式相反,CUA和AGC密码子的残余天然tRNA被证明足以用于sGFP表达。当报告肽的ORF序列中仅存在一个CUA或AGC密码子时,混合物中存在t7tRNA (UAG)和t7tRNA (GCU)分别使肽表达恢复至92和72%(表2)。这种不一致可能反映了与sGFP相比,肽转录本的浓度更高以及肽翻译的转换率(turnover rate)更高。在前一种情况下,延伸复合物的数量可能超过裂解物中保留的天然tRNA的数量,而在后一种情况下,相同的tRNA可能在相同的多聚核糖体单元(polysomal unit)中穿过^{37,60}。

[0149] 当在ORF中仅存在一个密码子时,sGFP表达显得对耗尽的裂解物中相应天然tRNA的残留量更敏感。与CUA和AGC密码子不同,当排除AGG的t7tRNA时,每个ORF只有一个AGG密码子的sGFP的表达水平下降约90%。这使得AGG成为最有希望的用于重新分配的密码子,即使不进一步优化天然tRNA的耗尽。

[0150] 在我们的无细胞系统中可以表达全长蛋白质的观察结果提示我们探索tRNA修饰在蛋白质翻译中的作用。如上所述,尽管我们系统中的大多数tRNA是合成的,但它们可能在有翻译活性的裂解物中被部分编辑或修饰。

[0151] 为了测试这种推定的修饰对t7tRNA的功能的影响,我们使用重构的PURE在体外翻译系统中重复了上述实验,该体外翻译系统可假定缺乏tRNA加工和修饰活性。我们发现,半合成的tRNA可以维持蛋白质的表达,但是我们观察到,补充了半合成tRNA的PURE系统的翻译效率只有天然tRNA套件的60%。这与我们观察到的天然的和合成的tRNA套件在耗竭的裂解物中表现得几乎一样好的结果相反。一方面,这表明未经修饰的合成t7tRNA能够维持蛋白质合成。另一方面,观察到的效率降低可能反映了S30提取物中至少一些t7tRNA的额外的转录后加工,但在PURE系统中没有这种额外的转录后加工。最终解决这个问题将需要测试PURE系统中各个t7tRNA的功能。由于污染性天然tRNA的水平出奇地高,所以这并不简单(图7)。

[0152] 讨论

[0153] 在这项工作中,我们建立了一个使用耗尽了内源tRNA的体外翻译系统来系统分析单个tRNA功能的方法。这是通过开发出一种非放射性试验来定量报道肽在tRNA耗尽的大肠杆菌无细胞翻译系统中的体外表达来实现的。我们证明了耗尽的裂解物保留了超过60%的亲本裂解物活性。尽管耗尽的裂解物中残留tRNA库不能维持eGFP和RGS-肽的表达,但我们发现一小部分天然tRNA没有完全耗尽。所开发的肽生物传感器试验允许我们估计tRNA同工受体相对于它们的密码子的耗尽水平,并揭示单个tRNA的耗尽效率与其丰度之间的相关性。重要的是,在加入缺乏单独tRNA的t7tRNA混合物的完全重组大肠杆菌PURE系统中也观察到了RGS1-肽的表达(数据未显示)。此外,没有外源tRNA组装的PURE系统可支持全长GFP的表达(图7),表明存在整个系列的污染性tRNA。这表明残留的tRNA最有可能与aaRS或翻译

机制的其他组成部分共纯化 (copurify)¹⁶。因此,从体外翻译系统中有效地耗尽tRNA仍然是一个挑战。

[0154] 为了区分t7tRNA活性与内源tRNA背景活性,我们利用了观察到的结果:两个相同的连续测试密码子显著地敏化了对特定tRNA的耗尽的试验。这使我们能够分析所有48种大肠杆菌tRNA的合成形式在我们的tRNA耗尽的裂解物中的功能。

[0155] 我们的研究结果表明大多数合成的tRNA在支持蛋白质/肽翻译方面是有效的(图4-7)。此外,大多数对应于2或4倍简并氨基酸的t7tRNA以高或中等效率解码其同源WC密码子,并以中等或低效率解码其摆动密码子(图4)。相反,我们发现Asn、Gln (CAA)、Ile (AUC)、Glu和Lys的t7tRNA在亲和力钳试验中是无功能的。

[0156] 尽管我们不能排除在该粗制 (crude) 翻译系统中合成tRNA的部分编辑和修饰,但是看起来不太可能有效地发生需要多种酶的活性的N34和N37修饰。这一观点得到了合成tRNAs不能支持Ile、Glu和Lys合成的观察结果的支持,并且符合先前的研究(修饰的核苷酸作为其同源aaRS²¹的关键分子识别特征)。之前有报道称,在反密码子环以外缺乏修饰的tRNA^{Lys} (UUU) 以140倍的更低效率进行氨酰化⁶²,然而,在第一个反密码子位置U到C的转化恢复了其对AAG密码子的活性。除了Lys (UUU) 之外,许多t7tRNA转录本或其相应的反密码子茎环(诸如Arg (UCU)、Ala (UGC)、Cys (GCA)、Glu (UUC) 和Gln (UUG)) 也在核糖体介导的密码子结合试验中未能成功结合⁵⁰。除了Glu (UUC) 和Gln (UUG) 之外,剩余的三个t7tRNA在亲和力钳试验中是有翻译活性的。特别地,t7tRNA^{Arg} (UCU) 显示出比其同源物高3倍的活性,其可能在总天然tRNA混合物中未被充分代表 (underrepresented) (图4)。总体而言,图4中所示的密码子-反密码子相互作用矩阵显示了天然的和t7tRNA之间高度相似的密码子识别模式,其最有可能在反密码子环内缺乏修饰。例如,t7tRNA的第一个反密码子中的U34不仅解码其同源性A和G,而且还解码第三密码子位置上的U和C,与Ser、Leu和Gly的天然tRNA中的cmo5U具有相似的阅读模式,并且与Val、Pro、Thr和Ala的tRNA具有相同的模式。从结构和动力学数据来看,密码子-反密码子螺旋的固有稳定性显得不如其合适的几何形状(由核糖体感测到)^{63,56,57}。因此,这些研究意味着密码子-反密码子复合物与核糖体之间的净亲和力促进了30S围绕解码中心闭合,从而促进tRNA入位并触发导致肽键形成的下游步骤⁵⁴。在目前的研究中,U34-N3相互作用的高度镶嵌模式以及在分割密码子框中缺乏非同源交叉识别支持了tRNA本体内具有更高级别脉络性 (contextuality) 的想法—可能存在额外的用于准确和有效解码的检查点^{56,57}。

[0157] 本文显示的解码偏好为从特定氨基酸的同义密码子中识别“正交的 (orthogonal)”和“天然的”密码子对提供了有价值的指导。这样的配对可以由不同的6倍简并氨基酸家族的密码子产生,或从衍生自具有高摆动限制的未分割密码子家族盒的密码子产生,如Arg、Ser和Leu和Pro、Thr和Gly (表7)。这项工作表明,天然tRNA几乎完全耗尽的AGG密码子可能比其他密码子更容易重新分配。

[0158] 我们在本文展示了除Ile、Glu和Asn之外的所有氨基酸都可以被合成的tRNA解码,并且这三个氨基酸的天然tRNA被以功能形式纯化到同质性。我们证明,用合成的tRNA和三种特定的天然tRNA重构的tRNA套件可以支持sGFP的体外合成,达到与天然tRNA混合物相当的水平。虽然完整的tRNA缺失仍然是一个挑战,但我们使用PURE系统的结果为污染性tRNA库的来源提供了线索。

[0159] 改进的tRNA耗尽方案与半合成的tRNA套件相结合可以重新分配肽和蛋白质中的有义密码子,从而显著扩大合成生物学家和蛋白质工程师的工具箱。此外,所开发的与PURE系统结合的肽表达试验能够解析各个tRNA修饰对其功能的影响。这反过来应当回答了一个长期存在的问题,即在构建最小细胞的过程中需要维持这种修改的程度⁶⁴。最后,所提出的方法不限于大肠杆菌,而是可以转移到真核无细胞表达系统中。

[0160] 额外的支持性信息

[0161] 48种t7tRNA转录本的制备

[0162] 我们使用3步PCR法来制备两种类型的用于tRNA体外转录的DNA模板。第一种类型是由2个正向(T7、F)和3个反向寡核苷酸(R1、R2、R3)组装而成的,用于以G、U和C开始的tRNA种类(图8A),而另一种类型由3个正向(T7、F1、F2)和3个反向寡核苷酸组装而成,用于以A和U开始的tRNA种类,以产生自加工锤头状核酶(HHRz)融合构建体(图S1B)²。F1引物含有T7启动子,接着是用于降低转录失败率的GGGAGA序列、与靶tRNA的5'部分互补的4-10个碱基、以及一段HHRz编码序列。F2含有HHRz-编码序列,接着是与tRNA的5'一部分互补的片段。在进行3步PCR法后,通过乙醇沉淀纯化产物,并用作材料和方法中所述的T7转录的模板以获得所需的tRNA。

[0163] t7tRNA转录本的3'和5'特性

[0164] 所产生的合成tRNA预期含有生理性的3'-羟基。取决于制备方法,5'-可以携带5'-单磷酸、5'-三磷酸或5'-羟基。非生理性的5'-三磷酸可以被aaRS和其他翻译机制接受^{67,68},或者可以通过内源性磷酸酶活性(例如RNA焦磷酸水解酶⁶⁹) 在粗制裂解物中被加工成单磷酸。

[0165] 在使用转录本的5'-核酶介导的加工的情况下,切割的产物含有5'-羟基,据报道,其不会影响Ser、Met、Phe、Tyr、Asp的tRNA转录本的氨酰化效率⁶⁶。尽管含有“CAT”反密码子的48种tRNA基因中的5种被指定到基因组tRNA数据库中的Met⁶,但是根据Modomics数据库的序列⁷¹,其中仅3种实际上编码Met(2个用于启动子,1个用于Met延伸密码子),而剩余的2种用于编码Ile。

[0166] 用于高效肽表达的RGS模板的优化

[0167] 为了测试不存在于RGS-肽序列中的密码子的t7tRNA种类的功能性,我们将测试密码子首先插入在RGS-肽序列和起始密码子之间。然而,启动子AUG附近的插入密码子可能影响翻译起始⁷²,从而使结果的解释变得复杂。因此,我们决定在起始密码子之后插入一个或两个不变的“绝缘子”密码子,以最小化下游密码子插入的作用。为此,我们测试了具有不同G/C含量的四种不同密码子,例如Thr的ACC、Ile的AUC、Gly的GGC、Asp的Asp,并将一个或两个密码子置于AUG起始密码子和肽编码框之间。其中,含有两个连续GAU密码子的模板显示出最佳的翻译效率,这通过亲和力钳试验可证明,而来自其他模板的信号很低或没有信号(图10A)。我们推测肽末端Asp促进肽从核糖体出口通道的释放,这是通过赋予前者的净负电荷实现的⁷³。我们构建了几个在两个不变GAU密码子之后具有可变密码子的DNA模板,所有可变密码子都在亲和力钳试验中表现良好,而不管使用的密码子如何(图10B)。因此,我们使用这个亲本模板来测试所有其他密码子。

[0168] 天然tRNA纯化策略的优化

[0169] 制备固定化的寡核苷酸矩阵

[0170] 常用的寡聚DNA的固定策略是基于生物素-链霉亲和素相互作用⁷⁴。然而,由于在用于tRNA洗脱的变性条件下矩阵不稳定(matrix instability),这种稳健和灵活的方法对于我们的目的而言不是最优的。因此,我们测试了基于胺-NHS和巯基-碘乙酰胺偶联的固定化学物质。具有3'-胺、3'-生物素或3'-巯基基团的寡聚DNA通过IDT合成。具有3'-巯基的寡聚DNA在固定之前需要还原,这增加了固定方案的复杂性。分别用112、150和55纳摩尔的胺、生物素和巯基官能化的寡聚DNA被固定,以产生1ml的各沉降树脂。由于成本高,方案复杂,结合效率低,放弃了用碘乙酰基树脂通过巯基反应的固定策略。由于胺/NHS偶联反应的N-羧基琥珀酰亚胺副产物在260nm处具有高吸光度并强烈影响未结合的寡聚DNA的定量,所以结合的与未结合的寡聚DNA的比率(作为偶联程度的量度)通过反应上清液的非变性PAGE来测定。

[0171] tRNA纯化

[0172] 据报道,四甲基氯化铵(TMA-Cl)可以增强tRNA-寡聚DNA杂交体的形成⁷⁵。因此,除了含有0.9M NaCl的标准杂交缓冲液之外,我们还测试了含有0.9M TMA-Cl的缓冲液。在胺偶联的基质上从100μg的天然tRNA总混合物中纯化出约2.25μg的tRNA^{Glu},相当于1nmol的固定的寡聚DNA,而链霉抗生物素基质(通过生物素固定的寡聚DNA)的产量低约3倍。

[0173] 在后一种情况下,洗脱的tRNA含有一部分高温共洗脱寡聚DNA,其在凝胶上迁移至tRNA前方的下面(在图15中用星号标记)。因此,NHS基质被用于纯化天然tRNA的所需要的部分。基于所报道的方法¹¹的最终纯化方案通过优化如材料和方法的相应部分中所述的洗涤步骤来进一步改进。除了tRNA^{Asp}以外,所需部分的所有天然tRNA以高产率和高纯度获得(图6)。

[0174] 实施例2

[0175] 使用吻合环选择性耗尽tRNA

[0176] 吻合环(kissing loop, KL)通过形成准连续双螺旋而以高亲和力(5-50nM)结合GCU Ser同工受体,该准连续双螺旋由核苷酸对平面之间的共同堆积柱(common stacking column)来稳定(图16A)。通过在乙醇胺琼脂糖凝胶上的层析,大肠杆菌S30裂解物可以在很大程度上耗尽总tRNA。耗尽后,如果补充了市售的总tRNA部分,裂解物保留其亲本的翻译效率的~60%(图16B)。

[0177] KL固定的基质可用于从粗制裂解物中直接耗尽特异性tRNA同工受体,或间接地通过从总tRNA混合物中提取同工受体,再用全tRNA耗尽的裂解物来重构(图(B))。使用任一方法制备的反应中的GFP合成显示在具有两个连续的AGC密码子的密码子偏好模板上的~100%暂停(图16C)。为此目的,将“吻合环”编码序列(大写)GGTAGTGAGGTAGTTAGCAGATACCTCACTACcaacacacacaacacacacacaagct与接头(小写)克隆到3'-侧接HindIII位点的T7启动子的下游,然后从HindIII消化的质粒进行runoff T7转录。为了获得KL-固定的基质,根据PMID:14652075,通过高碘化钠将KL-RNA氧化,并与己二酸二酰肼-琼脂糖珠(Sigma:A0802-50ML)偶联。在第一种方法中,将两批含有~20-25nmol KL的1g基质与2.5ml的用5mM Mg(OAc)₂、25mM HepesKOH(pH7.6)、0.1mM EDTA缓冲的粗制裂解物在RT下反复温育30分钟,然后在P-10凝胶过滤柱上进行缓冲液交换。在第二种方法中,在相同的条件下,将125mg的KL珠在200μl反应体积中重复孵育,然后通过乙醇沉淀tRNA,所述反应体积含有5μl的30mg/ml总大肠杆菌tRNA(Roche:10109541001)。

[0178] 我们发现tRNA (Ser) GCU同工受体耗尽高,虽然不完全,某些小的tRNA部分似乎是具有氨酰基-tRNA合成酶 (ARSase) 或EF-Tu的复合物的一部分,从而逃逸出KL陷阱 (trap)。我们试图通过采取“ARSase反冲”过程来实现更佳的耗尽效果 (参见图17, 下图)。在这种方法中,将翻译裂解物与上述反应混合物中的KL-固定的树脂一起温育,但另外含有23mM的Mg (OAc) 2、5mM的KOAc、5mM的KCl,并且存在a) 过量的反同工受体 (tRNA (Ser) gga) (15μM), 其负责解码密码子偏向模板中除目标AGC (由被耗尽的同工受体解码) 之外的所有Ser-密码子, b) 过量的其他两个丝氨酰氨酰基-tRNA合成 (SerRS) 底物,如ATP (20mM) 和丝氨酸 (5mM), 以及c) 1mM GDP以解离Aa-tRNA与EF-Tu的复合物。过量的SerRS底物加速了酶的转换动力学,导致有效的同工受体替代,其中感兴趣的同工受体tRNA (Ser) GCU被释放并且可用于KL陷阱。在翻译不含AGC-靶密码子 (4,5“cg”-“;4,5’tcc”-“)、包含2个连续的AGC-密码子 (4,5agc) 或包含在不同位置 (所有其他位置) 的单个靶密码子的Ser密码子偏爱性模板时 (图18, 上图), 与没有进行“反冲”的对照完成 (图13, 蓝色和红色条) 相比,进行了“ARSase反冲”程序的裂解物 (图13, 第3和第4图: 绿色和紫色条) 显示出更高的tRNA耗尽程度。

[0179] 实施例3

[0180] 体外氨酰化

[0181] 为了进行uAA的位点选择性并入,我们最初建立了三个体外tRNA氨酰化方案。第一种基于CysRS介导的突变tRNA (Cys) 的体外氨酰化方案。反应性巯基使得能够用马来酰亚胺或碘乙酰胺活化的试剂快速衍生酰化tRNA (Cys)。通过该途径将BODIPY-Cys并入到EGFP的独特AGG密码子中。我们还建立了吡咯赖氨酸RS (PylRS) 介导的吡咯赖氨酸^{CUA}tRNA的酰化和Flexizyme介导的tRNA氨酰化。我们证明,我们可以将已被反式环辛烯 (Lys-COC) 衍生化的Lysε-氨基并入大肠杆菌无细胞系统中重新分配的Ser AGC密码子中。重要的是,我们观察到,与琥珀抑制相比,将COC并入有义密码子对蛋白质表达产量的影响要小得多。

[0182] 实施例4

[0183] 半胱氨酸tRNA合成酶介导的氨酰化

[0184] 半胱氨酰-tRNA合成酶 (CysRS) 与tRNA (Cys) 的反密码子三联体具有接触区域。尽管酶对突变tRNA的亲合力降低,但是在体外高底物浓度下仍然可以支持高氨酰化速率,但是我们发现这样的突变tRNA在翻译反应过程中不能进行有效的再酰化,因为内源性CysRS浓度不足且野生型tRNA (Cys) 竞争活性高。这确保了只有uAA被并入到增长的多肽链中。

[0185] 将wt t7tRNACys GCA的反密码子改变为CCU或GCU以匹配相应的Arg-AGG或Ser-AGC密码子 (参见图19A)。使用具有CUA反密码子的琥珀密码子抑制剂t7tRNACys,通过监测含eGFP-ORF的翻译反应中的荧光产量来优化氨酰化条件,其中第153位的天冬氨酸密码子被转化为琥珀终止密码子。在包含50mM HepesKOH (pH8.0)、1mM半胱氨酸或2mM硒代半胱氨酸 (Secys)、10mM MgCl₂、5mM KOAc、5mM KCl、4mM ATP、22.5μM tRNA、10mM CysRS、20% DMSO、25mM DTT、5μM ZnCl₂、0.005u/ml酵母无机焦磷酸酶和50μg/ml BSA的反应液中,两种t7tRNA都通过重组CysRS在体外被负载了半胱氨酸。在反应前,在含有40mM Secis、50mM HepesKOH (pH8.0)、1M DTT的溶液 (通过100mM KOH使pH最终达到7.5) 中,将硒代胱氨酸 (Secis) 在37℃反应30分钟而还原出Secys。通过苯酚提取而纯化半胱氨酰化的或硒代半胱氨酰化的tRNA (Cys (Secys)-tRNA), 用乙醇沉淀,并在含有50mM TrisHCl (对于Secys-tRNA, pH8.5或7.2)、100mM NaCl、75% DMSO、80μM Cys (Secys)-tRNA和1mM碘乙酰胺或马来酰亚胺

衍生物的反应液中与感兴趣的碘乙酰胺(马来酰亚胺))化合物偶联,或直接用于在对照反应液中重新编码半胱氨酸的AGG(Arg)和AGC(Ser)密码子(参见图19A)。

[0186] 由于与半胱氨酸相比,硒代半胱氨酸残基的pKa和还原电位均较低,所以在偶联反应中,我们可以获得数倍高产量的BPFL-tRNA偶联产物,从而导致平均25%的用于氨酰化的原始tRNA材料被转化为BPFL-Se-tRNA偶联物。

[0187] 具有CCU-反密码子的半胱氨酸酰化的或硒代半胱氨酸酰化的t7tRNA(cys-tRNA(Cys)ccu)被显示以浓度依赖性方式维持eGFP生物合成,翻译反应液由tRNA全耗尽的裂解物和缺乏t7tRNA(Arg)ccu的半合成tRNA套件重构而成,并加入含有单个AGG或TAG密码子的eGFP编码模板(参见图19B,D)。添加非氨酰化的tRNACysccu(图19B)不能保持eGFP翻译,导致与缺失任何AGG密码子抑制剂的阴性对照反应中的背景翻译相比,使用10uM tRNACysccu时仅有轻微的作用(图B中的“半合成tRNA-1”)。后者表明具有改变的反密码子的翻译前氨酰化tRNACys是一种简单的正交系统。

[0188] 在具有使用KL直接耗尽内源性tRNA(Ser)GCU的裂解物的翻译反应液中(图19C),携带接枝的GCU-反密码子的半胱氨酸酰化的tRNACys(Cys-tRNA(Cys)gcu-参见图19A,C)有效抑制两个连续的AGC-密码子。

[0189] 实施例5

[0190] 大环状蛋白质设计和测试

[0191] 因为这个项目的一个方面是要创建具有分子内键的肽,我们将测试几个允许同质缩合和异质缩合的uAA。图20提供了一种方法的概览。在最简单的情况下,我们将使用硒代半胱氨酸(Se-Cys),它在生理条件下自发形成基本上不可逆的二硒键。现有的证据表明它可以通过CysRS被负载至tRNA(Cys)。我们将测试Se-Cys并入到一个或多个密码子中的效率,并使用质谱和内肽酶保护试验分析产物的自发环化。然后,我们将分析所开发的系统包含相互正交的反应基配对的能力。

[0192] 我们将构建测试mRNA,其编码携带Ser和Arg密码子和C端亲和性钼标签的合成10聚体肽。最初,两个AGC密码子将被包括在mRNA序列中。构建体将在含有CUAtRNA的无细胞体系中表达,其中CUAtRNA负载了Se-Cys或携带叠氮和炔或更多反应性反式环辛烯的氨基酸。

[0193] 预期含有Se-Cys的肽会自发地环化,反应的程度将通过质谱测定。在叠氮化和炔烃修饰的氨基酸的情况下,将催化性Cu(I)加入到翻译混合物中以诱导点击反应和内键的形成。在环辛烯的情况下,功能化的氨基酸环化将通过添加具有两个偶联的四嗪基团的试剂来实现(图20E)。我们由市售试剂合成了双四嗪(di-tetrazine),并证明它可以交联包含环辛烯-赖氨酸的EGFP分子(图20G)。随后,我们将合成具有由四嗪基团官能化的芳族支架的更刚性变体。虽然最初的实验将使用仅包含两个四嗪基团的试剂进行,但将开发具有三个基团的后续支架(图20F)。这样的试剂将提供多重交联并导致高度约束的大环结构。在对照实验中,缺乏重新分配密码子的基因将被翻译并经相同条件处理。所得产物将被亲和纯化,并通过LC-MS以天然形式或胰蛋白酶消化后形式进行分析,以鉴定新的分子间键。

[0194] 我们将生成携带Lys或Cys的重新分配tRNA库。具体来说,我们将重点放在叠氮修饰和炔修饰的氨基酸和用反式环辛烯和四嗪修饰的氨基酸。前者在Cu(I)存在下形成共价1,4-位置异构体,而后者则自发反应形成杂环。可以选择性地与上述基团偶联的各种荧光团和亲和试剂可商购获得,并将用于评估该方法的效率和选择性。

[0195] 实施例6

[0196] 大环肽和肽文库的构建与选择

[0197] 这里公开的技术为构建和选择大环肽奠定了基础。为了将我们的大环化方法与选择系统结合起来,我们将利用RNA或DNA展示,如CIS展示。CIS展示利用了DNA结合蛋白RepA的专门结合回其编码DNA的能力。该系统支持超过 10^{13} 的有效库大小,并利用大肠杆菌无细胞系统,这完美地适合我们的目的。或者,这可以用体外组装的病毒或噬菌体进行。

[0198] 我们将生成两个CIS展示兼容库,用于选择大环结合剂。我们首个10聚体文库将包含与酪氨酸交错的可用的重新分配密码子,所述酪氨酸已知介导形成高亲和力蛋白质:蛋白质界面。第二个库将通过计算机设计,代表5000个经选择的、覆盖所有类别的蛋白质折叠。密码子的使用将被优化以使该文库中重新分配的密码子的出现最大化。作为我们最初的目标,我们将选择之前被确定为MAP激酶途径的可靶向组分的citron激酶的CNH结构域。在重组CNH结构域上选择通过不同密码子重新分配而环化的文库,并使用下一代测序比较选择程序的结果。然后将肽和成为线性和环化形式,并用于AlphaScreen分析和单分子重合相互作用(single-molecule coincidence interaction)分析。最终,将在具有 10^8 个变体的天然Phylomer文库上测试该方法,并且将通过初始文库和已选文库的测序来比较文库的性能。

[0199] 所提出的项目将提供一个新的方法来构建和分析高度多样化的大环肽库。它允许存在所产生的肽的大量的翻译后结构性多样化,而不需要产生新的文库。这种方法很可能导致形成用于产生高效生物活性化合物的新平台。它还将为团队中的科学家提供生物正交化学和基于多样性的活性化合物开发方面的使用培训。

[0200]

氨基酸	M		R		G		S		I		D		T		W		V	
RGS1 模板中的密码子	AUG		CGG		GGC		UCC		AUC		GAC		ACC		UGG		GUC	
t7tRNA	iMcau		Rccg		Ggcc		Sgga		Igau		Dguc		Tggg		Wcca		Vgac	
t7tRNA 编号	c1	c2	a4	b2	c12	b5	a8	d4	d5	d7	d10	d11						
t7tRNA 浓度 (μ M)	0.4	1.2	0.8	0.8	0.8	0.8	0.8	0.8	0.4	0.8	0.8	0.4	0.4	0.4	0.4	0.4	0.4	0.4

[0201] 表1. 用于RGS1模板的t7tRNA套件。表中显示了RGS1模板中的密码子和相应的t7tRNA(参见图S1的命名)。无细胞翻译反应中引发剂(iMcau)和延伸tRNAs的终浓度分别为1.6和0.8 μ M。对于由2种遗传变异体例如iMcau(c1,2)、Tggg(d4,5)和Vgac(d10,11)编码的tRNA,将各种多态性同工受体以对应于其基因组拷贝数的比率混合。

[0202]

氨基酸	密码子	DE.	氨基酸	密码子	DE.	氨基酸	密码子	解码	氨基酸	密码子	DE.
Phe	UUU	90% ± 1%	Ser	UCU	68% ± 1%	Tyr	UAU	91% ± 1%	Cys	UGU	84% ±
Phe	UUC	81% ± 2%	Ser	UCC	79% ± 2%	Tyr	UAC	85% ± 1%	Cys	UGC	37% ±
Leu	UUA	86% ± 1%	Ser	UCA	92% ± 1%	Stop	UAA		Stop	UGA	
Leu	UUG	95% ± 1%	Ser	UCG	94% ± 5%	Stop	UAG		Trp	UGG	31% ±
Leu	CUU	87% ± 1%	Pro	CCU	95% ± 2%	His	CAU	89% ± 3%	Arg	CGU	46% ±
Leu	CUC	62% ± 9%	Pro	CCC	99% ± 10%	His	CAC	83% ± 9%	Arg	CGC	11% ±
Leu	CUA	92% ± 4%	Pro	CCA	94% ± 2%	Gln	CAA	61% ± 2%	Arg	CGA	9% ±
Leu	CUG	36% ± 4%	Pro	CCG	86% ± 1%	Gln	CAG	49% ± 1%	Arg	CGG	90% ±
Ile	AUU	3% ± 2%	Thr	ACU	53% ± 2%	Asn	AAU	82% ± 2%	Ser	AGU	64% ±
Ile	AUC	14% ± 10%	Thr	ACC	44% ± 5%	Asn	AAC	79% ± 4%	Ser	AGC	72% ±
Ile	AUA	90% ± 2%	Thr	ACA	93% ± 8%	Lys	AAA	68% ± 1%	Arg	AGA	78% ±
eMet	AUG	63% ± 3%	Thr	ACG	87% ± 8%	Lys	AAG	71% ± 2%	Arg	AGG	85% ±
Val	GUU	25% ± 3%	Ala	GCU	8% ± 9%	Asp	GAU	60% ± 5%	Gly	GGU	35% ±
Val	GUC	57% ± 9%	Ala	GCC	1% ± 8%	Asp	GAC	41% ± 8%	Gly	GGC	13% ±
Val	GUA	68% ± 4%	Ala	GCA	16% ± 3%	Glu	GAA	68% ± 8%	Gly	GGA	68% ±
Val	GUG	66% ± 7%	Ala	GCG	19% ± 3%	Glu	GAG	53% ± 12%	Gly	GGG	62% ±

[0203] 表2:使用单密码子RGS模板测试tRNA耗尽效率。使用图11所示的公式,基于通过单密码子模板和各种tRNA混合物进行的tRNA耗尽的裂解物中的肽表达水平,来计算对应于其同源密码子(密码子)的特定tRNA的耗尽效率(DE.)。

[0204]

寡聚物名称	序列	修饰
寡聚 DNA1-Asp-胺	GACAGGCAGGTATTCTAACCGACTGAA	3'-胺
寡聚 DNA1-Asn-胺	TTAACAGTCCGCCGTTCTACCGACT	3'-胺
寡聚 DNA1-Glu-胺	GTGAAAGGGCGGTGTCTGGGCCTCT	3'-胺
寡聚 DNA1-Ile-胺	TTATCAGGGGTGCGCTCTAACCACCT	3'-胺
寡聚 DNA2-Asn-胺	CTCCTCTGACTGGACTCGAACCAGTGACAT	3'-胺
寡聚 DNA1-Asp-生物素	GACAGGCAGGTATTCTAACCGACTGAA	3'-生物素-TEG
寡聚 DNA1-Asp-巯基	GACAGGCAGGTATTCTAACCGACTGAA	3'-巯基
寡聚 DNA1-Glu-生物素	GTGAAAGGGCGGTGTCTGGGCCTCT	3'-生物素-TEG

[0205] 表3:用于从总天然tRNA混合物中亲和纯化特定天然tRNA的DNA寡核苷酸(寡聚DNA)。寡聚DNA1和寡聚DNA2与目标tRNA序列互补,分别跨越D臂到反密码子环或跨越受体茎到可变环。寡核苷酸的3'末端用胺、生物素或巯基修饰以促进其在基质上的固定。TEG表示三甘醇。

[0206] 表4.sGFP_T1密码子组成及其半合成tRNA套件

[0207]

aa.	密码子	tRNA	密码子编号	tRNA 比率	aa.	密码子	tRNA	密码子编号	tRNA 比率
Ala	GCU/GCC	Aggc	0	0	iMet	1AUG	Mcau1	1	0.5
	7GCG	Augc	7	3.5			Mcau2		1.5
	1GCA	Augc	1	0.5	Ile	AUA	Icau1	0	0
Arg	CGU/1CGC/CGA	Racg	1	1			Icau2		0
	6CGG	Rccg	6	6	Met	5AUG	Mcau3	5	5
	1AGG	Rccu	1	1	Phe	10UUC/2UUU	Fga	12	8
Asn	AGA	Rucu	0	0		3CCG	Pcgg	3	3
	10AAC/3AAU	Nquu*	13	6.5	Pro	3CCC	Pggg	3	3
Asp	13GAC/5GAU	Dguc	18	9		4CCA/CCU	Pugg	4	4
Cys	2UGC	Cgca	2	2	Ser	8UCG	Scga	8	8
	7CAG	Qcug	7	7		1AGC/AGU	Sgcu	1	1
Gln	CAA	Quug	0	0		1UCC/UCU	Sgga	1	1
	13GAA/3GAG	Esuc*	16	8		UCA	Suga	0	0
Glu	3GGG	Gccc	3	1.5	Thr	8ACG	Tcgu1	8	4
	2GGU/5GGC	Ggcc	7	3.5			Tcgu2		4
His	12GGA	Gucc	12	6		5ACC/ACU	Tggg1	5	2.5
	8CAC/2CAU	Hgug	10	6			Tggg2		2.5
Ile	6AUC/5AUU	Igau*	11	5.5		5ACA	Tugu	5	5
	7UUG	Lcaa	7	7	Trp	1UGG	Wcca	1	1
Leu	5CUG	Lcag1	5	5		9UAC/UAU	Ygua1	9	6
		Lcag2	0	0	Tyr		Ygua2		3
	3CUC/3CUU	Lgag	6	6		2GUC	Vgac1	1	1
	UUA	Luca	0	0			Vgac2		1
	2CUA	Luag	2	2		6GUA/2GUU	Vuac	8	8
	20AAG	Kcuu	20	12		8GUG	Vuac	8	8

[0208] 表5.sGFP_T2密码子组成及其半合成tRNA套件

[0209]

aa.	密码子	tRNA	密码子编号	tRNA 比率	aa.	密码子	tRNA	密码子编号	tRNA 比率
Ala	GCU/GCC	Aggc	0	0	iMet	1AUG	Mcau1	1	0.5
	8GCG	Augc	8	4			Mcau2		1.5
	GCA	Augc	0	0	Ile	AUA	Icau1	0	0
Arg	CGU/CGC/CGA	Racg	0	0			Icau2		0
	7CGG	Rccg	7	7	Met	5AUG	Mcau3	5	5
	1AGG	Rccu	1	1	Phe	10UUC/2UUU	Fga	12	8
Asn	AGA	Rucu	0	0		CCG	Pcgg	0	0
	10AAC/3AAU	Nquu*	13	6.5	Pro	CCC	Pggg	0	0
Asp	13GAC/5GAU	Dguc	18	9		10CCA/CCU	Pugg	10	10
Cys	2UGC	Cgca	2	2	Ser	9UCG	Scga	9	9
	7CAG	Qcug	7	7		1AGC/AGU	Sgcu	1	1
Gln	CAA	Quug	0	0		UCC/UCU	Sgga	0	0
	13GAA/3GAG	Esuc*	16	8		UCA	Suga	0	0
Glu	GGG	Gccc	0	0	Thr	ACG	Tcgu1	0	0
	GGU/GGC	Ggcc	0	0			Tcgu2		0
His	22GGA	Gucc	22	11		ACC/ACU	Tggg1	0	0
	8CAC/2CAU	Hgug	10	6			Tggg2		0
Ile	6AUC/5AUU	Igau*	11	5.5		18ACA	Tugu	18	12
	19UUG	Lcaa	19	9.5	Trp	1UGG	Wcca	1	1
Leu	CUG	Lcag1	0	0		9UAC/UAU	Ygua1	9	6
		Lcag2	0	0	Tyr		Ygua2		3
	CUC/CUU	Lgag	0	0		GUC	Vgac1	0	0
	UUA	Luca	0	0			Vgac2		0
	1CUA	Luag	1	1		GUA/GUU	Vuac	0	0
	20AAG	Kcuu	20	12		18GUG	Vuac	18	10

[0210] 表6.sGFP_T3密码子组成及其半合成tRNA套件

[0211]

aa.	密码子	tRNA	密码子编号	tRNA 比率	aa.	密码子	tRNA	密码子编号	tRNA 比率
Ala	GCU/GCC	Aggc	0	0	iMet	1AUG	Mcau1	1	0.5
	7GCG	Augc	7	3.5			Mcau2		1.5
	1GCA	Augc	1	0.5	Ile	AUA	Icau1	0	0
Arg	CGU/CGC/CGA	Racg	0	0			Icau2		0
	7CGG	Rccg	7	7	Met	5AUG	Mcau3	5	5
	1AGG	Rccu	1	1	Phe	10UUC/2UUU	Fgaa	12	8
Asn	AGA	Rucu	0	0		3CCG	Pcgg	3	3
	10AAC/3AAU	Nguu*	13	6.5	Pro	3CCC	Pggg	3	3
Asp	13GAC/5GAU	Dguc	18	9		4CCA/CCU	Pugg	4	4
Cys	2UGC	Cgca	2	2	Ser	9UCG	Scga	9	9
	7CAG	Qcug	7	7		1AGC/AGU	Sgcu	1	1
Gln	CAA	Quug	0	0		UCC/UCU	Sgga	0	0
	13GAA/3GAG	Esuc*	16	8		UCA	Suga	0	0
Glu	3GGG	Gccc	3	1.5	Thr	8ACG	Tcgu1	8	4
	2GGU/5GGC	Gccc	7	3.5		5ACC/ACU	Tcgu2		4
Gly	12GGA	Gucc	12	6			Tggg1	5	2.5
	8CAC/2CAU	Hgug	10	6			Tggg2		2.5
His	6AUC/5AUU	Igau*	11	5.5		5ACA	Tugu	5	5
	7UUG	Lcaa	7	7	Trp	1UGG	Wcca	1	1
Leu	5CUG	Lcag1	5	5		9UAC/UAU	Ygua1	9	6
		Lcag2	0	0	Val	2GUC	Vgac1	1	1
	3CUC/3CUU	Lgag	6	6			Vgac2	1	1
	UUA	Luaa	0	0		6GUA/2GUU	Vuac	6	8
	2CUA	Luag	2	2		8GUG	Vuac	8	8
Lys	20AAG	Kcuu	20	12					

[0212] 表4-6. 三种sGFP模板的密码子组合物及其相应的半合成tRNA套件。表中显示对于每个密码子表现最好的t7tRNA(对于名称见图8)以及它在半合成tRNA套件中的比率(%)。通过在Kuuu反密码子中的U34至C34置换构建AAG密码子Kcuu的合成tRNA^{Lys}。Asn、Glu和Ile的特定天然tRNA(而不是非功能性t7tRNA对应物)以星号标记,用于半合成tRNA混合物。

氨基酸	"正交密码子"	"天然密码子"
Ser	AGC	UCU, UCC, UCA, UCG
	UCG	AGU, AGC, UCU, UCC, UCA
Arg	AGG	CGU, CGC, CGA, CGG
	CGG	AGA, AGG
Leu	UUG	CUU, CUC, CUA, CUG
	CUC	CUG, CUA, UUA, UUG
Pro	CCC	CCG
Thr	ACG	ACC
Gly	GGG	GGC

[0214] 表7. 从特定的天然tRNA缺失水平、t7tRNA的解码效率和它们对于分割和非分割密码子框的交叉识别倾向性的系统分析得到的潜在的“正交”与“天然”密码子对。最有希望的“正交”与“天然”密码子对包括来自混合密码子家族框(对于Ser、Arg和Leu)的密码子以及来自具有高摆动限制的非分割密码子家族框(对于Pro、Thr和Gly)的密码子。

[0215] 参考文献

- [0216] (1) O'Donoghue, P.; Ling, J.; Wang, Y.-S.; Söll, D. *Nat Chem Biol* 2013, 9, 594.
- [0217] (2) Liu, C.C.; Schultz, P.G. *Annu Rev Biochem* 2010, 79, 413.
- [0218] (3) Hong, S.H.; Ntai, I.; Haimovich, A.D.; Kelleher, N.L.; Isaacs, F.J.; Jewett, M.C. *ACS Synth Biol* 2014, 3, 398.
- [0219] (4) Wang, Y.S.; Fang, X.Q.; Chen, H.Y.; Wu, B.; Wang, Z.Y.U.; Hilty, C.; Liu, W.S.R. *Acs Chem Biol* 2013, 8, 405.
- [0220] (5) Albayrak, C.; Swartz, J.R. *Nucleic Acids Res* 2013, 41, 5949.
- [0221] (6) Kim, J.; Seo, M.H.; Lee, S.; Cho, K.; Yang, A.; Woo, K.; Kim, H.S.; Park, H.S. *Anal Chem* 2013, 85, 1468.
- [0222] (7) Wu, B.; Wang, Z.Y.; Huang, Y.; Liu, W.S.R. *Chembiochem* 2012, 13, 1405.
- [0223] (8) Wang, K.H.; Sachdeva, A.; Cox, D.J.; Wilf, N.W.; Lang, K.; Wallace, S.; Mehl, R.A.; Chin, J.W. *Nat Chem* 2014, 6, 393.
- [0224] (9) Chen, S.; Fahmi, N.E.; Wang, L.; Bhattacharya, C.; Benkovic, S.J.; Hecht, S.M. *J Am Chem Soc* 2013, 135, 12924.
- [0225] (10) Frankel, A.; Roberts, R.W. *RNA* 2003, 9, 780.
- [0226] (11) Kwon, I.; Kirshenbaum, K.; Tirrell, D.A. *J Am Chem Soc* 2003, 125, 7512.
- [0227] (12) Link, A.J.; Tirrell, D.A. *Methods* 2005, 36, 291.
- [0228] (13) Loscha, K.V.; Herlt, A.J.; Qi, R.H.; Huber, T.; Ozawa, K.; Otting, G. *Angew Chem Int Edit* 2012, 51, 2243.
- [0229] (14) Odoi, K.A.; Huang, Y.; Rezenom, Y.H.; Liu, W.R. *Plos One* 2013, 8, e57035.
- [0230] (15) Ma, C.; Kudlicki, W.; Odom, O.W.; Kramer, G.; Hardesty, B. *Biochemistry* 1993, 32, 7939.
- [0231] (16) Forster, A.C.; Tan, Z.P.; Nalam, M.N.L.; Lin, H.N.; Qu, H.; Cornish, V.W.; Blacklow, S.C. *P Natl Acad Sci USA* 2003, 100, 6353.
- [0232] (17) Hipolito, C.J.; Suga, H. *Curr Opin Chem Biol* 2012, 16, 196.
- [0233] (18) Grosjean, H.; de Crecy-Lagard, V.; Marck, C. *Febs Lett* 2010, 584, 252.
- [0234] (19) El Yacoubi, B.; Bailly, M.; de Crecy-Lagard, V. *Annu Rev Genet* 2012, 46, 69.
- [0235] (20) Jackman, J.E.; Alfonzo, J.D. *Wiley Interdiscip Rev RNA* 2013, 4, 35.
- [0236] (21) Giege, R.; Sissler, M.; Florentz, C. *Nucleic Acids Res* 1998, 26, 5017.
- [0237] (22) Amikura, K.; Sakai, Y.; Asami, S.; Kiga, D. *ACS Synth Biol* 2014, 3, 140.
- [0238] (23) Kawahara-Kobayashi, A.; Masuda, A.; Arais, Y.; Sakai, Y.; Kohda, A.; Uchiyama, M.; Asami, S.; Matsuda, T.; Ishitani, R.; Dohmae, N.; Yokoyama, S.; Kigawa, T.; Nureki, O.; Kiga, D. *Nucleic Acids Res* 2012, 40, 10576.
- [0239] (24) Takai, K.; Takaku, H.; Yokoyama, S. *Nucleic Acids Res* 1996, 24, 2894.
- [0240] (25) Takai, K.; Takaku, H.; Yokoyama, S. *Biochem Biophys Res Commun* 1999, 257, 662.
- [0241] (26) Claesson, C.; Samuelsson, T.; Lustig, F.; Boren, T. *Febs Lett* 1990, 273, 173.

- [0242] (27) Schwarz,D.;Junge,F.;Durst,F.;Frolich,N.;Schneider,B.;Reckel,S.;Sobhanifar,S.;Dotsch,V.;Bernhard,F.Nat Protoc 2007,2,2945.
- [0243] (28) Jackson,R.J.;Naphthine,S.;Brierley,I.RNA 2001,7,765.
- [0244] (29) Chan,P.P.;Lowe,T.M.Nucleic Acids Res 2009,37,D93.
- [0245] (30) Beckert,B.;Masquida,B.Methods Mol Biol 2011,703,29.
- [0246] (31) Fechter,P.;Rudinger,J.;Giege,R.;Theobald-Dietrich,A.Febs Lett 1998,436,99.
- [0247] (32) Stein,V.;Alexandrov,K.P Natl Acad Sci USA 2014,111,15934.
- [0248] (33) Huang,J.;Makabe,K.;Biancalana,M.;Koide,A.;Koide,S.J Mol Biol 2009,392,1221.
- [0249] (34) Huang,J.;Koide,A.;Makabe,K.;Koide,S.P Natl Acad Sci USA 2008,105,6578.
- [0250] (35) Huang,J.;Nagy,S.S.;Koide,A.;Rock,R.S.;Koide,S.Biochemistry 2009,48,11834.
- [0251] (36) Dong,H.J.;Nilsson,L.;Kurland,C.G.J Mol Biol 1996,260,649.
- [0252] (37) Rosenberg,A.H.;Goldman,E.;Dunn,J.J.;Studier,F.W.;Zubay,G.J Bacteriol 1993,175,716.
- [0253] (38) Samuelsson,T.;Boren,T.;Johansen,T.I.;Lustig,F.J Biol Chem 1988,263,13692.
- [0254] (39) Spenkuch,F.;Motorin,Y.;Helm,M.RNA Biol 2014,11,1540.
- [0255] (40) Kruger,M.K.;Pedersen,S.;Hagervall,T.G.;Sorensen,M.A.J Mol Biol 1998,284,621.
- [0256] (41) Himeno,H.;Hasegawa,T.;Ueda,T.;Watanabe,K.;Miura,K.;Shimizu,M.Nucleic Acids Res 1989,17,7855.
- [0257] (42) Kirino,Y.;Yasukawa,T.;Ohta,S.;Akira,S.;Ishihara,K.;Watanabe,K.;Suzuki,T.Proc Natl Acad Sci U S A 2004,101,15070.
- [0258] (43) Agris,P.F.;Vendeix,F.A.;Graham,W.D.J Mol Biol 2007,366,1.
- [0259] (44) Nureki,O.;Niimi,T.;Muramatsu,T.;Kanno,H.;Kohno,T.;Florentz,C.;Giege,R.;Yokoyama,S.J Mol Biol 1994,236,710.
- [0260] (45) Sylvers,L.A.;Rogers,K.C.;Shimizu,M.;Ohtsuka,E.;Soll,D.Biochemistry 1993,32,3836.
- [0261] (46) Shimizu,M.;Asahara,H.;Tamura,K.;Hasegawa,T.;Himeno,H.J Mol Evol 1992,35,436.
- [0262] (47) Stuart,J.W.;Gdaniec,Z.;Guenther,R.;Marszalek,M.;Sochacka,E.;Malkiewicz,A.;Agris,P.F.Biochemistry 2000,39,13396.
- [0263] (48) Sundaram,M.;Durant,P.C.;Davis,D.R.Biochemistry 2000,39,15652.
- [0264] (49) Serra,M.J.;Turner,D.H.Methods Enzymol 1995,259,242.
- [0265] (50) Agris,P.F.Nucleic Acids Res 2004,32,223.
- [0266] (51) Yokoyama,S.;Watanabe,T.;Murao,K.;Ishikura,H.;Yamaizumi,Z.;

- Nishimura, S.; Miyazawa, T. *Proc Natl Acad Sci U S A* 1985, 82, 4905.
- [0267] (52) Yarian, C.; Townsend, H.; Czestkowski, W.; Sochacka, E.; Malkiewicz, A.J.; Guenther, R.; Miskiewicz, A.; Agris, P.F. *J Biol Chem* 2002, 277, 16391.
- [0268] (53) Wohlgemuth, I.; Pohl, C.; Mittelstaet, J.; Konevega, A.L.; Rodnina, M.V. *Philos T R Soc B* 2011, 366, 2979.
- [0269] (54) Kothe, U.; Rodnina, M.V. *Mol Cell* 2007, 25, 167.
- [0270] (55) Sprinzl, M.; Horn, C.; Brown, M.; Ioudovitch, A.; Steinberg, S. *Nucleic Acids Res* 1998, 26, 148.
- [0271] (56) Cantara, W.A.; Bilbille, Y.; Kim, J.; Kaiser, R.; Leszczynska, G.; Malkiewicz, A.; Agris, P.F. *J Mol Biol* 2012, 416, 579.
- [0272] (57) Miyauchi, K.; Ohara, T.; Suzuki, T. *Nucleic Acids Res* 2007, 35, e24.
- [0273] (58) Yokogawa, T.; Kitamura, Y.; Nakamura, D.; Ohno, S.; Nishikawa, K. *Nucleic Acids Res* 2010, 38, e89.
- [0274] (59) Pedelacq, J.D.; Cabantous, S.; Tran, T.; Terwilliger, T.C.; Waldo, G.S. *Nat Biotechnol* 2006, 24, 1170.
- [0275] (60) Godinic-Mikulcic, V.; Jaric, J.; Greber, B.J.; Franke, V.; Hodnik, V.; Anderluh, G.; Ban, N.; Weygand-Durasevic, I. *Nucleic Acids Res* 2014, 42, 5191.
- [0276] (61) Shimizu, Y.; Ueda, T. *Methods Mol Biol* 2010, 607, 11.
- [0277] (62) Tamura, K.; Himeno, H.; Asahara, H.; Hasegawa, T.; Shimizu, M. *Nucleic Acids Res* 1992, 20, 2335.
- [0278] (63) Ogle, J.M.; Murphy, F.V.; Tarry, M.J.; Ramakrishnan, V. *Cell* 2002, 111, 721.
- [0279] (64) Forster, A.C.; Church, G.M. *Mol Syst Biol* 2006, 2, 45.
- [0280] (65) Goto, Y.; Katoh, T.; Suga, H. *Nat Protoc* 2011, 6, 779.
- [0281] (66) Fechter, P.; Rudinger, J.; Giege, R.; Theobald-Dietrich, A. *Febs Lett* 1998, 436, 99.
- [0282] (67) Sampson, J.R.; Uhlenbeck, O.C. *Proc Natl Acad Sci U S A* 1988, 85, 1033.
- [0283] (68) Noren, C.J.; Anthony-Cahill, S.J.; Suich, D.J.; Noren, K.A.; Griffith, M.C.; Schultz, P.G. *Nucleic Acids Res* 1990, 18, 83.
- [0284] (69) Deana, A.; Celesnik, H.; Belasco, J.G. *Nature* 2008, 451, 355.
- [0285] (70) Chan, P.P.; Lowe, T.M. *Nucleic Acids Res* 2009, 37, D93.
- [0286] (71) Czerwoniec, A.; Dunin-Horkawicz, S.; Purta, E.; Kaminska, K.H.; Kasprzak, J.M.; Bujnicki, J.M.; Grosjean, H.; Rother, K. *Nucleic Acids Res* 2009, 37, D118.
- [0287] (72) Bentele, K.; Saffert, P.; Rauscher, R.; Ignatova, Z.; Bluthgen, N. *Mol Syst Biol* 2013, 9, 675.
- [0288] (73) Roberts, R.G. *Plos Biol* 2013, 11, e1001509.
- [0289] (74) Miyauchi, K.; Ohara, T.; Suzuki, T. *Nucleic Acids Res* 2007, 35, e24.
- [0290] (75) Yokogawa, T.; Kitamura, Y.; Nakamura, D.; Ohno, S.; Nishikawa, K. *Nucleic Acids Res* 2010, 38.

[0291] 实施例7

[0292] 通过有义密码子和琥珀抑制密码子进行双重蛋白质标记

[0293] 材料和方法

[0294] 材料

[0295] 非天然氨基酸、BCN(双环[6.1.0]壬炔-赖氨酸)、TCO*(反式-环辛-2-烯-L-赖氨酸)、TCO(反式-环辛烯-赖氨酸)、PrK(N-炔丙基-赖氨酸)购自Sirius Fine Chemicals SiChem GmbH。AzF(对叠氮基苯丙氨酸)购自SynChem。TAMRA DIBO炔烃购自Life Technologies Australia Pty Ltd。BODIPY®FL碘代乙酰胺来自Molecular Probes®。乙醇胺-琼脂糖是通过将乙醇胺偶联在环氧化物活化的Sephacrose 6B(GE Healthcare)上制备的。用水(~200ml每克基质)充分洗涤环氧化物活化的琼脂糖6B以除去添加剂,然后在室温下用1M乙醇胺温育,温和搅拌过夜。进行另一次水洗直至pH接近7。将得到的乙醇胺-琼脂糖基质以含2mM NaN₃的缓冲液A(100mM KOAc、25mM NaOAc(pH5.2)、0.25mM EDTA)中以20%浆液(slurry)储存。8克基质产生100ml 20%浆液。

[0296] 正交tRNA合成酶的克隆与纯化

[0297] AzFRS的蛋白质序列来源于之前报道的突变体7³⁹。Py1RS变体MbBCNRS、MbTCORS36和MmPy1RSAF的蛋白质序列从各自的文献中获得^{19,29}。通过IDT合成在N端含有6个His标签的这些蛋白质的gBlock编码密码子优化的ORF,并亚克隆到NcoI-NotI双重消化的pOPINE质粒中。消化的质粒和所需蛋白质基因的组装通过Gibson Assembly® Master Mix(NEB)实现。通过测序鉴定正确的构建体。在Rosetta细胞中用0.5mM异丙基-β-D-硫代吡喃半乳糖苷(IPTG)在20℃诱导蛋白表达过夜。使用标准缓冲液通过Ni²⁺亲和层析纯化AzFRS,然后在PBS缓冲液中在Superdex 200(GE Healthcare)上进行凝胶过滤。使用改良的结合缓冲液(50mM磷酸钠,pH7.4,0.5M NaCl,5%甘油(体积/体积),10mM MgCl₂,0.1%吐温,1mM DTT,20mM咪唑)和洗脱缓冲液(结合缓冲液加0.5M咪唑)通过Ni²⁺亲和层析纯化MbBCNRS和MbTCORS。通过凝胶过滤进一步纯化蛋白质,并在-80℃保存于40mM Hepes pH7.4、150mM NaCl、10mM MgCl₂、10%(体积/体积)甘油、0.5mM TCEP中。

[0298] T7tRNA合成

[0299] MjY1、MjY2和MjY3tRNA在一份早期报道中被鉴定为用于AzF并入的最佳三种tRNA³⁷,而具有六个突变的MjY4tRNA描述于后来的报道³⁸。这四个tRNA的序列比对显示在SI中(图27)。优化的Py1T tRNA是U25C突变体³⁸。除了将“GUC”替换为“CUA”反密码子之外,tRNATyr_{cua}的序列与野生型tRNATyr的序列相似。所有上述tRNA的t7转录本如之前报道的那样制备²⁴。基本上,使用3步PCR法由5或6个寡核苷酸(表8)组装每个DNA模板,包括正常模板或转化酶构建体。然后通过乙醇沉淀法纯化PCR产物并将其用于T7RNA聚合酶介导的run-off转录。转录的tRNA通过使用乙醇胺-琼脂糖基质的亲和层析来纯化。具体而言,使用10倍储存液将2ml转录反应物调整为缓冲液A条件。加入与转录反应体积相同体积的20%的基质,并在4℃温育过夜。用10ml缓冲液A洗涤基质4次后,用2ml缓冲液B(2M NaOAc(pH5.2)、0.25mM EDTA、2.5mM Mg(OAc)₂)洗涤tRNA两次,然后用乙醇沉淀。

[0300] 总tRNA制备

[0301] 通过改进的Zubay方法从大肠杆菌(金)菌株中分离总tRNA混合物⁵⁹。简言之,总tRNA是从过夜培养物中提取的。将细胞沉淀(20g)重悬于40ml 10mM Mg(OAc)₂、1mM Tris

pH7.4缓冲液中。然后用34.4ml液化苯酚提取核苷酸,在冷室中剧烈搅拌2小时。通过加入10ml液化苯酚再次重复萃取。然后通过18,000g离心30分钟收集水相。在-20℃用0.05体积的4M乙酸钾和2体积的100%乙醇沉淀RNA过夜。通过5,000g离心10分钟收集沉淀并重悬于20ml 1M冷NaCl中。在冷室中剧烈搅拌1小时以分散沉淀。收集上清液,用1M NaCl重复提取。将两次NaCl提取的上清液合并,加入2体积乙醇来沉淀。重复两次之后,将粗制的tRNA部分溶于水。

[0302] 总tRNA混合物中的特定tRNA的耗尽

[0303] 按文献所述²⁴,通过DNA杂交色谱从总混合物中耗尽特定的tRNA种类。设计并使用来自IDT的3'-胺基团(表8)合成与反密码子环至受体茎片段互补的DNA寡核苷酸。根据制造商的方案,将这些DNA寡核苷酸固定在NHS-活化的琼脂糖凝胶(GE)上,每100μl树脂使用100μl的150μM寡核苷酸。将200μl的5μg/μl未分离tRNA混合物与相同体积的2×杂交缓冲液(20mM Tris-HCl (pH7.6)、1.8M NaCl、0.2mM EDTA)混合,然后与基质上的寡核苷酸杂交。对于每种目标tRNA,使用125μl固定化了寡聚DNA的沉降树脂。将tRNA混合物和基质的悬浮液在65℃下热变性10分钟。在室温下缓慢冷却样品后,通过在略低于400μl的离心条件下收集未杂交的tRNA。然后使用40μl的1x杂交缓冲液洗涤基质并与从最后一步收集的tRNA混合。将所有未结合的tRNA合并在一起进行乙醇沉淀并溶于水。

[0304] tRNA氨酰化及荧光团结合

[0305] 如上所述,通过t7转录制备具有CCU或CUA反密码子的合成半胱氨酸tRNA。通过重组半胱氨酰-tRNA合成酶(CysRS, SI)在这些tRNA变体上负载半胱氨酸(Cys)或硒代半胱氨酸(Secys)。首先,将50mM Cys与50mM TCEP在37℃下预孵育15min,以在氨酰化反应之前将其转化为完全还原形式。硒代胱氨酸(Secis)在含有40mM Secis、50mM HepesKOH (pH8.0)、1M DTT的溶液(用100mM KOH使终点pH达到7.5)中于37℃还原30min成为Secys。同时,tRNA在78℃变性,并在5mM Mg²⁺存在下以50μM浓度室温缓慢重新折叠以保持最大活性。然后在100mM Hepes-KOH (pH8.0)、2mM Cys或Secys、2mM TCEP或25mM DTT (Secys)、10mM MgCl₂、5mM KCl、5mM KOAc、0.1mM CTP、4mM ATP、22.5μM t7tRNACys变体、10μMCysRS、20%DMSO、5μM ZnCl₂、0.005u/ml酵母无机焦磷酸酶和50μg/ml BSA中在37℃孵育1h来进行氨酰化反应。

[0306] 氨酰化后,将反应混合物稀释4倍至1×缓冲液A中,用于乙醇胺-琼脂糖凝胶纯化,接着进行苯酚提取(通过用1V氯仿/异戊醇(24:1)进行2次连续提取而出去残留苯酚)和乙醇沉淀。在50mM Tris-HCl (对于Secys-tRNA为pH8.5或7.2)、80μM Cys-tRNA、75%DMSO、1mM BPFL-1A和100mM NaCl中进行与BODIPY®FL碘乙酰胺(BPFL-1A)的偶联反应。使用乙醇使偶联产物(BPFL-tRNA)沉淀并溶解在最小体积的tRNA缓冲液(0.5mM MgCl₂、0.5mM NaOAc pH5.0)中。然后将样品调节至1×HPLC缓冲液(0.1M TEAA\1%ACN),然后使用以下缓冲液在POROS®110μm柱(Applied Biosystems)上通过HPLC纯化:缓冲液C为0.1M TEAA、1%ACN;缓冲液D是0.1M TEAA、90%ACN,以1-20%线性梯度运行9min。HPLC纯化后,用乙醇沉淀BPFL-tRNA部分,重悬于tRNA缓冲液中并保存于-80℃。

[0307] 体外蛋白质翻译试验

[0308] 按文献所述通过乙醇胺-琼脂糖亲和层析制备tRNA耗尽的裂解物²⁴。使用具有两个GFPORF的DNA模板:一个具有eGFP的优化密码子组成(模板A)和另一个具有sGFP的简化密码子组成(模板B)。通过IDT合成全部基因片段作为gBlock,将所有6个Arg密码子同时改变为

AGG,或使用适当的引物进行PCR以在相同位置引入少于3个核苷酸的突变,来引入对琥珀密码子或AGG密码子的突变。按照标准Gibson组装克隆程序(NEB),将获得的片段克隆到基于pLTE或pOPINE的质粒中。通过Midiprep试剂盒(Qiagen)纯化具有所需突变的正确质粒,以确保无细胞蛋白质翻译的高质量。SI中提供了本文使用的所有开放阅读框的序列信息。

[0309] 用于GFP生产的无细胞翻译反应按照标准方案⁶⁰进行,优化的 Mg^{2+} 浓度为10mM。在反应中使用0.35体积的耗尽的裂解物(标准S30裂解物相同)。每个实验都记录了补充其他化合物。在荧光板酶标仪(Synergy)上于30℃在485nm激发波长和528nm发射波长下监测GFP的产生3-5小时。

[0310] 蛋白质标记

[0311] 具有AzF和BP-FL的CaM在无细胞翻译反应物中的表达:将pOPINE-CaM模板(在CaMORF的位置1和149处分别具有TAG和AGG密码子;图27;SI)以100ng/ μ l加入无细胞反应物中,其含有35%(v/v)耗尽的裂解物、0.5 μ g/ μ l耗尽的tRNA混合物、10 μ M MjY2tRNA、10 μ M AzFRS、1mM AzF;10 μ M BP-tRNACysccu。反应在32℃下进行4小时。将200 μ l无细胞反应物在12-14kDa截留透析管中以600ml PBS缓冲液透析3小时,进行两次以除去翻译反应中的AzF。透析后,反应体积增加至250 μ l。在反应物中加入16 μ l 1mM DIBO-TAMRA,使最终浓度达到60 μ M。标记反应在室温下进行3小时。

[0312] 标记的CaM在亲和钳基质上的纯化:在以5000rpm离心反应混合物2分钟后,将20 μ l 50%亲和钳基质(参见SI)加入到标记样品中。基质在室温下与样品温育30分钟。将基质用1.5mlPBS-Triton(0.1%)缓冲液洗涤,然后用7.5ml PBS洗涤,最后用3.2ml剪切缓冲液(50mM Tris-HCl pH8、150mM NaCl、1mM DTT)洗涤。随后将40 μ l Precision蛋白酶(0.5mg/ml)加入到基质中并在16℃温育过夜。用2 $\times$ 75 μ l剪切缓冲液洗脱蛋白质。

[0313] 进行两个单独的反应以制备单标记的CaM蛋白质。为了表达CaM-BPFL,将负载了BPFL的tRNAccu补充到缺乏天然AGG tRNA种类的无细胞翻译系统中以解码AGG密码子,同时添加10 μ M tRNATyrcua以将UAG密码子解码成Tyr。为了获得CaM-TAMRA,使用MjY2通过琥珀抑制将AzF并入到蛋白质序列中,并将AzFRS并入UAG密码子的第一位置,同时补充5 μ M合成tRNAArgccu以将AGG密码子解码为Arg。然后,AzF通过如上所述的无铜点击化学由DIBO-TAMRA修饰。

[0314] smFRET实验

[0315] 在50mM Tris-HCl (pH8)、150mM NaCl、1mM DTT中进行smFRET实验,其中 $\sim$ 100pM双标记的CaM浓度。在smFRET实验中记录是否存在2mM Ca^{2+} 或10mM EDTA。对于每个实验,将20 μ l样品上样在由70 $\times$ 80mm玻璃盖保持的自制硅酮板(SYLGARD®)上。在装有ConfoCor3模块的Zeiss LSM 710显微镜上使用Apochromat 40 $\times$ 1.2NA水浸物镜(Zeiss)在室温下进行smFRET测量。用488nm激光激发BP-FL荧光。发射光穿过针孔,并使用565nm二向色镜分裂成供体组分和受体成分;供体和受体信号分别通过505-540和580-610nm带通滤波器进一步过滤。

[0316] 在每次实验中,供体和受体信号同时记录30s,重复20次,仓时间(bin time)为1ms。将来自3个实验的数据组合在一起以获得足够的事件以用于分析。在20 μ M荧光团浓度的单独实验中估计供体发射光到游离TAMRA染料的受体通道的泄漏率,并计算为14%以在FRET分析之前校正信号。对于供体和受体信号的总和,将阈值设定为20个计数,以过滤背景

噪声。FRET效率计算为[强度(受体)-来自供体通道的泄漏]/总强度(强度(受体)+强度(供体))并绘制为直方图。通过Origin软件(OriginLab Corp.)使用高斯函数来拟合数据。

[0317] 结果

[0318] 创建用于nnAA并入的“AGG”空白密码子

[0319] 我们最近证实,大肠杆菌裂解物中AGG密码子的天然精氨酸同工受体能够几乎被完全耗尽²⁴。我们证明这种裂解物不能支持含有单个AGG密码子的sfGFP ORF的翻译,而加入合成的tRNA^{Accu}可以恢复其翻译。原先方法的主要缺点是需要从各种tRNA配制半合成tRNA混合物。因此,我们决定测试一种替代方法,通过DNA杂交层析从总天然tRNA混合物中选择性地耗尽目标tRNA²⁵。

[0320] 为此,从BL21(金)大肠杆菌菌株中分离出总tRNA或商购获得总tRNA。与tRNA^{Accu}/ucu的核苷酸互补的DNA寡核苷酸被用于从总tRNA混合物中实现色谱法耗尽。通过是否支持具有一个或六个AGG密码子的GFP模板的翻译来测试tRNA^{Accu}/ucu耗尽前后tRNA混合物的翻译效率(图21和图27)。在前一种情况下,AGG解码同工受体耗尽的总tRNA混合物中剩余的精氨酸密码子导致对于1AGG模板仅观察到可忽略的背景翻译,对于6AGG模板没有可检测的翻译(图21和图27)。向反应物中补充t7tRNA^{Accu}使翻译的水平恢复至相当于由总tRNA混合物耗尽前提供的翻译水平。我们将tRNA耗尽效率定义为 $1 - (\text{RFU}(\text{Depl. tRNA}) / \text{RFU}(\text{Depl. tRNA} + \text{t7tRNA}^{\text{Accu}}))$ 。使用这个评估标准,自制的tRNA混合物的耗尽效率在89%到94%之间,而对于商业购买的tRNA混合物,耗尽效率为62%至88%,分别基于1AGG模板和6AGG模板计算。自制tRNA混合物用于所有后续实验。

[0321] 比较用于nnAA并入的o-tRNA/aaRS对的效率

[0322] 在证明AGG密码子可用于重新分配后,我们想要鉴定最佳的tRNA负载系统以及适用于AGG重新分配和双重蛋白质标记的nnAAs。因此,我们决定在体外翻译反应中测试可用的o-tRNA/aaRS对基于抑制琥珀密码子的能力,以寻找最佳的tRNA负载系统以及适合未来与AGG重新分配以及双重蛋白质标记目的的nnAAs。最近开发的应变(strained)烯烃/炔烃和四嗪之间的生物正交反应,被称为反向电子需求Diels-Alder环加成反应,其速率常数比这个反应的早期版本高几个数量级²⁶。其中一种生物正交反应物四嗪是发色团淬灭剂,其衍生的BODIPY探针与蛋白质结合后显示超过1000倍的开启荧光(turn-on fluorescence)²⁷,这使得未反应染料的后续去除比常规染料更容易且要求更低。因此,应变烯烃/炔烃的基因整合将是用四嗪荧光团实现快速、选择性和“开启”蛋白质标记的理想策略。Py1RS的活性位点的可塑性使得能够将大量的nnAA并入蛋白中^{19,28-35},包括应变的炔和烯(BCN、TCO或TCO*) (图22A)。已经报道了三种演变的Py1RS变体BCNRS36、TCORS36和Py1RSAF19,其具有许多独特的突变,以介导这些nnAA在体内的并入。我们着手比较这些酶在体外翻译系统对几种底物的效率。

[0323] 我们测试了重组形式的Py1RS变体支持由在151位含有琥珀密码子的模板(A_{151X}, S1)体外合成eGFP的能力,所述合成在存在其反密码子茎中带有U25C突变的进化o-tRNA(Py1T)¹⁴和nnAA底物BCN、TCO或TCO*的情况下进行。另一种也用于叠氮化物偶联的nnAA PrK(图22A)也被测试,已知其可入位并通过wt MbPy1RS30负载到tRNA中,并且ORF属于CGN密码子家族框MmPy1RS5。令人惊讶的是,BCN、TCO和TCO*抑制UAG的效率非常低(图22B)。三个Py1RS突变体的任何一个都不能以高于5%的效率并入这些大的脂肪族炔烃或烯烃。

BCNRS显示BCN和TCO*为3%和2%时抑制效率最高。Py1RSAF与3%的TCO结合显示出最好的抑制效率。令人惊讶的是,虽然BCNRS和TCORS基本上无法并入PrK,但是Py1RSAF以13%的效率负载它,尽管它为大脂环族基团而设计的扩大的氨基酸结合口袋证实了早期的报道²⁹;后者证实了工程化Py1RS变体对综合研究中报道³⁵的各种底物的混杂性(promiscuity)。

[0324] 来自M. jannaschii的TyrRS是另一种广泛使用的正交tRNA负载酶,其被工程化而介导在一个或多个位点并入包括四嗪衍生物在内的苄基侧链类似物^{2,13,22}。基于它们在体内和体外对琥珀密码子抑制的成功应用的报道,我们选择了p-叠氮基-L-苯丙氨酸(AzF,图22A)和MjTyrRS-AzFRS39的相应工程化形式。在测试的底物中,AzF在与flexizyme(eFx)负载系统组合使用时显示最佳的琥珀密码子并入效率(图29)。这间接地不仅指示酰化率高,还指示酰化酪氨酸类似物与EF-Tu的相容性。后者作为nnAA并入的限速步骤的重要性经常被低估。在体内选择试验中发现了AzFRS的几种优化的o-tRNA,因此我们决定在体外比较它们的活性。我们发现MjY2以81%的最高效率支持AzF并入(图22C)。它比MjY4略好,比MjY1和MjY3分别高1.8和4.6倍。尽管MjY1在选择实验中表现得最好³⁷,但与我们的试验中效率比MjY2低。

[0325] 在单个表征之后,我们系统地比较了MjY2/AzFRS/AzF和Py1T/Py1RSAF/PrK。首先,将o-tRNA和aaRS滴定到翻译反应中,以确定其浓度对由在151位具有琥珀密码子(A₁₅₁X)的模板进行eGFP表达的产量的影响。当tRNA浓度保持恒定在20μM时,GFP产量随着AzFRS浓度增加而增加(直至10μM),而Py1RS即使在30μM也不饱和(图22D)。当AzFRS或Py1RS的浓度分别保持恒定在20μM和30μM时,蛋白质产量在5-10μM的MjY2tRNA或20μM的Py1TtRNA时达到最高峰,并且在更高浓度时两者下降(图23A)。支持半最大蛋白质产量的o-tRNA浓度对于MjY2tRNA和Py1T分别是1.5μM和8μM。使用20μM MjY2tRNA和10μM AzFRS的MjY系统的最大产量达到8570RFU,而30μM的Py1RSAF和Py1T仅能够支持GFP翻译达到MjTyrRS系统达到的水平的43%。

[0326] 由于琥珀密码子的抑制效率是背景依赖性的(context dependence),所以我们设计了三个额外的编码GFP的ORF,其具有位于不同位置的单个TAG密码子,并比较了两个抑制系统之间的GFP产量(图22D)。使用缺乏nnAA-底物、o-tRNA或酶的反应来评估非特异性抑制的水平(图30;SI)。使用具有嫁接的CUA反密码子的大肠杆菌T7转录的酪氨酸tRNA(EctRNA^{Tyr}(cua),以其有效的琥珀密码子抑制而闻名)用作阳性对照。由o系统(o-systems)提供的相对活性被计算为阳性对照中荧光的百分比(图23E)。

[0327] 如图22E所示,无论模板背景(template context)如何,MjY系统在TAG密码子抑制中始终优于Py1T系统。对于模板亚组A,两个系统中的相对抑制效率对于近端TAG密码子(位置1)来说始终较高。令人惊讶的是,对于B亚组,位置偏好被逆转,与近端琥珀密码子相比其更偏好远端(位置151)琥珀密码子,其对于Py1RS o-系统来说表明完全没有抑制(图23B)。总体而言,MjY o-系统的平均活性比Py1RS高1.2-3倍(不考虑Py1RS系统在某些情况下的失效)并且具有更小的背景依赖性(图22E)。

[0328] 将AGG密码子重新分配至AzF

[0329] 我们之前的分析结果表明,正交的MjY tRNA负载o-系统介导在琥珀密码子位的高水平的AzF并入。MjTyrRS缺失缺乏一个主要的反密码子结合域,表明反密码子三联体是不是有效tRNA识别的必要条件⁴⁰。然而,据报道将第一反密码子位置(琥珀抑制剂)的G转变成C

或者突变所有三个反密码子核苷酸(ATG-抑制剂)分别导致氨酰化效率损失约100倍和900倍⁴⁰。由于基于Mj o-tRNA构建的琥珀抑制剂提供非常高的抑制效率,我们决定改变其反密码子来测试其相应的AGG密码子抑制剂,通过改变反密码子至CCU从而将其靶向AGG密码子,同时这可能对在负载反应中的o-tRNA/酶浓度提出了更高的要求。

[0330] 我们着手通过将缺乏AGG密码子同工受体的裂解物与MjY2(ccu)/AzFRS和AzF结合来评估重新分配AGG密码子到AzF的可能性。为此,我们在密码子偏好的模板亚组A和B(SI)中将单个琥珀密码子转换成AGG,而ORF中剩余的精氨酸由CGN家族密码子家族编码,以同时监测背景效应(序列信息参考图23和SI)。A_{151R}、A_{1R}、B_{151R}、B_{1R}的天然AGG同工受体的耗尽程度分别为75%、80%、93%和84%(等式1)。对于151R和1R,向反应物中补充MjY2(ccu)/AzFRS导致A亚组的荧光增加4.2和5.2倍,而B亚组增加18.4和7.2倍。这表明在所有情况下超过80%的GFP蛋白被AzF修饰。翻译产物的质谱分析表明,大部分蛋白质在AGG限定的位置具有AzF,而在该位点具有非常小的含Arg成分(图32B-C)。令人惊奇的是,即使GUA=>CCU转化导致在每个位置产生与wt不同的反密码子,并且伴随着U至C的改变导致可能氢键损失^{35,41}。AGG抑制剂在此如其琥珀抑制剂对应物一样保持高抑制活性。

[0331] Bodipy-FL nnAA衍生物的遗传编码

[0332] 通过MjY o-系统将AGG成功地重新分配至AzF为我们提供了第二个正交密码子,该密码子一旦可以与UAG组合就可能用于nnAA的双重两位点并入。然而,与AzF化学上正交的第二个nnAA的选择是有问题的,因为应变炔烃和烯烃的并入效率低并且作为替代点击化学标签的PrK和AzF之间存在交叉反应性。因此,我们选择测试一种新的策略,其结合了荧光nnAA的直接并入和点击化学相容的nnAA-标签以进一步与形成FRET的荧光团偶联。这种仅用一个翻译后偶联/纯化衍生化步骤的“假一步法”标记方法将在整体上大大简化所有先前公开的双重标记程序。

[0333] 由于它们的尺寸限制,带有荧光发色团的氨基酸类似物几乎不能被aaRS氨酰化,因而排除了其共翻译应用。因此,补充预负载有荧光氨基酸的tRNA是必要的。已经开发出了用nnAAs预负载tRNA的几种策略。一种策略涉及将化学制备的氨酰化二核苷酸连接到缺乏3'-CA二核苷酸的截短tRNA⁴²。该方法涉及多个步骤,因而在技术上是具有挑战性的。在另一种方法中,flexizyme(一种tRNA酰化核酶)被开发用于在体外将nnAAs负载到tRNA上⁴³。尽管原则上这种方法是直接的,但是它依赖于多步骤化学合成并且对不同的nnAAs负载效率变化显著,并且这种方法非常昂贵,因为多步骤化学合成需要大量荧光染料,并且染料的体积庞大可能导致空间位阻而妨碍反应基团的正确取向。另一种方法涉及通过反应性氨基酸侧链(例如半胱氨酸的巯基⁴⁴或赖氨酸残基的ε-氨基⁴⁵)实现与酶促氨酰化的tRNA的官能团偶联。我们选择使用半胱氨酸残基作为偶联标签,因为其修饰导致产生最短的脂肪族间隔(aliphatic spacer)并在其相应的半胱氨酰-tRNA合成酶(CysRS)中缺乏编辑结构域^{46,47}。

[0334] 在广泛的市售巯基反应性荧光化合物中,我们选择使用最小的Bodipy FL荧光团(BPFL)碘乙酰胺。由于EF-Tu或其工程化形式对酯化nnAA的不适当入位常常妨碍其有效并入(在三元复合物形成水平或核糖体上nnAA-tRNA解离的水平上)⁴⁸,所以热力学补偿假说要求在tRNA本体和酯化氨基酸之间具有相容性以实现有效的并入⁴⁹。因此,我们进行了EF-Tu结构的计算机分析以建立可接受的接头长度和荧光团尺寸。例如,虽然通过8聚体间隔物连接至tRNA(Lys)氨基-酯骨架的商业上可获得的BPFL看起来在蛋白质标记中表现良好,但是

其他相同的BPFL-吡咯基类似物的EF-Tu亲和力比wt Lys-tRNA (Lys) 低30倍⁴⁸。通过17聚体间隔物连接的生物素的并入效率通过交换到紧密结合EF-Tu的tRNA (Ala) 而可被挽救⁵⁰。由于预测环大小和间隔物长度如何影响酯化nnAA与EF-Tu的结合似乎很困难,我们着手粗略估计EF-Tu有关调节BPFL/半胱氨酰-tRNA偶联物的能力,其中5聚体间隔物将BPFL环与氨基酯骨架分开(图32, SI)。与此相适应,我们手动将在Trp位置的BPFL置于已知结构中⁵¹,将间隔长度调整为5聚体(图33, SI)。在所获得的模型中,BPFL占据口袋,没有任何明显的立体冲突,与Trp相比具有略有不同的取向,而且似乎还留有一些空间可供间隔物延伸。尽管从模型上来看,BPFL-半胱氨酸衍生物似乎很可能至少在某种程度上入位EFTu,从而提供了一些负的自由结合能,但是这种能量很难相对于Cys本身或亲和力分布非常广泛的芳香族氨基酸残基排列^{52,53}。因此,为了与BPFL耦合,我们选择保持wt T-茎和tRNA (Cys) 对EFTu的天然亲和力(从中度到紧密,取决于所使用的算法),以相互补偿弱到中度的Cys结合。

[0335] 将含有CUA或CCU反密码子的t7tRNA (Cys) 变体氨酰基并偶联至BPFL而产生BPFL-cys-tRNA,然后通过HPLC纯化偶联产物,最终产率约15-20% (图24A)。接下来,我们将纯化的BP-cys-tRNA_{accu}补充到具有半合成t7tRNA混合物但缺乏各自同工受体的反应物中,通过使用sGFP_T2模板²⁴,实现经由AGG抑制的共翻译单位点蛋白标记。反应通过具有单个AGG密码子的eGFP模板来引发,并且比较了在由不同tRNA混合物参与的未耗尽“亲本”裂解物和tRNA耗尽裂解物中所产生的标记蛋白质相对于总蛋白质的产量(图24B)。与wt裂解物相比,标记蛋白质的产量提高了1.4倍,而相对标记效率比tRNA-重构系统高10倍(图24B,泳道3和4具有BP-tRNA)。由于存在天然AGG同工受体,BPFL-tRNA与天然tRNA较少竞争以解码正常裂解物中的AGG密码子,尽管它们在天然tRNA库中丰度较低⁵⁴。在缺乏AGG同工受体的具有半合成t7tRNA套件的无细胞系统中几乎不产生sGFP(图24B,泳道3没有BP-tRNA)。加入BPFL-Cys-tRNA_{accu}后,sGFP的表达可以恢复,表明几乎所有表达的多肽都掺入了BP-荧光团。我们得出结论,BP-FL负载的tRNA与宿主翻译机制(包括延伸因子和核糖体)是相容的。在翻译过程中从nnAA释放的Cys-tRNA_{accu}进行可忽略的与半胱氨酸的共翻译再负载(图34, SI)。该系统的正交性可能是由于内源性CysRS的浓度相对较低,AGG抑制剂的k_{cat}/K_m大大降低,导致不倾向与内源性wt tRNA (Cys) 竞争,因为CysRS依赖于反密码子(特别是G35C36),它是经由CysRS识别的tRNA识别中的重要特性元件(identity element)⁴⁷。

[0336] 我们通过使用在GFP ORF的第1位和第151位具有单一琥珀密码子或AGG密码子的四个模板B_1X (151X) 和B_1R (151R) 比较了琥珀密码子介导的和AGG密码子介导的BP-FL并入效率(详细讨论参见图24C, SI)。由缺乏AGG同工受体的总tRNA混合物与tRNA耗尽的裂解物重构翻译反应物,向其中补充纯化的BPFL-cys-tRNA抑制剂并用其各自的模板引发。AGG抑制相比于琥珀抑制使BP-FL在第1位和第151位的并入量分别提高了1.8倍和4倍。这表明,与释放因子相比,AGG密码子的天然tRNA同工受体对于在琥珀密码子上的经介导的翻译终止具有更小的竞争。我们得出结论,在我们测试的密码子背景下,负载了BP-FL的tRNA与AGG更加相容。

[0337] 通过AGG和琥珀密码子重新分配的组合进行双重蛋白质标记

[0338] 如引言中所述,当前的用FRET形成染料对来进行位点特异性双重蛋白质标记的策略依赖于引入两个生物正交基团(其被遗传整合到蛋白质序列中),接着经由两个独立的反应而偶联到另一个形成FRET荧光团的基团^{7,19}。这种方法需要仔细控制这些nnAAs和荧光团

中官能团的反应性,使得在技术上具有挑战性。由于BP-FL可以共翻译并入蛋白质序列,因此安装第二荧光团将变得更直接,而不会对第二种遗传编码的nnAAs的选择设置太多限制。我们通过合成钙调素(CaM)(一种细胞功能的调节蛋白)测试了这种方法,该合成是通过分别经琥珀密码子重新分配和AGG密码子重新分配而将生物正交的反应基(AzF)和含荧光团氨基酸(BPFL-Cys)装配在第1位和第149位来实现的。然后通过无铜点击化学将TAMRA(一种与BPFL形成FRET形成染料对的物质)直接引入蛋白质中。该反应可以在纯蛋白质或无细胞反应物的条件下进行,并且避免由铜催化的叠氮化物-炔烃环加成(CuAAC)中所需要的金属催化剂引起的蛋白质聚集或氧化的副作用。荧光凝胶扫描确认了两个FRET探针在CaM蛋白中的成功安装(图25A,第一泳道)。

[0339] 双标记的CaM在488nm的激发显示荧光发射光谱中有两个不同的峰(图25B)。512nm峰对应于BP-FL的发射,而575nm的另一个峰则反映了TAMRA的发射。相比之下,用BP-FL或TAMRA单分子标记的CaM蛋白在488nm激发时,荧光发射光谱在575nm处有较低的发射。预先调整用TAMRA单标记的CaM蛋白的浓度,以使在543nm(受体激发波长)激发时其最大发射荧光与双标记蛋白质的最大发射荧光相均衡(equalize)。当两种样品在488nm处激发时,双标记蛋白在575nm处的发射荧光比TAMRA单标记蛋白的发射荧光高得多。这表明在双标记的CaM中从供体荧光团BP-FL向受体荧光团TAMRA的能量转移并且证实了成功的双重标记。

[0340] 为了证明smFRET在蛋白构象变化分析中的应用,我们用 Ca^{2+} 和EDTA滴定双标记的CaM蛋白(图25D-F)。双标记的CaM从在没有 Ca^{2+} 的缓冲液条件下的无细胞翻译反应物中纯化。其smFRET直方图显示两个峰值,一个以0.2为中心,另一个以0.55为中心。它们代表了纯化样品中的两种构象,我们猜想其中一种是无 Ca^{2+} 形式,而另一种是 Ca^{2+} 结合形式。无细胞表达的CaM可能被微量的钙污染,引起样品中一些CaM蛋白的构象变化。样品中加入2mM Ca^{2+} 后,第二个峰消失,出现第一个峰。这证实在0.2左右的第一个峰是钙结合形式。样品中添加EDTA时,第一个峰明显下降,表明部分 Ca^{2+} 被从CaM蛋白中去除。这个数据与CaM的第1位和第149位的两个荧光团在无 Ca^{2+} 形式(高FRET效率)中比在 Ca^{2+} 结合形式中更接近的结构信息是一致的(图25C)。

[0341] 讨论

[0342] 我们在这里描述了一个新的用于在大肠杆菌体外翻译系统中将Arg AGG有义密码子重新分配为uAA的策略。为实现这一目标,我们利用了最近开发的耗尽了内源性tRNA的大肠杆菌翻译系统²⁴。我们现在开发了通过DNA杂交色谱从总tRNA混合物中去除所选择的tRNA的新方法。与之前报道的使用半合成tRNA套件的重构方法不同,所开发的tRNA耗尽方法简单且有效。我们能够实现AGG密码子的天然同工受体的几乎完全耗尽,使其适合重新分配。如果没有补充tRNA_{Accu},则使用具有6AGG密码子的GFP模板没有观察到蛋白质翻译,而在含有一个1AGG密码子的模板上仅观察到最小的GFP表达。在加入合成的tRNA_{Accu}后,GFP蛋白表达增加到与野生型系统相当的水平。这验证了有义密码子重新分配策略,并且开辟了创建潜在的大量新正交密码子的方式。

[0343] 无细胞翻译系统的开放性使得能够直接控制其成分的浓度和特性。我们利用这个功能来比较几种正交tRNA/aaRS系统(包括Py1和MjTyr系统)在支持将nnAA并入到体外翻译的蛋白中的能力。我们评估了几种报道的Py1RS变体介导nnAA(包括BCN、TCO、TCO*和PrK)的并入的能力。我们对前三种非天然氨基酸特别感兴趣,因为它们具有高反应性和生物正交

性²⁶。令人惊讶的是,它们中没有一个显得是所测试的Py1RS变体的良好底物。这些结果显示与UAG抑制的体内数据出乎意料的不一致性,MS分析证实了这些nnAA可通过UAG抑制而被有效并入^{19,36}。相比之下,Py1RSAF能够以较高的效率利用PrK有效解码UK密码子,并且可以通过铜催化的点击化学而被进一步修饰。这表明通过体内选择而进化出的氨酰-tRNA合成酶特别是对于大脂肪族炔烃或烯烃可能不是最佳的。Py1RS变体总体上来说是多底物酶,其具有高度的底物诱导的结构域间动力学重排。因此,另一方面,底物结合口袋的突变可能导致总氨基酸化动力学降低数百倍³⁵。我们在纯化步骤期间观察到这些基于Py1RS的酶的沉淀。这些Py1RS变体在缓冲条件下的不稳定性对于体外应用造成额外的障碍,尽管镁、高浓度盐和甘油可以在一定程度上帮助维持这些蛋白质的可溶形式。这与一些体内数据一致,这些数据表明Py1RS表达差,易于错误折叠,并且需要优化密码子使用以及大肠杆菌中的表达条件以实现nnAA并入¹⁴。

[0344] 尽管MjY1被筛选确定为在体内最有效³⁷,并且最近的工作中使用了MjY4 (tRNA^{Opt}CUA)³⁸,但当分析MjTyr系统时,我们发现MjY2在编码AzF功能中是最有效的。进化压力确保在天然系统中所有的aa-tRNA通过将氨基酸侧链与不同的tRNA结合而显示出对延伸因子(EF-Tu)的类似亲和力^{49,55}。因此,为了达到最佳的nnAA并入效率,需要优化野生型o-tRNA序列以与所选择的目标nnAA协同工作以实现EF-Tu的最佳亲和力。

[0345] 此外,我们通过两个正交翻译系统的最佳组合MjY2/AzFRS/AzF和Py1T/Py1RSAF/PrK比较了琥珀密码子背景的解码效率。在我们测试的所有四种密码子背景下,MjTyr系统始终比Py1系统表现好。在解码UAG密码子方面,MjTyr系统的相对效率为55%到97%。相比之下,Py1T/Py1RS/PrK的抑制效率变化不定并且在一些情况下不能介导GFP翻译。这些结果与先前的发现一致,即ORF的某些位置比其他位点更容易被抑制^{56,57},并且表明当使用高效的o-aaRS/tRNA对时,背景效应降低。

[0346] 异源tRNA/aaRS/nnAA的高正交性是决定密码子重新分配的特异性和保真度的关键问题。我们的结果表明,在大肠杆菌中Py1抑制系统的正交性高于MjY系统(图30),因为在后者中我们观察到在没有AzF的情况下典型芳香族氨基酸(Tyr、Phe、Trp)的并入。然而,在翻译系统中补充AzF后,这些典型氨基酸的错误并入变得可忽略(图31),表明针对特定nnAA优化的MjTyrRS在接受其它底物方面具有低效率。尽管广泛的底物特异性对于并入具有相似结构的nnAAs具有优势,但是当需要并入不止一种nAAA时,它也增加了错误并入。每个成熟tRNA与其同源aaRS的比例都通过进化而平衡(evolutionary balanced),以确保翻译保真度和效率。

[0347] 我们证实了MjY抑制系统不局限于琥珀抑制,而且当与我们新开发的不含AGG-天然同工受体的无细胞系统结合时,与AGG重新分配到AzF也很好地兼容。MjTyrRS对MjTyrRS的tRNA反密码子三联体的低敏感性使其未来能够被用于重新分配至其他有义密码子。除了AzF,AGG密码子可以通过使用预负载的BPFL-Cys-tRNA^{Accu}而直接被重新分配到BP-FL荧光基团,从而可以显著简化标记过程。所释放的tRNA^{Accu}是由于反密码子改变而成为CysRS的较差底物,并且进行可忽略的与半胱氨酸的共翻译负载。这维持了系统的保真度。在酯化BPFL-Cys与序列相互作用的背景下进一步优化tRNA以提高其与酯化BPFL-Cys的相容性从而更好地入位EF-Tu而提高解码效率。

[0348] 我们利用上述观察的结果开发了一种新的方法来构建用两个FRET形成探针标记

的位点选择性蛋白质。为了避免与平衡两种标记反应的选择性相关的困难,我们决定在蛋白质序列中通过共翻译而并入一个荧光团。我们将最有效的MjTyr抑制系统与BPFL预负载tRNA组合,从而分别重新分配琥珀密码子和AGG密码子。通过无铜点击化学,与叠氮基在几分钟内完成反应,将TAMRA染料进一步安装到蛋白质序列中。通过使用这种策略,我们位点特异性地将BP-FL和TAMRA安装到钙调蛋白中。单分子FRET分析显示,钙与钙结合后其构象发生显著变化,证实了有效的和位点选择性的染料并入。所开发的方法潜在地允许创建大量的正交密码子并允许用各种官能团进行修饰。此外,这种方法可以直接转移到真核细胞系统中。这使得我们的方法不同于基于使用完全重构的体外翻译系统的方法,例如PURE系统,该PURE系统尽管相对容易产生而用于原核翻译,但是当转移到真核系统时带来几乎难以管理的复杂性。然而,我们最近的分析表明,大肠杆菌无细胞系统只能以非聚集形式表达10%的人类胞质蛋白,而真核系统则表现好得多⁵⁸。因此,将我们的方法转移到真核系统上,可以对复杂的多域真核蛋白质和蛋白质复合物进行smFRET分析,而这是目前这种方法很难适用的。

[0349] 补充信息(SI)

[0350] 编码蛋白质的核苷酸序列

[0351] (1) 模板A(eGFP模板,GenBank编号为KJ541667.1,基于pLTE载体)

[0352] ATGACAGTAATGTATAAAGTCTGTAAAGACATTAAACACGTAAGTGAAACCATGGAGATCXXTCGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAAGTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTATACAAGGAGCAGAAGCTGATCTCGGAGGAGGATCTGCAAGCTTGTCGACCTCTAGAGGATCCCCGGGGCTAA

[0353] 模板A从N至C末端包含SITS蛋白质、eGFP构建体和myc标签。它被用作亲本模板来制备下列衍生物:

[0354] (a) 具有6AGG密码子的GFP(该构建体中所有6个Arg密码子被换成AGG);

[0355] (b) A_{1X}:TAG密码子被引入到eGFP构建体的N末端(第1个红色密码子)

[0356] (c) A_{151X}:TAT(Tyr密码子)被换成在eGFP构建体的第151位的TAG密码子(第2个红色密码子)

[0357] (d) A_{1R}:AGG密码子被引入至eGFP构建体的N末端(第1个红色密码子)

[0358] (e) A_{151R}:TAT(Tyr密码子)被换成在eGFP构建体的第151位的AGG密码子(第2个红色密码子)

[0359] (2) 模板B(sGFP模板,基于pOPINE载体,GenBank编号为EF372397.1.)

[0360] ATGACAGAACAGAAGTTGATCTCGGAAGAAGATTTGXXTCGAAGGGAGAAGAATTGTTTACAGGAGTG

GTGCCAATCTTGGTGAATTGGACGGAGATGTGAACGGACACAAGTTCTCGGTGCGTGGAGAAGGAGAAGGAGACGC
GACAAACGGAAAGTTGACATTGAAGTTCATCTGCACAACAGGAAAGTTGCCAGTGCCATGGCCAACATTGGTGACAA
CATTGACATACGGAGTGCAGTGCTTTTCGCGTTACCCAGATCACATGAAGCGTCATGACTTCTTCAAGTCGGCGATG
CCAGAAGGATACGTGCAGGAACGTACAATCTCGTTCAAGGACGACGGAACATACAAGACACGTGCGGAAGTGAAGTT
CGAAGGAGACACATTGGTGAATCGTATTGAATTGAAGGGAATCGACTTCAAGGAGGATGGAAACATTTTGGGACACA
AGTTGGAGTACAACCTCAACTCGCACAACGTGXXXATCACAGCGGACAAGCAGAAGAACGGAATCAAGGCGAACTTC
AAGATTTCGTACAACGTGGAGGACGGATCGGTGCAGTTGGCGGATCACTACCAGCAGAATACACCAATTGGAGATGG
ACCAGTGTGTTGTTGCCAGACAATCACTACTTGTGCGACACAGTCGGTGTGTCGAAGGACCCAAACGAAAAGCGTGACC
ACATGGTGTGTTGGAATTTGTGACAGCGGCGGGAATTACACATGGAATGGACGAATTGTACAAGGGAGGATTGGAG
GTGTTGTTCCAGGGACCAGGACGTGGATCGATCGACACATGGGTGTGATGATGA

[0361] 模板A从N至C末端包含myc标签、sGFP构建体、prescission酶剪切位点和RGS肽标签。它被用作亲本模板来制备以下衍生物:

[0362] (a) B_{1X}: TAG密码子被引入在sGFP构建体的N末端(第1个红色密码子)

[0363] (b) B_{151X}: TAT (Tyr密码子) 被换成在sGFP构建体的第151位的TAG密码子(第2个红色密码子)

[0364] (c) b_{1R}: AGG密码子被引入在sGFP构建体的N末端(第1个红色密码子)

[0365] (d) B_{151R}: TAT (Tyr密码子) 被换成在sGFP构建体的第151位的AGG密码子(第2个红色密码子)

[0366] (3) sGFP_{T2}

[0367] ATGTCGAAGGGAGAAGAATTGTTACAGGAGTGGTGCCAATCTTGGTGAATTGGACGGAGATGTGAAC
GGACACAAGTTCTCGGTGCGGGAGAAGGAGAAGGAGACGCGACAAACGGAAAGTTGACATTGAAGTTCATCTGCAC
AACAGGAAAGTTGCCAGTGCCATGGCCAACATTGGTGACAACATTGACATACGGAGTGCAGTGCTTTTCGCGGTACC
CAGATCACATGAAGCGGCATGACTTCTTCAAGTCGGCGATGCCAGAAGGATACGTGCAGGAACGACAATCTCGTTC
AAGGACGACGGAACATACAAGACAAGGGCGGAAGTGAAGTTCGAAGGAGACACATTGGTGAATCGGATTGAATTGAA
GGGAATCGACTTCAAGGAGGATGGAAACATTTTGGGACACAAGTTGGAGTACAACCTCAACAGCCACAACGTGTACA
TCACAGCGGACAAGCAGAAGAACGGAATCAAGGCGAACTTCAAGATTCCGCACAACGTGGAGGACGGATCGGTGCAG
TTGGCGGATCACTACCAGCAGAATACACCAATTGGAGATGGACCAGTGTGTTGCCAGACAATCACTACTTGTGCGAC
ACAGTCGGTGTGTCGAAGGACCCAAACGAAAAGCGGGACCACATGGTGCTATTGGAATTTGTGACAGCGGCGGAA
TTACACATGGAATGGACGAATTGTACAAG

[0368] 模板sGFP_{T2}¹仅含有sGFP蛋白的开放阅读框。该蛋白质中第108位的精氨酸由AGG编码(红色),其余由CGG密码子编码。

[0369] (4) CaM

[0370] ATGGAACAGAAGCTGATTTCCGAAGAGGATCTGGGATAGGATCAGCTGACGGAAGAGCAGATCGCGGAA
TTCAAGGAAGCGTTCTCCCTGTTTCGACAAGGACGGAGACGGAACGATTACGACGAAGGAAGTGGGAACGGTGATGCG
TTCCCTGGGACAGAATCCGACGGAAGCGGAGCTGCAGGACATGATCAACGAAGTGGATGCGGACGGAACGGAACGA
TTGACTTCCCGGAATTCCTGACGATGATGGCGCGTAAGATGAAGGATACGGACTCCGAAGAGGAAATCCGTGAAGCG
TTTCGTGTGTTTCGACAAGGATGGAAATGGATACATTTCCGCGGCGGAGCTGCGTCACGTGATGACGAACCTGGGAGA
GAAGCTGACGGACGAAGAGGTGGACGAAATGATTCGTGAAGCGGACATTGACGGAGACGGACAGGTGAATTACGAAG
AGTTTCGTGCAGATGATGACGGCGAGGGGAGGACTGGAGGTGCTGTTCAGGGACCGGGACGTGGATCCATCGACACG

TGGGTGTGATGATGA

[0371] 如载体图谱所示,CaM模板从N到C末端含有myc标签、CaM蛋白、prescission酶剪切位点和RGS肽标签。

[0372] 材料和方法

[0373] 重组半胱氨酰-tRNA合成酶生产:通过在大肠杆菌BL21 (DE3) 菌株上的菌落PCR扩增CysRS的基因片段,并将其克隆到具有C端His标签的pET28载体中,并通过标准亲和层析纯化。用0.5mM异丙基- β -D-硫代吡喃半乳糖苷 (IPTG) 在BL21 (DE3) RIL细胞中于20℃诱导蛋白质表达过夜。通过使用标准缓冲液的Ni²⁺亲和层析纯化CysRS,然后在Su-perdex 200 (GEHealthcare) 上在缓冲液:50mM Hepes-KOH、pH8、150mM NaCl和1mM DTT中进行凝胶过滤。

[0374] 亲和钳基质制备:将C末端Cys添加到亲和钳蛋白 (pdz-纤连蛋白融合蛋白,PDB:3CH8_A) 的序列中。使用制造商的方案将纯化的具有C端Cys的蛋白质固定在碘乙酰凝胶 (UltraLink®) 上。

[0375] 通过flexizyme进行tRNA负载而用于琥珀抑制:按之前报道进行酰化反应⁶²。1体积的250 μ M o-tRNA与相同体积的50 μ M Fx (250 μ M) 和3体积的水混合。然后将混合物在95℃加热5分钟,同时以1200rpm恒速搅拌,并在室温下缓慢冷却。加入1倍体积的0.5M Hepes (pH7.9) 和2倍体积的3M MgCl₂后,将混合物置于冰上2分钟,然后加入2V的nAAA底物 (100mM) (在100%DMSO或100%ACN中)。将反应物在搅拌下于10°温育给定的时间段 (2.5、5或10小时)。停止Flexizyme酰化反应并将其用于引发翻译反应。

[0376] MS分析:将修饰的蛋白质在纳米陷阱 (nano-trap) 基质上纯化,然后用1xSDS PAGE上样缓冲液洗脱。在凝胶上电泳后,切下所需的条带并用胰蛋白酶消化。然后通过LC-MS/MS分析所得的肽。

[0377] 结果和讨论

[0378] 尽管商业化的tRNA混合物在支持1AGG sGFP模板的翻译方面显示出比本文产生的tRNA混合物的效率低~2倍,但其在支持具有6AGG密码子的eGFP模板在饱和浓度下的翻译方面显示出3倍高的活性 (图21和图27)。这表明商业混合物中天然的罕见AGG-密码子同工受体的比例较高。将t7tRNA^{Accu}加入到耗尽的自制tRNA混合物中导致翻译活性与原始的总tRNA混合物相比增加5倍 (图21)。这证实了我们的假设,即在适当的细胞培养时段,用于解码AGG密码子的低丰度tRNA的含量偏向有利于密码子重新分配的目的。

[0379] 为了确保翻译保真度,外源性tRNA/aaRS对应与内源性tRNA/aaRS对正交。我们评估了大肠杆菌无细胞翻译系统中MjY2/AzFRS/AzF和Py1T/Py1RSAF/PrK的翻译保真度。通过补充o-tRNA/aaRS/nnAA的不同组合来在无细胞反应物中翻译模板A₁₅₁X。对于MjTyr系统,排除tRNA或aaRS时没有翻译出蛋白质。但是,当从反应物中排除AzF时产生了约10%的蛋白质。这表明AzFRS可以用典型氨基酸对MjY2tRNA进行负载以实现琥珀抑制。我们还发现,随着AzFRS浓度的增加,用典型氨基酸对MjtRNA进行负载的能力增加 (数据未显示)。为了检验这个问题,通过在纳米陷阱基质上使用亲和层析纯化使用或未使用AzF生产的蛋白质,并上样进行SDS-PAGE凝胶电泳。切下相应的条带,用胰蛋白酶消化。然后使用LC-MS/MS分析所得的肽以评估期望位置处的氨基酸的特性。没有AzF补充的GFP产物的MS谱分析显示,一些芳族氨基酸Phe、Tyr和Trp被并入在TAG密码子上 (数据未显示)。在GFP151AzF样品中,野生型

肽可以忽略不计,并且超过80%的蛋白质含有AzF(图32A)。AzF并入也可以通过经由DIBO-TAMRA的TAMRA并入来确认。

[0380] Py1T/Py1RSAF对的正交性优于MjY2/AzFRS对。排除nnAA、Py1T或Py1RS,没有观察到可检测的荧光。然而,当用MjY2tRNA补充反应时,蛋白质产量略有增加,表明MjYtRNA可能被Py1RS用Prk错误地酰化。每个成熟tRNA与其同源aaRSs的比例被通过进化而平衡,以确保翻译保真度和效率。因此,将第21个和第22个tRNA/aaRS对引入翻译系统时,必须注意避免交叉识别。

[0381] 我们通过使用在第1位和第151位GFP ORF具有单个AGG或单个琥珀密码子的四个模板,比较了琥珀密码子介导的和AGG密码子介导的BP-FL并入效率(图25C)。1TAG上BP-FL标记蛋白的产量是151TAG的3.7倍,而1AGG上BP-FL标记蛋白的产量是151AGG的1.2倍。对于琥珀密码子和AGG密码子而言,在第1位的抑制效率始终都高于第151位。蛋白质印迹上的相对总蛋白质产量由带强度表示。如果我们假设151TAG抑制的全长蛋白质(图25C,泳道3,抗GFP抗体)直接归因于BP-FL的并入,那么应该将BP-FL/抗GFP抗体(上方条带)的带强度比作为100%标记效率。1AGG抑制的荧光/总蛋白比例(图25C,泳道2)明显高于151TAG,表明超过100%的标记效率,这是不可能的。因此,在151TAG抑制的全长条带中存在一些非特异性并入。它不仅包含具有BP-FL的产物,而且含有具有其他一些非荧光氨基酸的产物。最大的可能性是Cys,它在源于BPFL-tRNA抑制的蛋白质翻译后期通过释放的tRNA^{Cys}_{scua}并入蛋白质中。这种非特异性并入是背景依赖性的,因为1TAG的荧光/总蛋白比率($3.7/1=3.7$)比151TAG的($1/(1+1.1)=0.48$)高约8倍。因此,从由全长/(全长加截短)比率推导的TAG标记效率来估计AGG标记效率并不是直接的。然而,这里1TAG:1AGG:151TAG:1AGG的标记效率是4:7:1:4。AGG抑制比琥珀抑制提供了持续较高的BP-FL并入,在第1位和第151位分别是琥珀抑制的1.8和4倍。这表明,与释放因子相比,AGG密码子的天然tRNA同工受体对于在琥珀密码子上的经介导的翻译终止具有更小的竞争。我们得出结论,在我们测试的密码子背景下,负载了BP-FL的tRNA与AGG更加相容。

[0382] 表8.用于合成的tRNA制备及天然同工受体耗尽的寡聚物

[0383]

trRNA 名称	寡聚物名称	序列
MjY1	T7	GTAATACGACTCACTATA GG
	F1	GTAATACGACTCACTATAGGGAGACTAAGGCGGGCTGATGAGTCCGTGAGGACGAAACGG
	F2	CCGTGAGGACGAAACGGTACCCGGTACCGTCCCCGCCTTAGTTTCAGAGGGCAGAACGGCG
	R1	GAACCCGTGCCATGCGGATTAGAGTCCGCCGTTCTGCCCTC
	R2	TGGTCCCGCCTACGGGATTGAACCCGTGCCATGCGGA
	R3	TGGTCCCGCCTACGGGATT
MjY2	F1	GTAATACGACTCACTATAGGGAGACTACGGCCGGCTGATGAGTCCGTGAGGACGAAACGG
	F2	CCGTGAGGACGAAACGGTACCCGGTACCGTCCCGGCCGTAGTTTCAGCAGGGCAGAACGGC
	R1	GAACCCATGCCATGCGGATTAGAGTCCGCCGTTCTGCCCTGC
	R2	TGGTCCGGCCGATGGGATTGAACCCATGCCATGCGGA
	R3	TGGTCCGGCCGATGGGATT
MjY3	F1	GTAATACGACTCACTATAGGGAGACTAGGGCGGGCTGATGAGTCCGTGAGGACGAAACGG
	F2	CCGTGAGGACGAAACGGTACCCGGTACCGTCCCCGCCCTAGTTTCAGAGGCAGAACGGCGG
	R1	GAACCCGAGCCATGCGGATTAGAGTCCGCCGTTCTGCCCTC
	R2	TGGTCCCGCCCTCGGGATTGAACCCGAGCCATGCGGA
	R3	TGGTCCCGCCCTCGGGATT
MjY4	F1	GTAATACGACTCACTATAGGGAGACTACCGCCGGCTGATGAGTCCGTGAGGACGAAACGG
	F2	CCGTGAGGACGAAACGGTACCCGGTACCGTCCCGCGGGTAGTTTCAGCAGGGCAGAACGGC
	R1	GAACCCCTGCCATGCGGATTAGAGTCCGCCGTTCTGCCCTGC
	R2	TGGTCCGGCGGAGGGGATTGAACCCCTGCCATGCGGA
	R3	TGGTCCGGCGGAGGGGATT
PylT(u25c)	F1	CTAATACGACTCACTATAGGAAACCTGATCATGTAGATCGAACGGAC
	R1	TAACCCGGCTGAACGGATTAGAGTCCGTTTCGATCTACATG
	R2	TGGCGGAAACCCGGGAATCTAACCCGGCTGAACGG
	R3	TGGCGGAAACCCGGGAATC
tRNA ^{Tyr} (cua)	F1	GTAATACGACTCACTATAGGTGGGGTTCCCGAGCGGCCAAAGGGAGCAGA
	R1	GAACCTTCGAAGTCGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTC
	R2	TGGTGGTGGGGGAAGGATTCTGAACCTTCGAAGTCGATGACGGCAGA
	R3	TGGTGGTGGGGGAAGGATT
RNACys(ccu)	F1	GTAATACGACTCACTATAGGCGCGTTAACAAGCGGTTATGTAGCGGA
	R1_ccu	GAACCGGACTAGACGGATTAGGAATCCGCTACATAACCGCTT
	R2	TGGAGGCGCGTTCGGGAGTCGAACCGGACTAGACGGA
	R3	TGGAGGCGCGTTCGGGAGTC
tRNACys(cua)	R1_cua	GAACCGGACTAGACGGATTAGGAATCCGCTACATAACCGCTT
用于 tRNA ^{Arg} (CCU)耗尽		TGTCCCCTGCAGGAATCGAACCTGCAATTA(3'-氨基)
用于 tRNA ^{Arg} (UCU)耗尽		AGAAGGTCGTTGCTCTATCCAACCTGAGCTA(3'-氨基)

[0384] 注意:2个正向(T7,F)和3个反向寡核苷酸(R1,R2,R3)用于装配标准tRNA模板,而3个正向(T7,F1,F2)和3个反向寡核苷酸用于产生自加工的锤头状核酶(HHRz)融合构建体(REF)。用于耗尽总混合物中的天然同工受体的寡聚物含有3'-胺以固定在NHS活化的基质上。

[0385] 参考文献

- [0386] (1) Crick, F.H.J. Mol Biol 1968, 38, 367.
- [0387] (2) Liu, C.C.; Schultz, P.G. Annu Rev Biochem 2010, 79, 413.
- [0388] (3) O'Donoghue, P.; Ling, J.; Wang, Y.S.; Soll, D. Nature chemical biology 2013, 9, 594.
- [0389] (4) Odoi, K.A.; Huang, Y.; Rezenom, Y.H.; Liu, W.R. Plos One 2013, 8, e57035.
- [0390] (5) Wan, W.; Huang, Y.; Wang, Z.Y.; Russell, W.K.; Pai, P.J.; Russell, D.H.; Liu, W.R. Angew Chem Int Edit 2010, 49, 3211.

- [0391] (6) Neumann, H.; Wang, K.H.; Davis, L.; Garcia-Alai, M.; Chin, J.W. *Nature* 2010, 464, 441.
- [0392] (7) Wang, K.H.; Sachdeva, A.; Cox, D.J.; Wilf, N.W.; Lang, K.; Wallace, S.; Mehl, R.A.; Chin, J.W. *Nat Chem* 2014, 6, 393.
- [0393] (8) Wu, B.; Wang, Z.; Huang, Y.; Liu, W.R. *Chembiochem* 2012, 13, 1405.
- [0394] (9) Krishnakumar, R.; Prat, L.; Aerni, H.R.; Ling, J.Q.; Merryman, C.; Glass, J.I.; Rinehart, J.; Soll, D. *Chembiochem* 2013, 14, 1967.
- [0395] (10) Lee, B.S.; Shin, S.; Jeon, J.Y.; Jang, K.S.; Lee, B.Y.; Choi, S.; Yoo, T.H. *ACS chemical biology* 2015, 10, 1648.
- [0396] (11) Mukai, T.; Yamaguchi, A.; Ohtake, K.; Takahashi, M.; Hayashi, A.; Iraha, F.; Kira, S.; Yanagisawa, T.; Yokoyama, S.; Hoshi, H.; Kobayashi, T.; Sakamoto, K. *Nucleic acids research* 2015, 43, 8111.
- [0397] (12) Zeng, Y.; Wang, W.; Liu, W.S.R. *Chembiochem* 2014, 15, 1750.
- [0398] (13) Dumas, A.; Lercher, L.; Spicer, C.D.; Davis, B.G. *Chem. Sci.* 2015, 6, 50.
- [0399] (14) Chatterjee, A.; Sun, S.B.; Furman, J.L.; Xiao, H.; Schultz, P.G. *Biochemistry* 2013, 52, 1828.
- [0400] (15) Sasmal, P.K.; Carregal-Romero, S.; Han, A.A.; Streu, C.N.; Lin, Z.; Namikawa, K.; Elliott, S.L.; Koster, R.W.; Parak, W.J.; Meggers, E. *Chembiochem* 2012, 13, 1116.
- [0401] (16) Iwane, Y.; Hitomi, A.; Murakami, H.; Katoh, T.; Goto, Y.; Suga, H. *Nat Chem* 2016.
- [0402] (17) Quast, R.B.; Mrusek, D.; Hoffmeister, C.; Sonnabend, A.; Kubick, S. *Febs Lett* 2015, 589, 1703.
- [0403] (18) Michalet, X.; Weiss, S.; Jager, M. *Chem Rev* 2006, 106, 1785.
- [0404] (19) Nikic, I.; Plass, T.; Schraidt, O.; Szymanski, J.; Briggs, J.A.G.; Schultz, C.; Lemke, E.A. *Angew Chem Int Edit* 2014, 53, 2245.
- [0405] (20) Sachdeva, A.; Wang, K.H.; Elliott, T.; Chin, J.W. *Journal of the American Chemical Society* 2014, 136, 7785.
- [0406] (21) Albayrak, C.; Swartz, J.R. *Nucleic acids research* 2013, 41, 5949.
- [0407] (22) Loscha, K.V.; Herlt, A.J.; Qi, R.; Huber, T.; Ozawa, K.; Otting, G. *Angewandte Chemie* 2012, 51, 2243.
- [0408] (23) Wan, W.; Tharp, J.M.; Liu, W.R. *Biochimica et biophysica acta* 2014, 1844, 1059.
- [0409] (24) Cui, Z.; Stein, V.; Tnimov, Z.; Mureev, S.; Alexandrov, K. *J Am Chem Soc* 2015, 137, 4404.
- [0410] (25) Yokogawa, T.; Kitamura, Y.; Nakamura, D.; Ohno, S.; Nishikawa, K. *Nucleic acids research* 2010, 38, e89.
- [0411] (26) Lang, K.; Chin, J.W. *ACS chemical biology* 2014, 9, 16.
- [0412] (27) Carlson, J.C.; Meimetis, L.G.; Hilderbrand, S.A.; Weissleder,

R.Angewandte Chemie 2013,52,6917.

[0413] (28) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. *Biochem Biophys Res Commun* 2008, 371, 818.

[0414] (29) Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. *Chem Biol* 2008, 15, 1187.

[0415] (30) Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W. *J Am Chem Soc* 2009, 131, 8720.

[0416] (31) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C. J.; Deiters, A.; Chin, J. W. *Nat Chem* 2012, 4, 298.

[0417] (32) Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J Am Chem Soc* 2012, 134, 10317.

[0418] (33) Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A. *Journal of the American Chemical Society* 2012, 134, 2898.

[0419] (34) Wang, Y. S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. *Journal of the American Chemical Society* 2012, 134, 2950.

[0420] (35) Guo, L. T.; Wang, Y. S.; Nakamura, A.; Eiler, D.; Kavran, J. M.; Wong, M.; Kiessling, L. L.; Steitz, T. A.; O'Donoghue, P.; Soll, D. *Proc Natl Acad Sci U S A* 2014, 111, 16724.

[0421] (36) Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *Journal of the American Chemical Society* 2012, 134, 10317.

[0422] (37) Guo, J. T.; Melancon, C. E.; Lee, H. S.; Groff, D.; Schultz, P. G. *Angew Chem Int Edit* 2009, 48, 9148.

[0423] (38) Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J Mol Biol* 2010, 395, 361.

[0424] (39) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Journal of the American Chemical Society* 2002, 124, 9026.

[0425] (40) Steer, B. A.; Schimmel, P. J. *Biol Chem* 1999, 274, 35601.

[0426] (41) Tsunoda, M.; Kusakabe, Y.; Tanaka, N.; Ohno, S.; Nakamura, M.; Senda, T.; Moriguchi, T.; Asai, N.; Sekine, M.; Yokogawa, T.; Nishikawa, K.; Nakamura, K. T. *Nucleic acids research* 2007, 35, 4289.

[0427] (42) Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. *The Journal of biological chemistry* 1978, 253, 4517.

[0428] (43) Goto, Y.; Katoh, T.; Suga, H. *Nat Protoc* 2011, 6, 779.

[0429] (44) Gubbens, J.; Kim, S. J.; Yang, Z.; Johnson, A. E.; Skach, W. R. *Rna* 2010, 16, 1660. (45) Johnson, A. E.; Woodward, W. R.; Herbert, E.; Menninger, J. R. *Biochemistry* 1976, 15, 569.

[0430] (46) Newberry, K. J.; Hou, Y. M.; Perona, J. J. *Embo J* 2002, 21, 2778.

[0431] (47) Hauenstein, S.; Zhang, C. M.; Hou, Y. M.; Perona, J. J. *Nat Struct Mol Biol* 2004, 11, 1134.

- [0432] (48) Mittelstaet, J.; Konevega, A.L.; Rodnina, M.V. *Journal of the American Chemical Society* 2013, 135, 17031.
- [0433] (49) Schrader, J.M.; Chapman, S.J.; Uhlenbeck, O.C. *Proc Natl Acad Sci U S A* 2011, 108, 5215.
- [0434] (50) Jeong, K.W.; Pavlov, M.Y.; Kwiatkowski, M.; Ehrenberg, M.; Forster, A.C. *Rna—a Publication of the Rna Society* 2014, 20, 632.
- [0435] (51) Voorhees, R.M.; Schmeing, T.M.; Kelley, A.C.; Ramakrishnan, V. *Science* 2010, 330, 835.
- [0436] (52) Asahara, H.; Uhlenbeck, O.C. *P Natl Acad Sci USA* 2002, 99, 3499.
- [0437] (53) Asahara, H.; Uhlenbeck, O.C. *Biochemistry* 2005, 44, 11254.
- [0438] (54) Dong, H.J.; Nilsson, L.; Kurland, C.G. *J Mol Biol* 1996, 260, 649.
- [0439] (55) LaRiviere, F.J.; Wolfson, A.D.; Uhlenbeck, O.C. *Science* 2001, 294, 165.
- [0440] (56) Bossi, L.; Ruth, J.R. *Nature* 1980, 286, 123.
- [0441] (57) Pott, M.; Schmidt, M.J.; Summerer, D. *ACS chemical biology* 2014, 9, 2815.
- [0442] (58) Gagoski, D.; Polinkovsky, M.E.; Mureev, S.; Kunert, A.; Johnston, W.; Gambin, Y.; Alexandrov, K. *Biotechnology and bioengineering* 2016, 113, 292.
- [0443] (59) Zubay, G. *J Mol Biol* 1962, 4, 347.
- [0444] (60) Schwarz, D.; Junge, F.; Durst, F.; Frolich, N.; Schneider, B.; Reckel, S.; Sobhanifar, S.; Dotsch, V.; Bernhard, F. *Nat Protoc* 2007, 2, 2945.
- [0445] (61) Cui, Z.; Stein, V.; Tnimov, Z.; Mureev, S.; Alexandrov, K. *Journal of the American Chemical Society* 2015, 137, 4404.
- [0446] (62) Goto, Y.; Katoh, T.; Suga, H. *Nat Protoc* 2011, 6, 779.
- [0447] (63) Guo, J.T.; Melancon, C.E.; Lee, H.S.; Groff, D.; Schultz, P.G. *Angew Chem Int Edit* 2009, 48, 9148.
- [0448] (64) Young, T.S.; Ahmad, I.; Yin, J.A.; Schultz, P.G. *J Mol Biol* 2010, 395, 361.
- [0449] 实施例8
- [0450] 通过两个有义密码子的双重蛋白标记
- [0451] AGG和AGC天然tRNA抑制剂的总tRNA混合物的特异性耗尽
- [0452] 基于反密码子环特异性tRNA适配体的特异性耗尽(按照实施例2),接着是基于tRNA特异性DNA寡核苷酸的特异性耗尽(按照实施例7“总tRNA混合物中的特异性tRNA耗尽”),组合用于特异性tRNA耗尽的这两种方法,我们制备了丝氨酸tRNA GCU同工受体受体(c11)和精氨酸CCU/UCU同工受体(a05/06)高度耗竭的总tRNA混合物。在抑制实验中,两种tRNA都显示出高度耗尽(图35)。
- [0453] 通过两个有义密码子的双重蛋白标记
- [0454] 在具有不同密码子偏好的相应模板引发的无细胞翻译反应中,通过AGC和AGG密码子将AzF和BP-FL位点特异性并入CaM和eGFP中。在CaM ORF的第1位(1和2)和151位分别具有单一或双重AGC密码子和单个AGG密码子的pOPINE-CaM模板以100ng/μl被补充到无细胞反应物中,该反应物含有35% (v/v) 耗尽裂解物、0.5μg/μl耗尽的tRNA混合物、12.5μM MjY2tRNA_{gcu}、10μM AzFRS、1mM AzF; 16μM BP-tRNA_{Cysccu}。反应在28℃下进行3小时。将来

自200 μ l无细胞反应的翻译产物固定在亲和力树脂珠上(参见实施例7)。将具有固定有翻译产物的树脂用25mM Tris (pH7.5)、0.25M NaCl、0.1%Triton X-100缓冲液洗涤以从翻译反应中除去游离的AzF,接着加入DIBO-TAMRA,使终浓度在两倍原始反应体积的PBS中达到1mM。在室温下偶联1h后,洗涤珠以去除未结合的TAMRA-染料,并且用95℃加热的x2SDS-PAGE样品缓冲液洗脱对应于从10 μ l翻译反应获得的蛋白质的小珠等份(the aliquote of beads),并在SDS-PAGE上解析(图37)。

[0455] 整个说明书的目的是描述本发明的优选实施例,而不将本发明限制为任何一个实施例或特征的特定集合。在不脱离本发明的情况下,可以对描述和示出的实施例进行各种改变和修改。

[0456] 本说明书中引用的每个专利和科学文献的公开内容、计算机程序和算法通过引用整体并入。

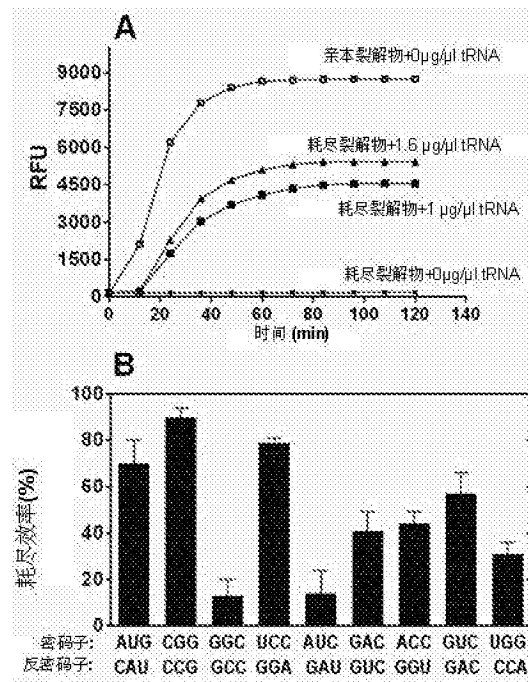


图3

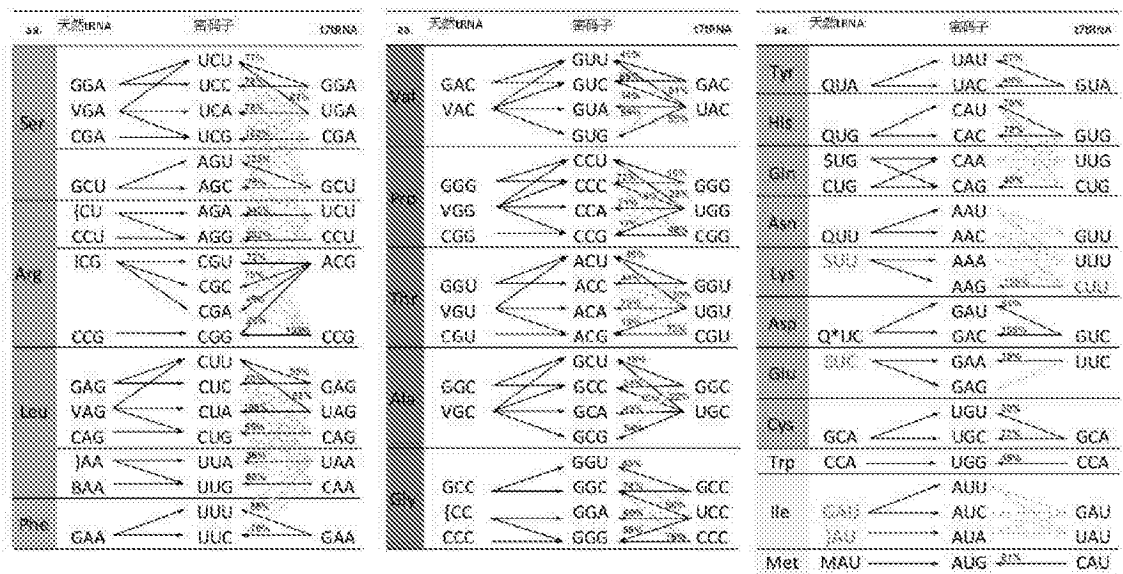


图4

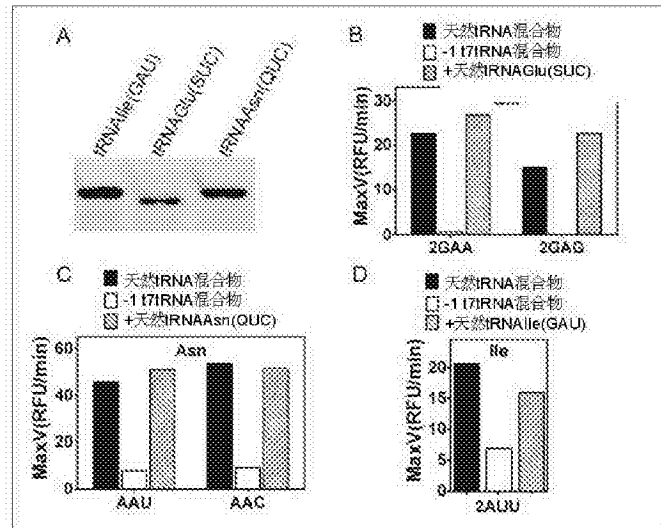


图5

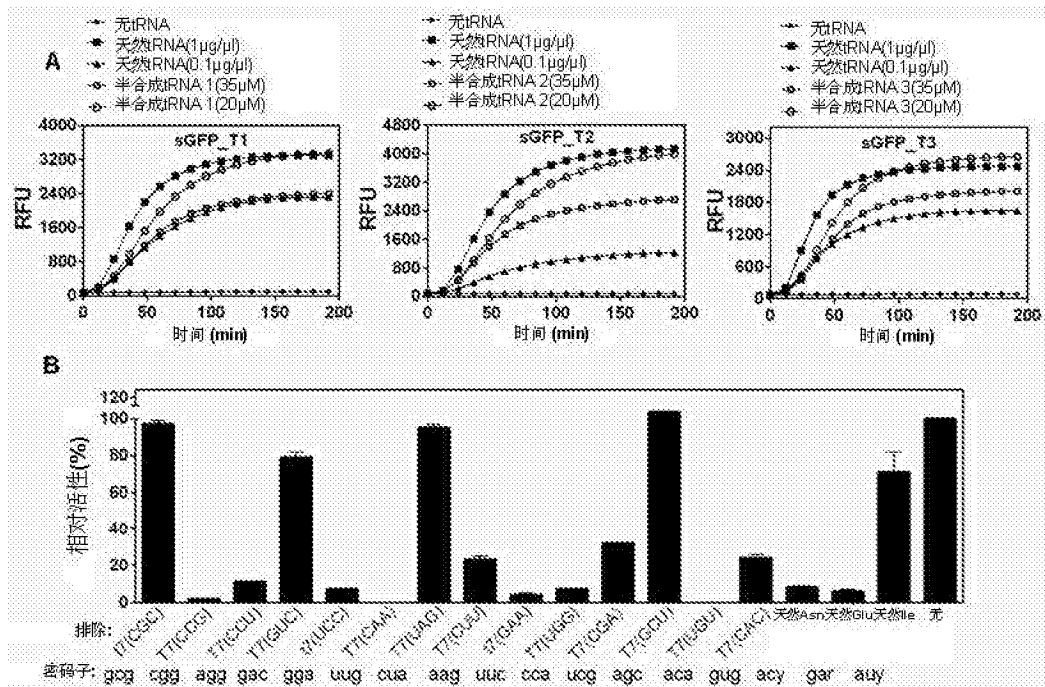


图6

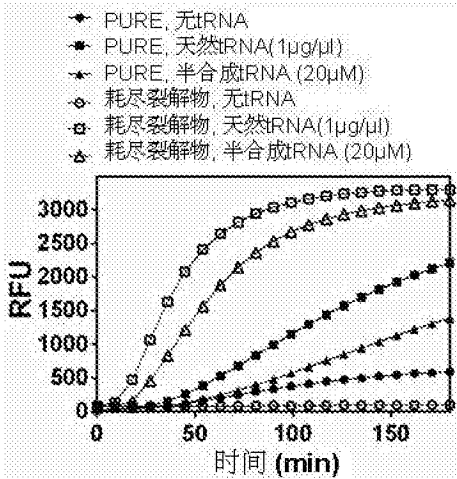


图7

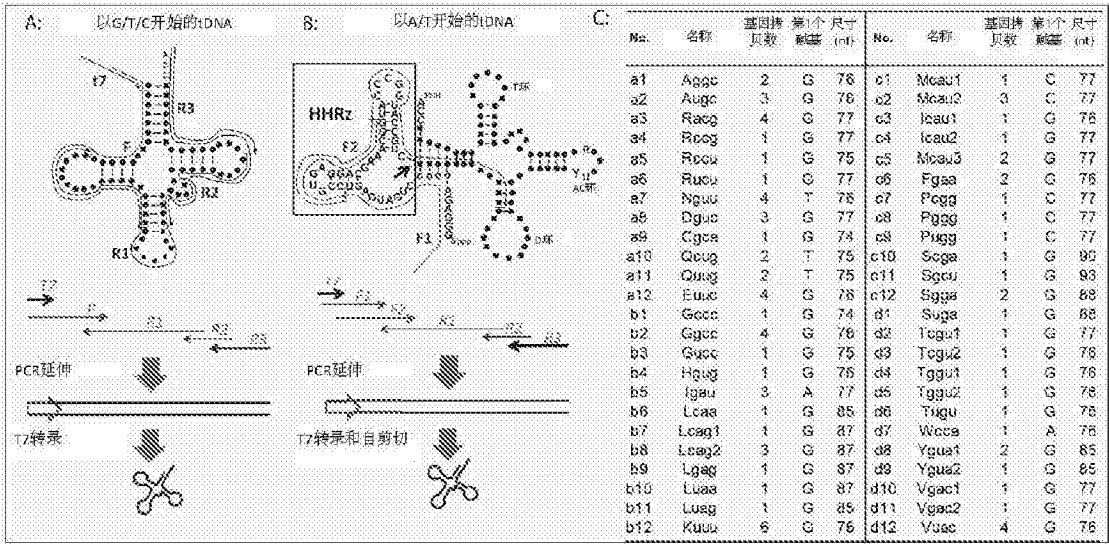


图8

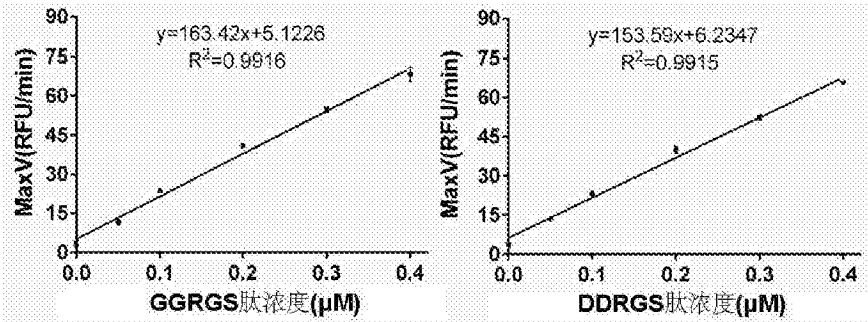


图9

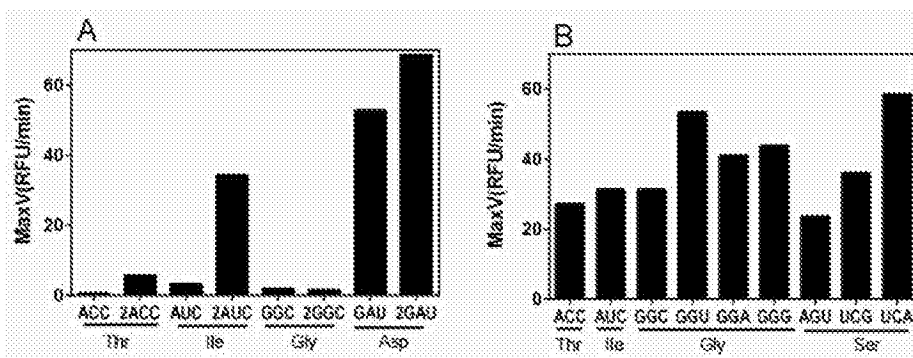


图10

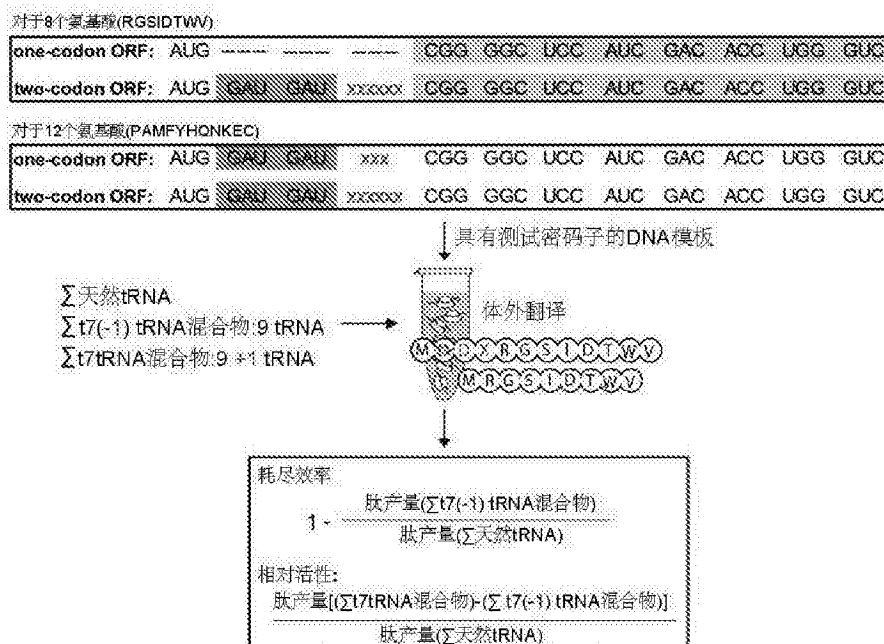


图11

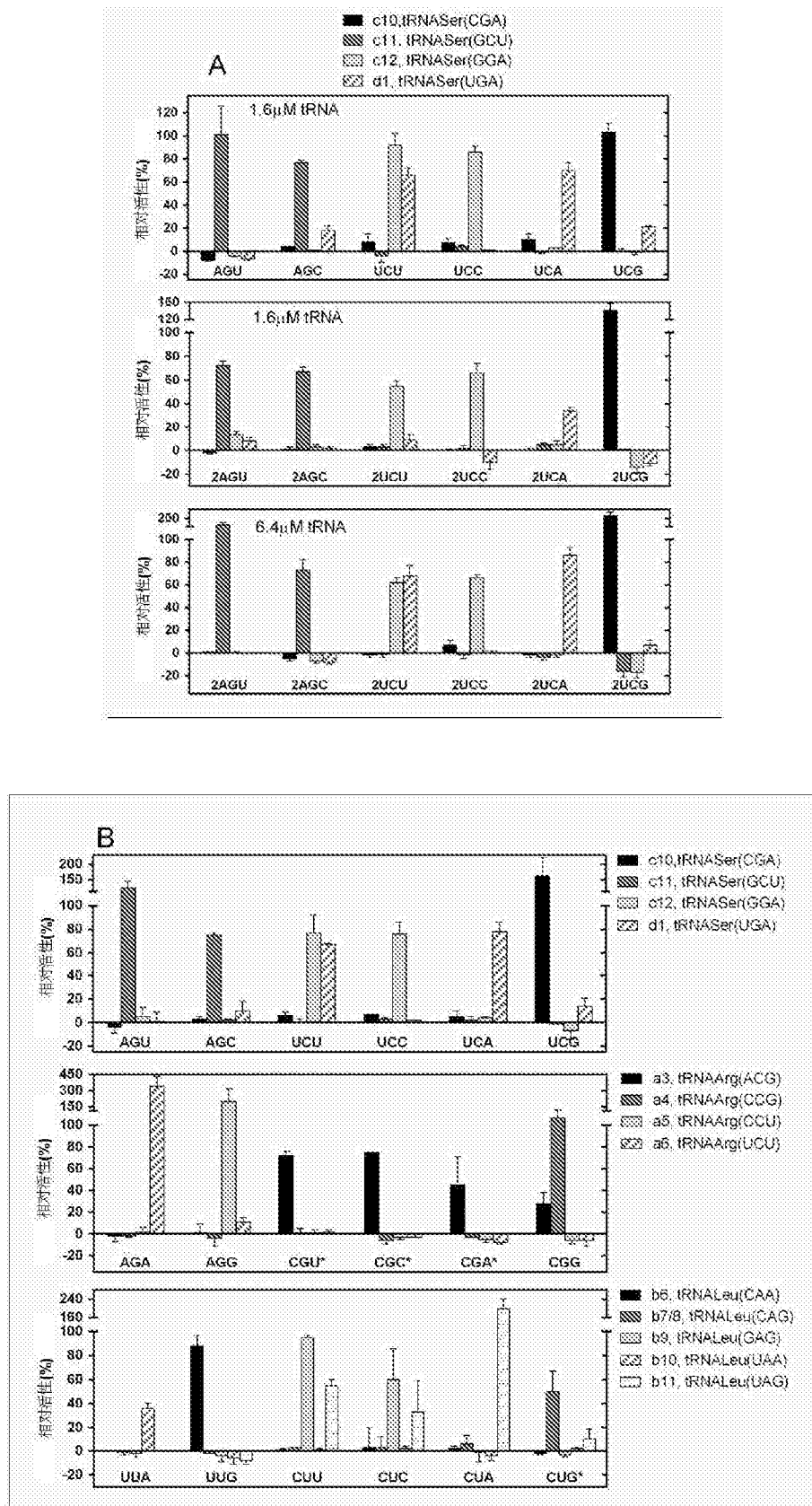


图12

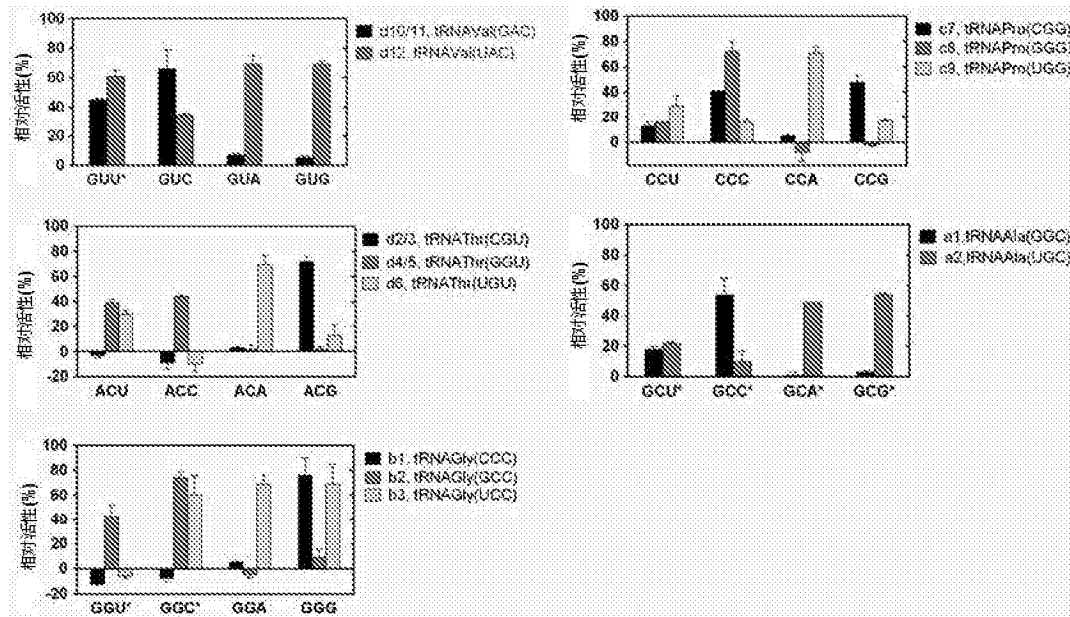


图13

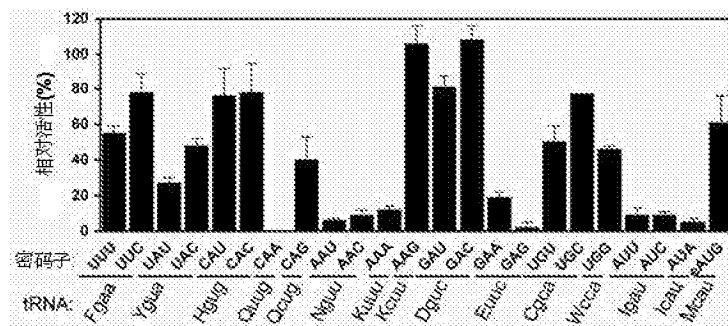


图14

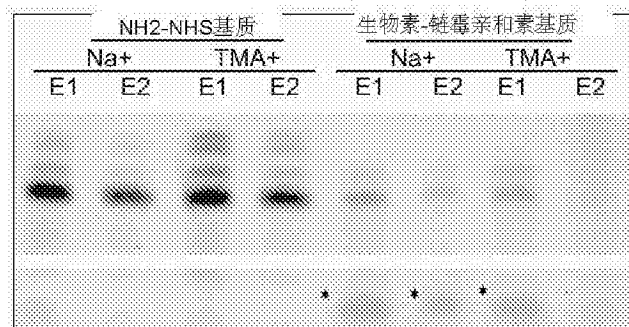


图15

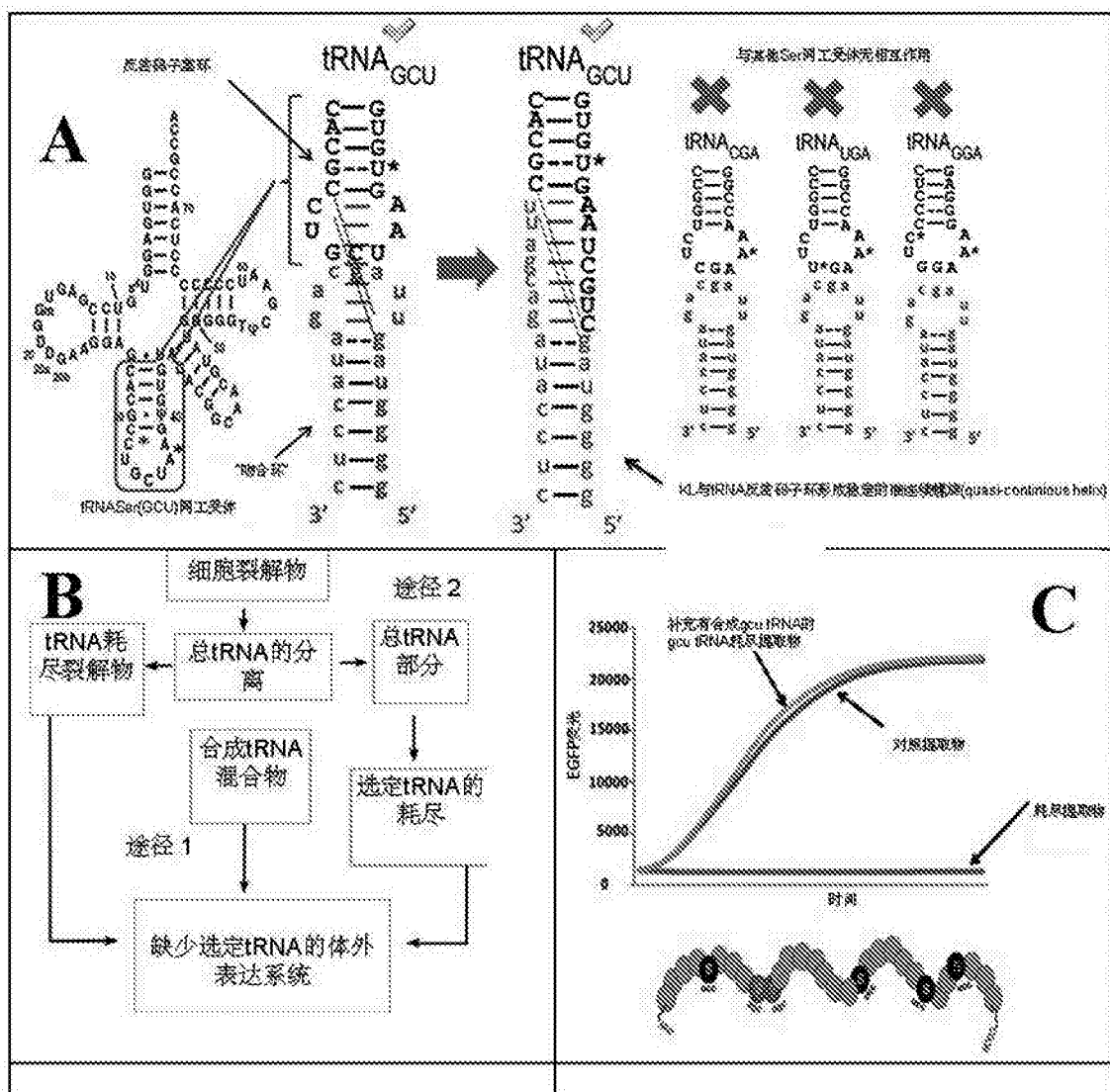


图16

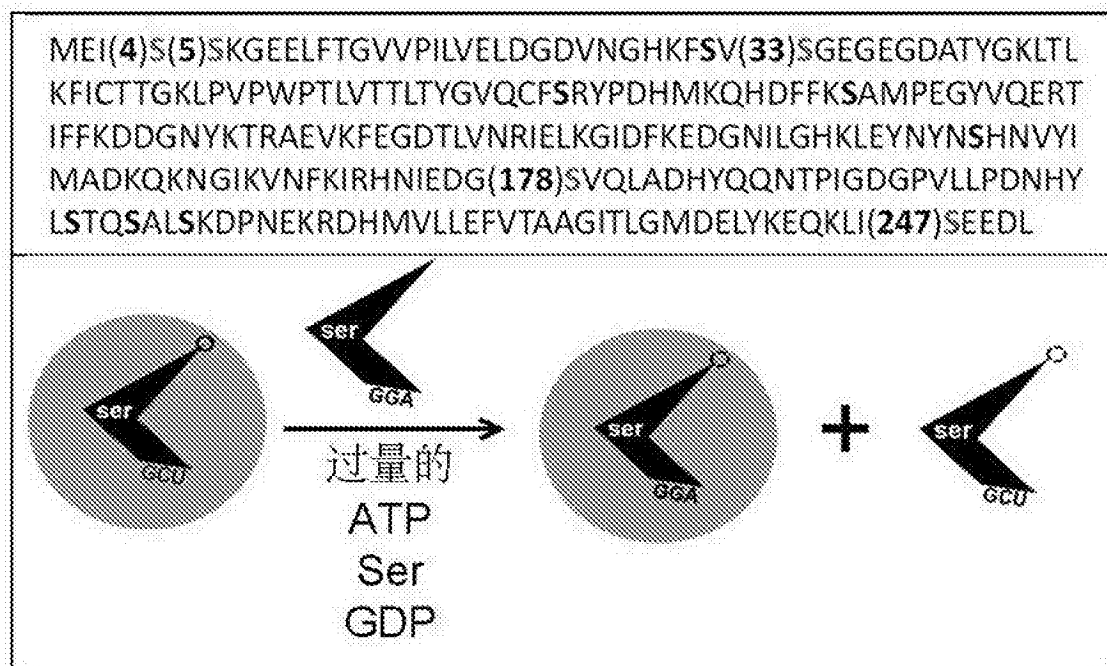


图17

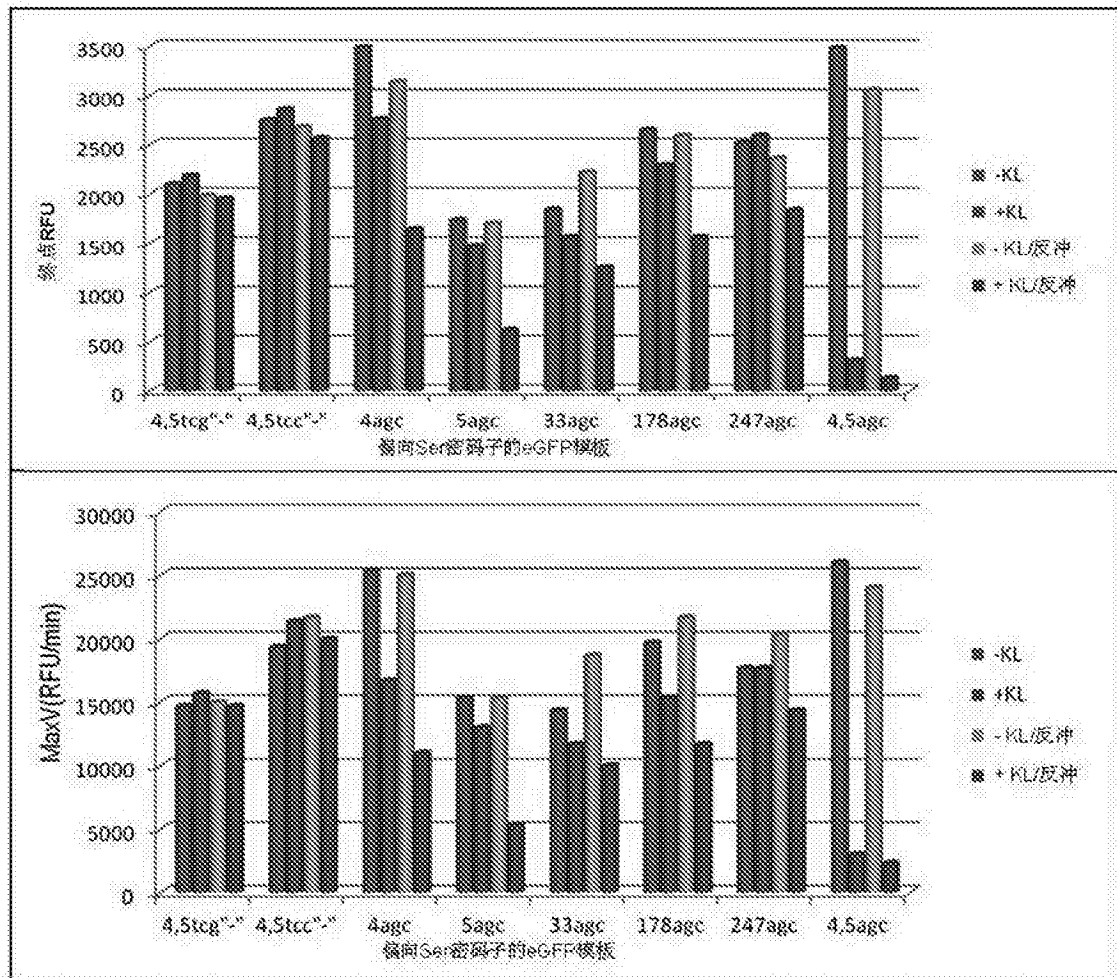


图18

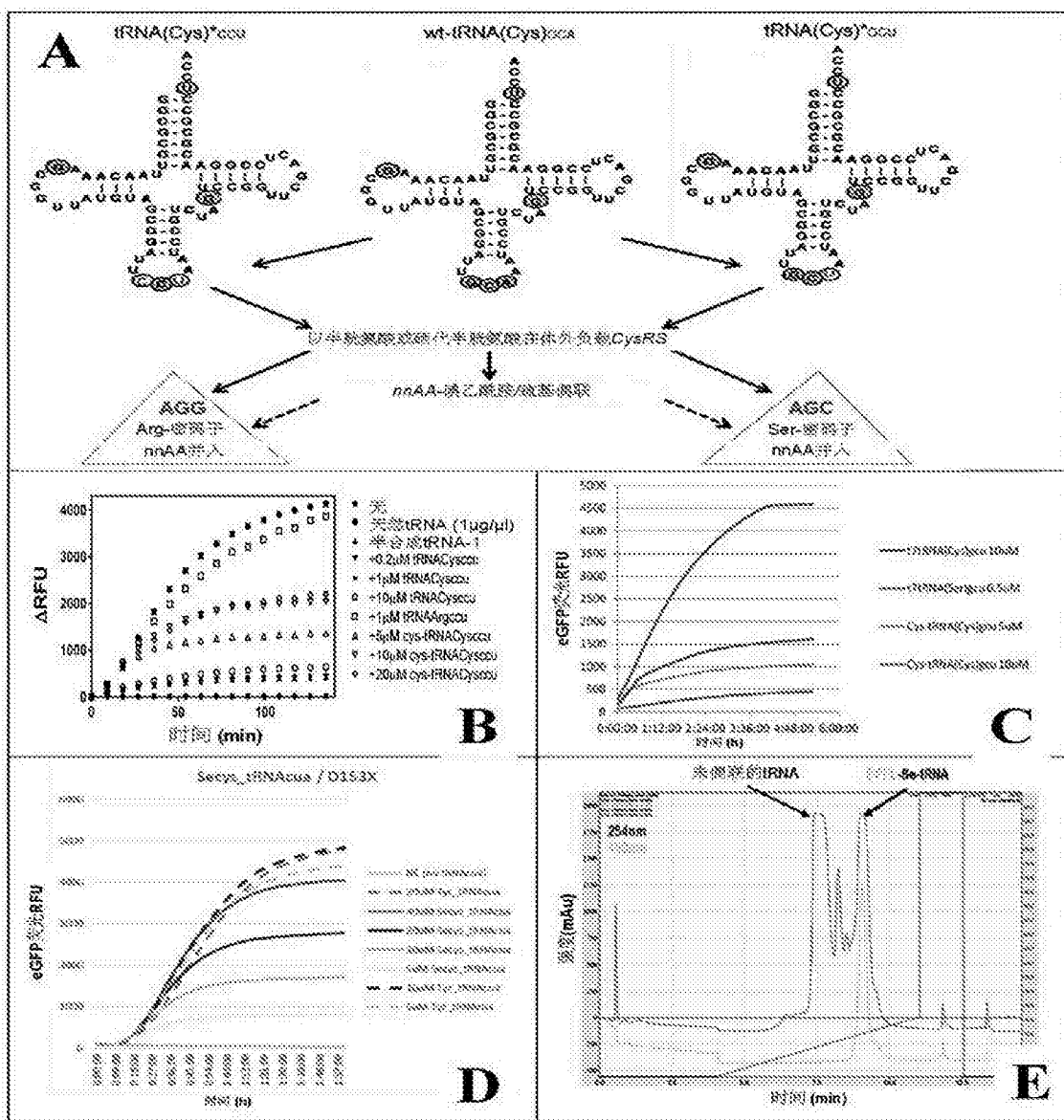


图19

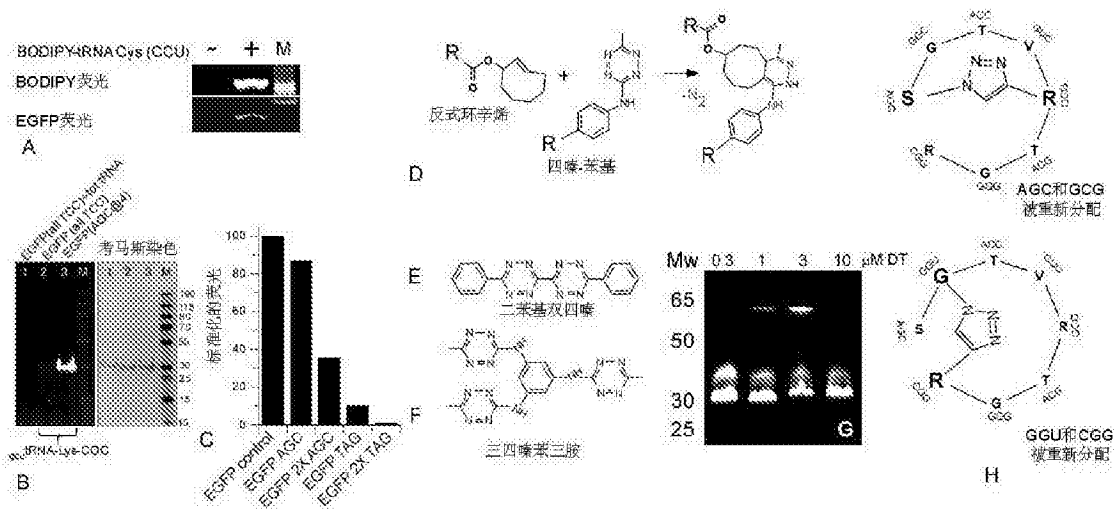


图20

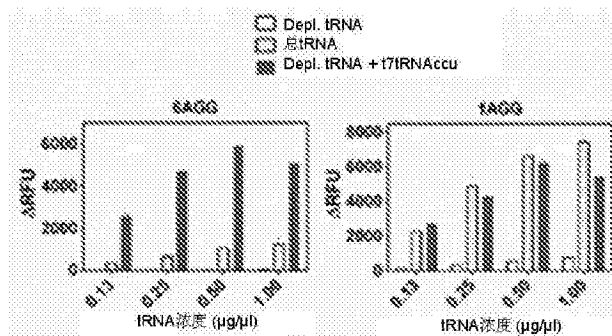


图21

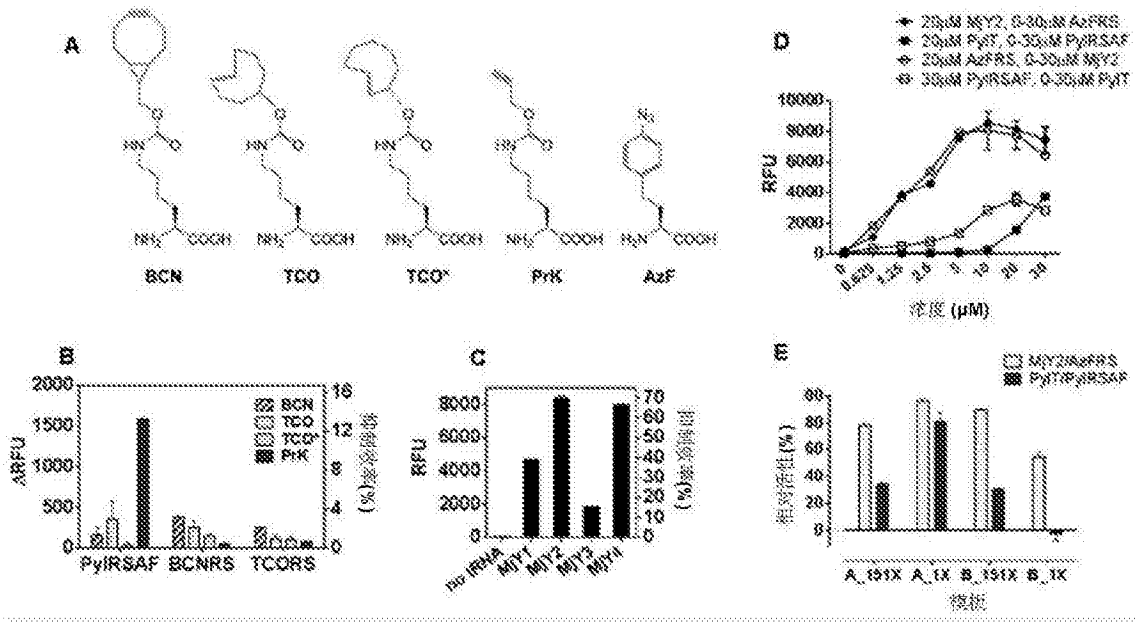


图22

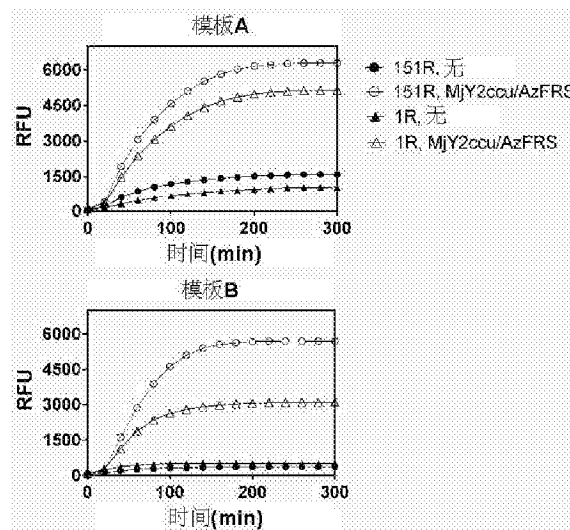


图23

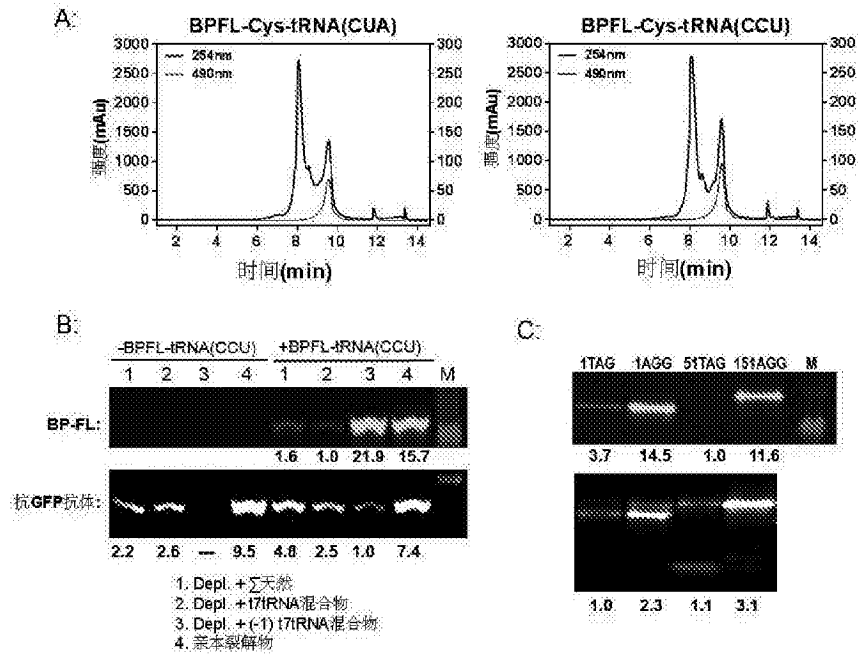


图24

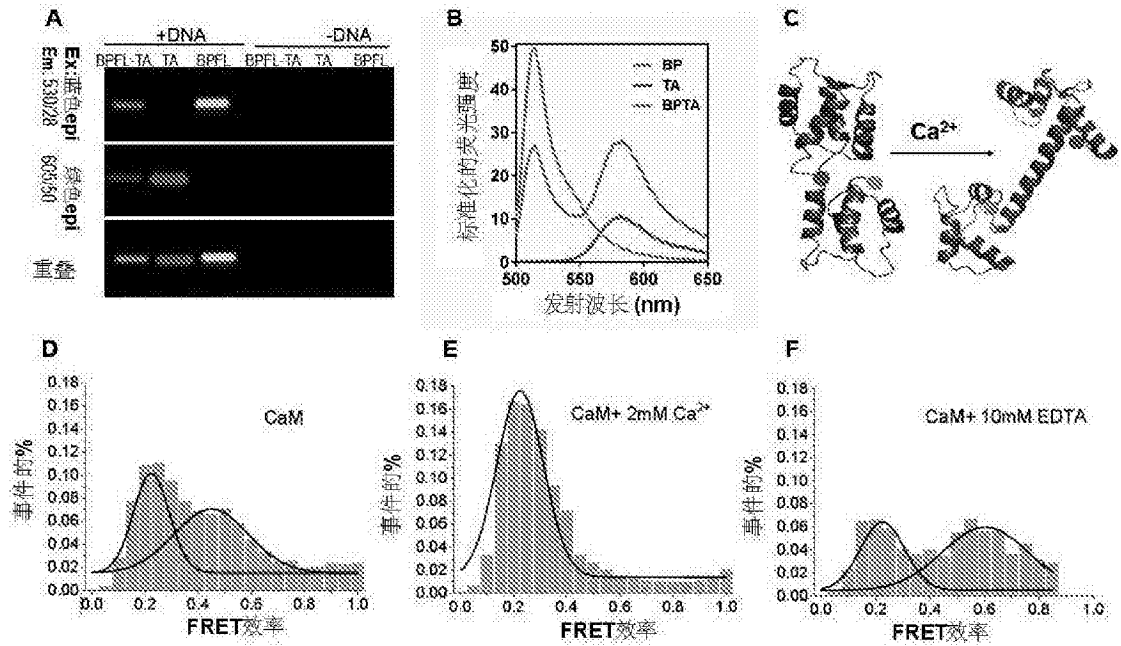


图25

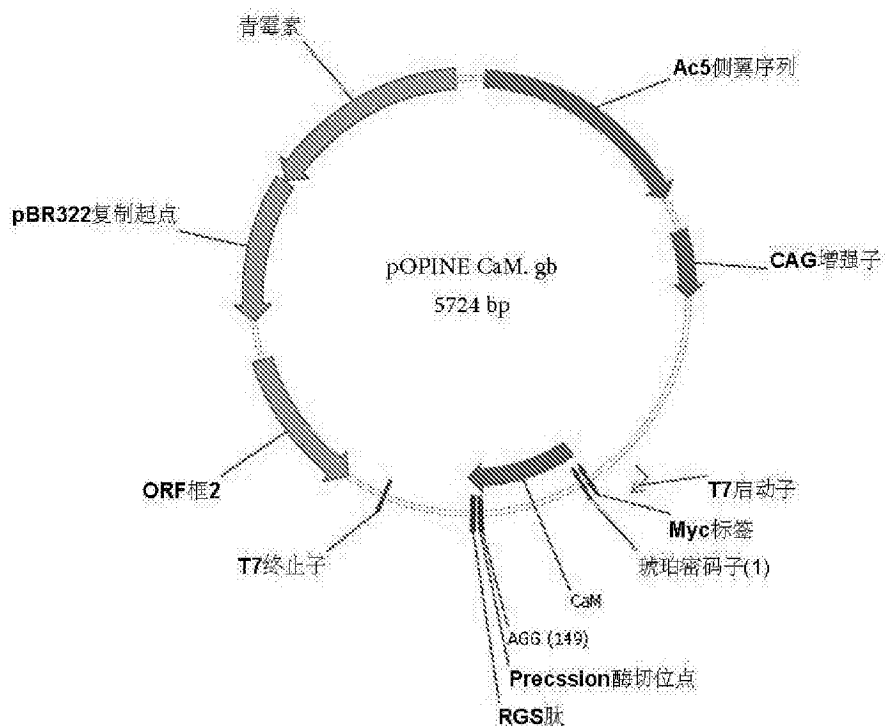


图26

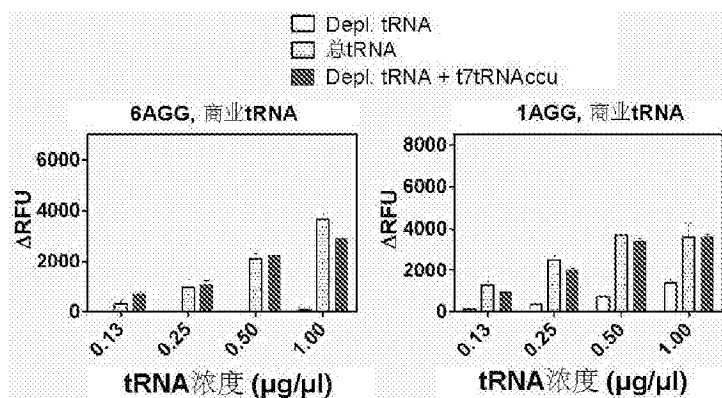


图27

	(1)	1	10	20	30	40	50	60	77															
Miy1_tRNA	(1)	CCG	CCG	TAG	TCAG	AGG	GCAG	ACCG	CGG	ACTCT	TAA	ATCC	GCAT	GGC	AC	GGG	TTCA	AA	TCC	CGT	GG	CG	GG	ACCA
Miy2_tRNA	(1)	CCG	CCG	TAG	TCAG	AGG	GCAG	ACCG	CGG	ACTCT	TAA	ATCC	GCAT	GGC	AC	GGG	TTCA	AA	TCC	CGT	GG	CG	GG	ACCA
Miy3_tRNA	(1)	CCG	CCG	TAG	TCAG	AGG	GCAG	ACCG	CGG	ACTCT	TAA	ATCC	GCAT	GGC	AC	GGG	TTCA	AA	TCC	CGT	GG	CG	GG	ACCA
Miy4_tRNA	(1)	CCG	CCG	TAG	TCAG	AGG	GCAG	ACCG	CGG	ACTCT	TAA	ATCC	GCAT	GGC	AC	GGG	TTCA	AA	TCC	CGT	GG	CG	GG	ACCA

图28

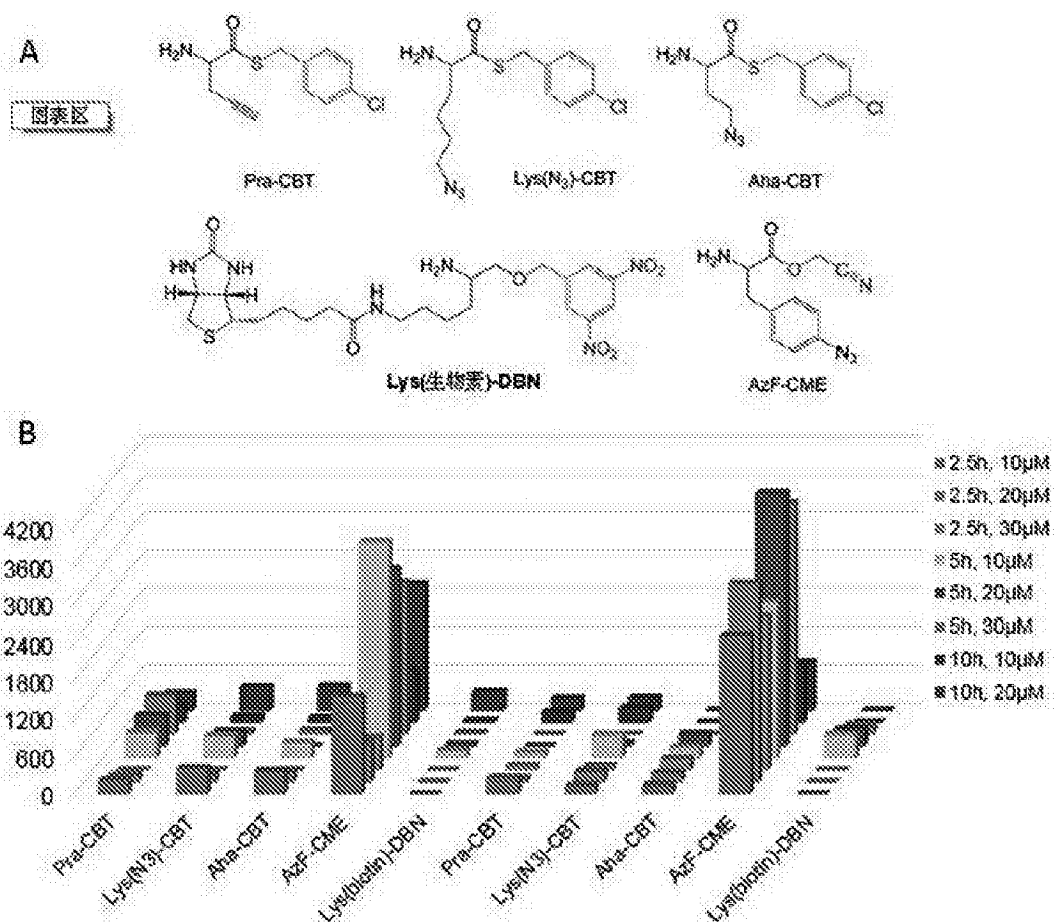


图 29

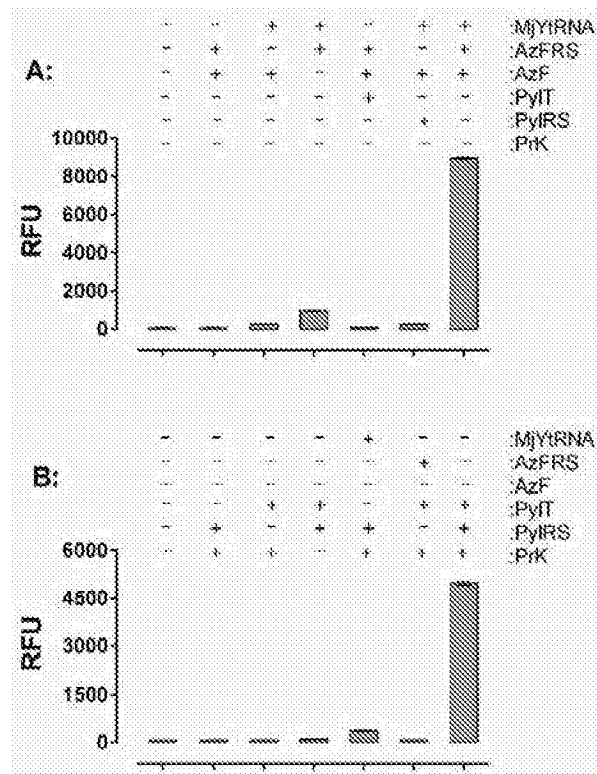


图30

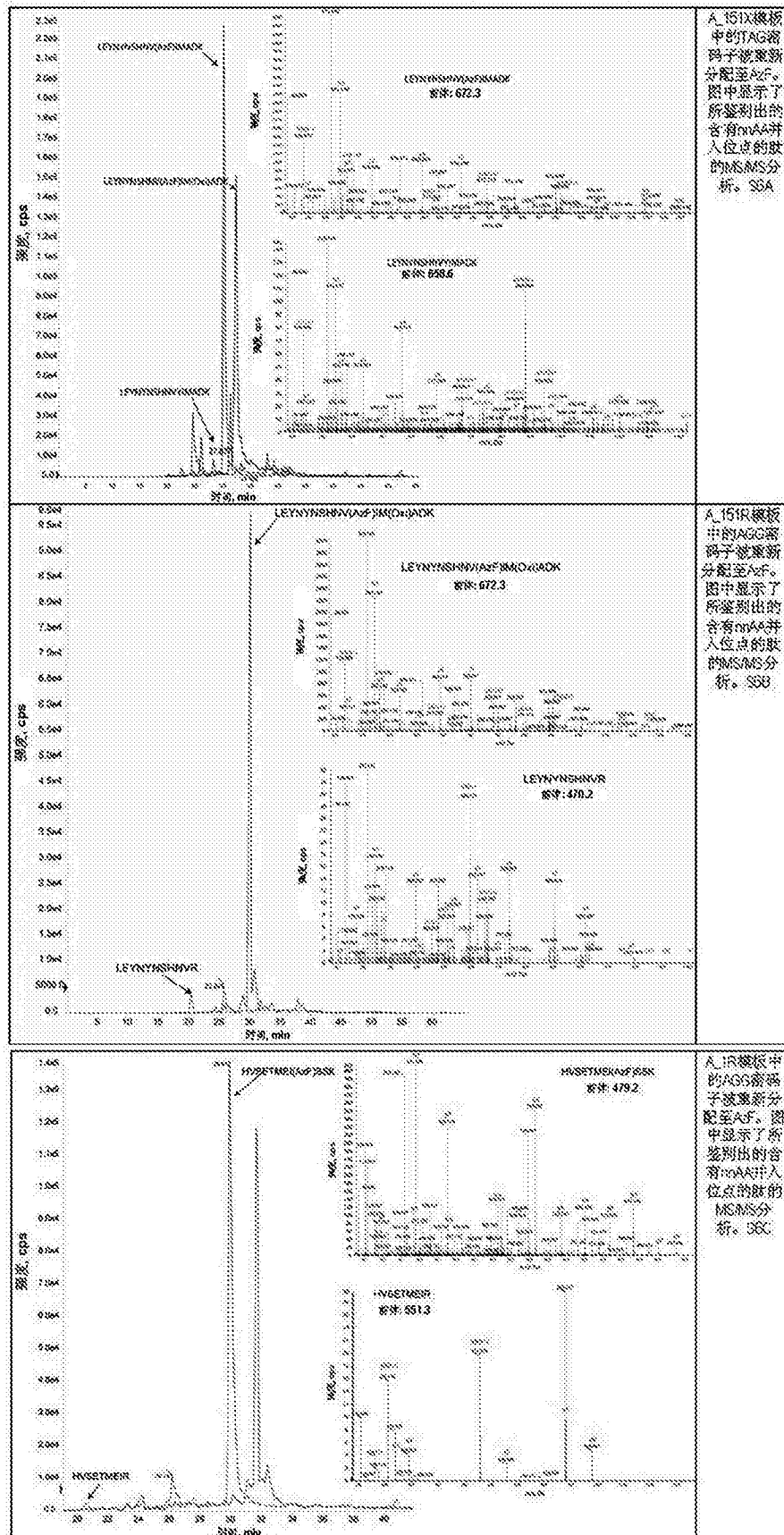


图31

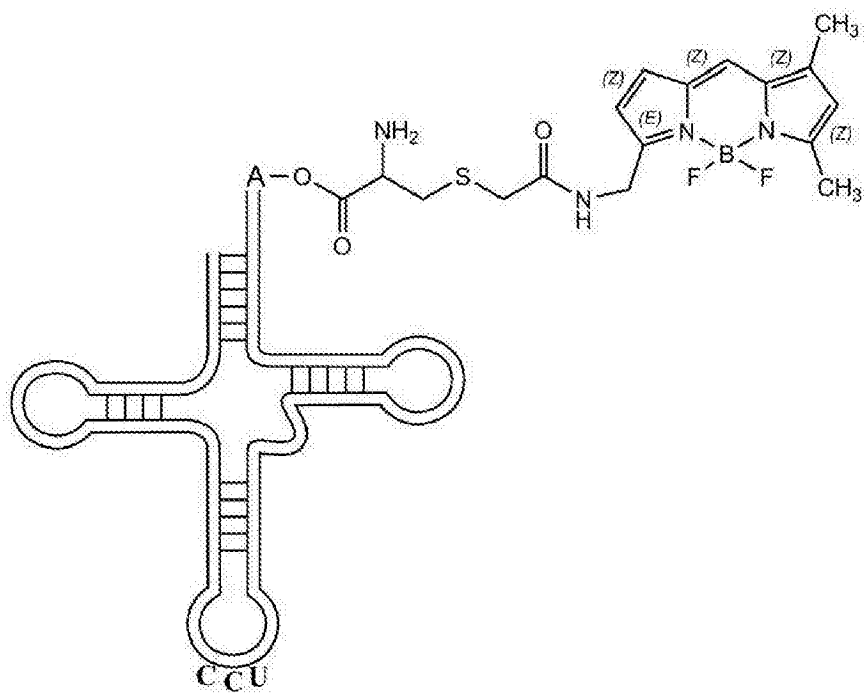


图32

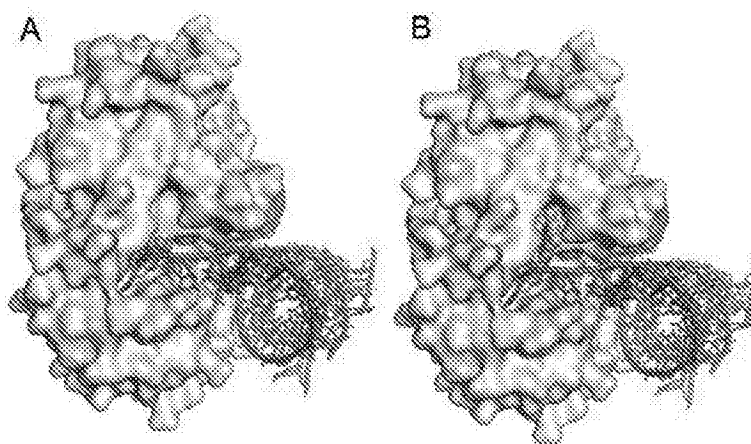


图33

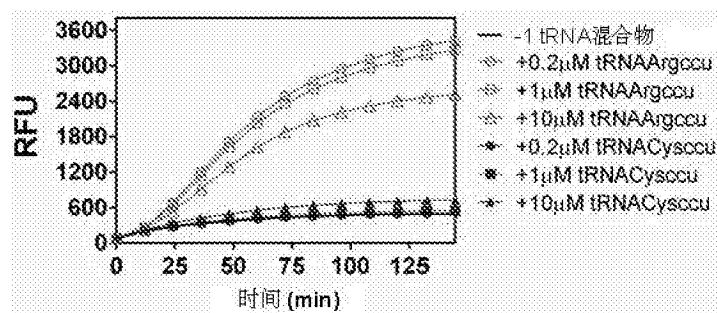


图34

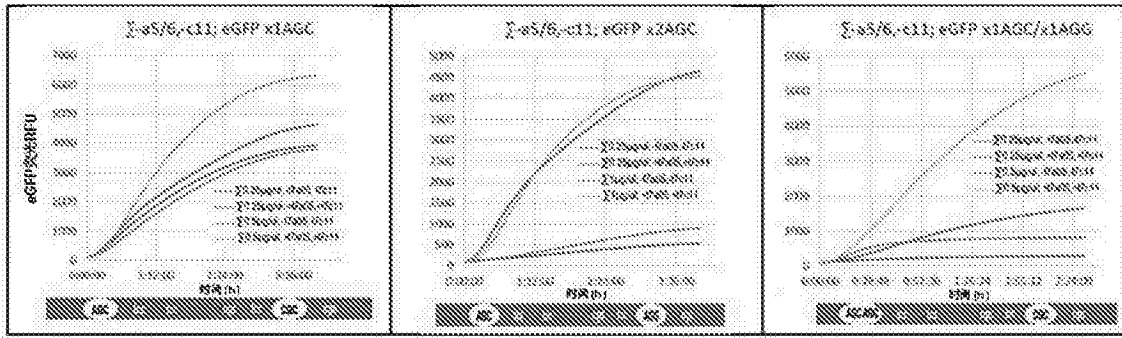


图35

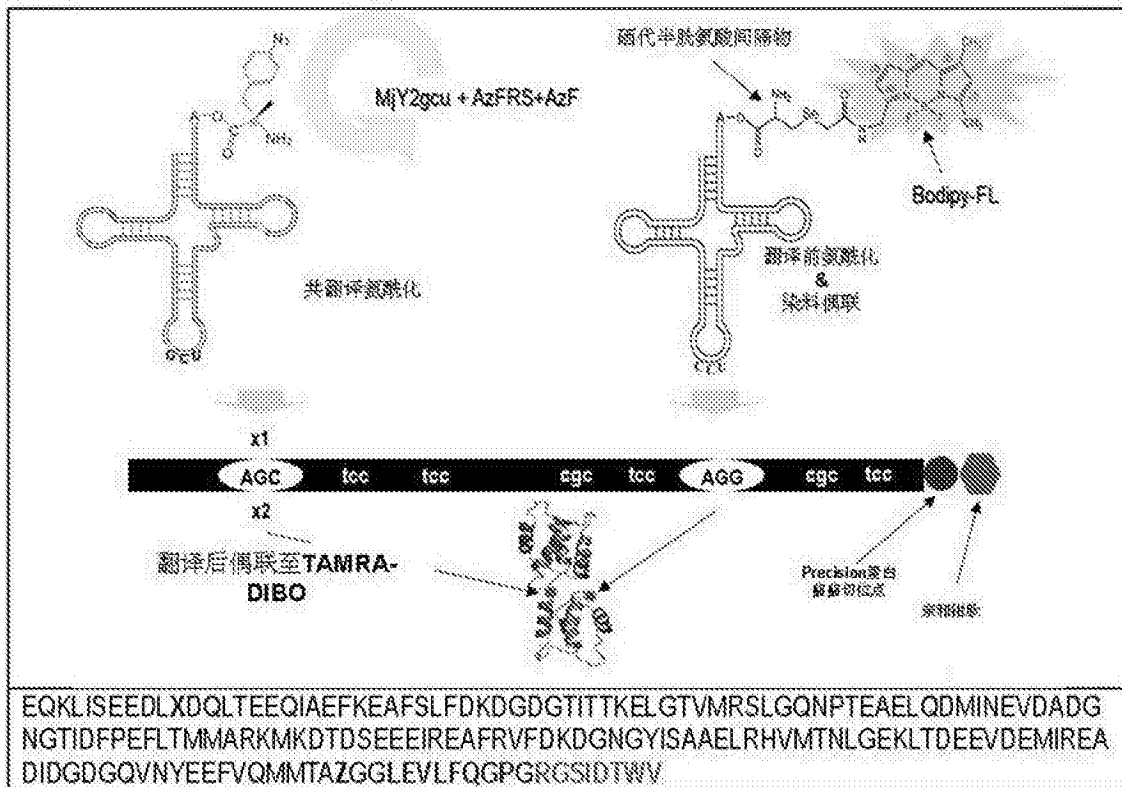


图36

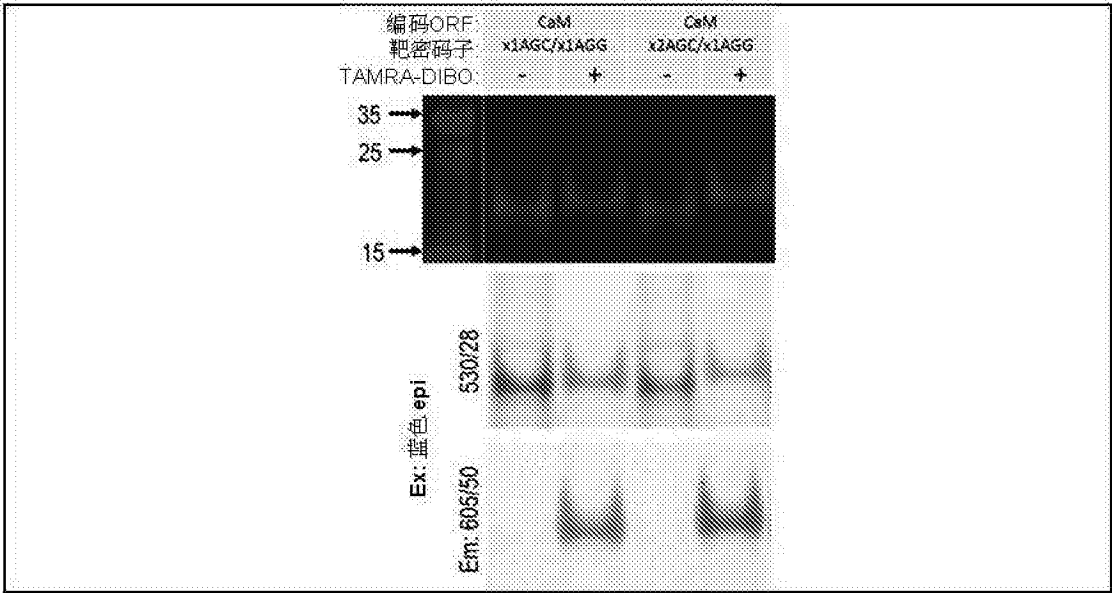


图37