

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 98803511.1

[51] Int. Cl.

C07H 21/02 (2006.01)

C07K 4/00 (2006.01)

C07K 14/00 (2006.01)

C12Q 1/68 (2006.01)

[45] 授权公告日 2006 年 1 月 25 日

[11] 授权公告号 CN 1238366C

[22] 申请日 1998.1.14 [21] 申请号 98803511.1

[30] 优先权

[32] 1997.1.21 [33] US [31] 60/035,963

[32] 1997.11.6 [33] US [31] 60/064,491

[86] 国际申请 PCT/US1998/000807 1998.1.14

[87] 国际公布 WO1998/031700 英 1998.7.23

[85] 进入国家阶段日期 1999.9.20

[71] 专利权人 综合医院公司

地址 美国马萨诸塞州

[72] 发明人 J·W·索斯塔克 R·W·罗伯茨

刘日河

审查员 王青华

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 谭明胜

权利要求书 5 页 说明书 55 页 附图 16 页

[54] 发明名称

利用 RNA - 蛋白融合体筛选蛋白

[57] 摘要

本发明披露了用于筛选蛋白分子的方法和试剂，所述方法和试剂利用了 RNA - 蛋白融合体。

1. 一种用于筛选需要的蛋白的方法，包括以下步骤：

- a) 提供一群候选 RNA 分子，每一个分子包括可操作地连接于一个候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；
- 5 b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和
- c) 筛选所需要的 RNA-蛋白融合体，从而筛选所需要的蛋白。

2. 一种用于筛选编码一种需要的蛋白的 DNA 分子的方法，该方法包
10 括以下步骤：

- a) 提供一群候选 RNA 分子，每一个分子包括可操作地连接于一个候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；
- 15 b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；
- c) 筛选所需要的 RNA-蛋白融合体；和
- d) 由所述融合体的 RNA 部分制备一种编码所需要的蛋白的 DNA 分子。

3. 一种用于筛选相对参考蛋白而言具有改变了的功能的蛋白的方法，该方法包括以下步骤：

- 20 a) 由一群 DNA 模板产生一群候选 RNA 分子，所述候选 DNA 模板各有一个不同于所述参考蛋白编码序列的候选蛋白编码序列，所述 RNA 分子各自包括可操作地连接于所述候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；
- b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和
- 25 c) 筛选具有改变了的功能的 RNA-蛋白融合体，从而筛选具有改变了的功能的蛋白。

4. 一种用于筛选编码一种相对参考蛋白而言具有改变了的功能的蛋白的 DNA 分子的方法，该方法包括以下步骤：

- 30 a) 由一群候选 DNA 模板生产一群候选 RNA 分子，所述候选 DNA 模板各自具有不同于所述参考蛋白编码序列的候选蛋白编码序列，所述 RNA

分子各自包括可操作地连接于所述候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且各自在其 3' 末端可操作地连接一个肽受体；

b) 体外或原位翻译所述候选蛋白编码序列，以产生一群 RNA-蛋白融合体；

5 c) 筛选具有改变了的功能的 RNA-蛋白融合体；和

d) 由所述融合体的 RNA 部分制备一种编码具有改变了的功能的蛋白的 DNA 分子。

5. 一种用于筛选需要的 RNA 的方法，该方法包括以下步骤：

a) 提供一群候选 RNA 分子，其中的每一个 RNA 分子包括可操作地连接于所述候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；

b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和

15 c) 筛选所需要的 RNA-蛋白融合体，从而筛选所需要的 RNA。

6. 如权利要求 1-5 中任一项的方法，其中，所述肽受体是嘌呤霉素。

7. 如权利要求 1-5 中任一项的方法，其中，所述每一个候选 RNA 分子还包括一个间歇序列或还包括一个共价连接于该 RNA 分子 3' 末端的 DNA 或 DNA 类似物序列。

20 8. 如权利要求 1-5 中任一项的方法，其中，所述候选 RNA 分子群至少包括 10^{13} 个不同的 RNA 分子。

9. 如权利要求 1-5 中任一项的方法，其中，所述体外翻译反应是在由真核细胞或其部分制备的裂解物中进行的。

25 10. 如权利要求 9 中的方法，其中，所述体外翻译反应是在网织红细胞裂解物中进行的。

11. 如权利要求 9 中的方法，其中，所述体外翻译反应是在小麦胚裂解物中进行的。

12. 如权利要求 1-5 中任一项的方法，其中，所述体外翻译反应是在由细菌细胞或其部分制备的裂解物中进行的。

30 13. 如权利要求 1-5 中任一项的方法，其中，所述筛选步骤包括将所述需要的蛋白结合到一种固定的结合配偶体上。

14. 如权利要求 1-5 中任一项的方法，其中，所述筛选步骤包括测定

所述所需蛋白的功能活性。

15. 如权利要求 2 或 4 的方法，其中，所述 DNA 分子是扩增的。

16. 如权利要求 1、3 或 5 的方法，其中，所述方法还包括重复步骤 (a) - (c)。

5 17. 如权利要求 2 或 4 的方法，其中，所述方法还包括由所述 DNA 分子转录 RNA 分子，并重复步骤 (a) - (d)。

18. 如权利要求 1-5 中任一项的方法，其中，所述 RNA 通过一个酰胺键共价连接于所述 RNA-蛋白融合体中的蛋白上。

10 19. 如权利要求 1-5 中任一项的方法，其中，所述 RNA 共价连接于所述 RNA-蛋白融合体中的蛋白上，所述共价键能够抗核糖体的裂解。

20. 如权利要求 1-5 中任一项的方法，其中，在体外翻译步骤之后，在有 50-100mM Mg²⁺的条件下进行培养。

21. 如权利要求 1-5 中任一项的方法，其中，所述 RNA-蛋白融合体还包括一个靠近所述肽受体的核酸或核酸类似物序列，它能提高其柔性。

15 22. 一种通过权利要求 1-5 中任一项的方法筛选的 RNA-蛋白融合体，其中 RNA-蛋白融合体中的蛋白部分通过肽受体共价连接于 RNA-蛋白融合体中的 RNA 部分上，并且所述蛋白部分是由 RNA-蛋白融合体中的 RNA 部分编码的。

20 23. 一种分子，包括通过酰胺键共价连接于蛋白上的核糖核酸，其中酰胺键位于所述蛋白和可操作地连接于该核糖核酸的肽受体之间。

24. 如权利要求 23 分子，其中，所述蛋白是由所述核糖核酸编码的，并且所述肽受体可操作地连接于所述核糖核酸的 3' 末端。

25 25. 如权利要求 23 分子，其中，所述核糖核酸是 mRNA。

26. 一种分子，包括通过肽受体共价连接于蛋白上的核糖核酸，所述蛋白完全是由所述核糖核酸编码的，其中所述肽受体连接于所述核糖核酸的 3' 末端。

27. 如权利要求 26 分子，其中，所述核酸是 mRNA。

28. 一种分子，包括通过肽受体共价连接于蛋白上的核糖核酸，其中所述肽受体选自嘌呤霉素、与腺嘌呤连接的氨基酸和与腺嘌呤样化合物连接的氨基酸，所述共价键能抗核糖体的裂解。

30 29. 如权利要求 28 分子，其中，所述核糖核酸是 mRNA。

30. 一种核糖核酸，包括可操作地连接于候选蛋白编码序列上的翻译

起始序列和起始密码子，所述核糖核酸于所述候选蛋白编码序列 3' 末端共价连接于肽受体上。

31. 一种筛选需要的蛋白或需要的 RNA 的方法，该方法包括以下步骤：

5 (a) 提供一群候选 RNA 分子，其中的每一个 RNA 分子包括可操作地连接于一个候选蛋白编码序列上的一个翻译起始序列和一个起始密码子，而且各自通过所述候选蛋白编码序列的 3' 末端可操作地连接于一个肽受体上；

10 (b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；

(c) 让所述 RNA-蛋白融合体群体与一种对所述 RNA-蛋白融合体的 RNA 部分或蛋白部分专一的结合配偶体接触，接触条件为基本上能将所述结合配偶体- RNA-蛋白融合体复合体与所述群体的未结合的成员分离；

(d) 从所述复合体上释放出结合的 RNA-蛋白融合体；和

15 (e) 让来自步骤 (d) 的 RNA-蛋白融合体群体与一种对所需要的 RNA-蛋白融合体的蛋白部分具有专一性的结合配偶体接触，接触条件为基本上能将所述结合配偶体- RNA-蛋白融合体的复合体与所述群体的未结合的成员分离，从而筛选需要的蛋白和需要的 RNA。

32. 如权利要求 31 的方法，其中，所述方法还包括重复步骤 (a) -
20 (e)。

33. 如权利要求 31 的方法，其中，所述肽受体是嘌呤霉素。

34. 如权利要求 31 的方法，其中，所述每一个候选 RNA 分子还包括一个间歇序列或还包括一个共价连接于所述 RNA 分子的 3' 末端的 DNA 或 DNA 类似物序列。

25 35. 如权利要求 31 的方法，其中，所述候选 RNA 分子群体至少包括 10^{13} 种不同的 RNA 分子。

36. 如权利要求 31 的方法，其中，所述体外翻译反应是在由真核细胞或其部分制备的裂解物中进行的。

30 37. 如权利要求 36 的方法，其中，所述体外翻译反应是在网织红细胞裂解物或小麦胚裂解物中进行的。

38. 如权利要求 31 的方法，其中，所述体外翻译反应是在由原核细胞或其部分制备的提取物中进行的。

39. 如权利要求 31 的方法，其中，将所述结合配偶体中的至少一个固定在固体支持物上。

40. 如权利要求 31 的方法，其中，在体外翻译步骤之后，在有 50-100mM Mg²⁺的条件下进行培养。

5 41. 如权利要求 31 的方法，其中，所述 RNA-蛋白融合体还包括一个靠近所述肽受体的核酸或核酸类似物序列，它能提高其柔性。

42. 一种小嵌片，包括一系列固定化的单链核酸，所述核酸能与 RNA-蛋白融合体杂交，其中 RNA-蛋白融合体中的蛋白部分通过肽受体共价连接于 RNA 部分上，并且所述蛋白部分是由所述 RNA 部分编码的。

10 43. 如权利要求 42 的小嵌片，其中，所述蛋白是由所述 RNA 编码的。

利用 RNA-蛋白融合体筛选蛋白

发明背景

5 本发明涉及蛋白筛选方法。

本发明的完成得到了政府的支持，授权号为 F32 GM17776-01 和 F32 GM17776-02。政府拥有对本发明的某些权利。

现在已经存在根据其功能提取 RNA 和 DNA 分子的方法。例如， Ellington 和 Szostak (自然 346: 818 (1990)；和自然 355: 850 (1992)) 和 Tuerk 和 Gold (科学 249: 505 (1990)；和分子生物学杂志 222: 739 (1991)) 的实验业已证实，通过反复筛选和扩增可以从复杂的分子库中提取到极少的（即少于 $1/10^{13}$ ）具有所需特性的核酸分子。与传统遗传学筛选方法相比，上述方法的优点是 (i) 可以筛选极大的候选库 ($>10^{15}$)，(ii) 与宿主生活力和体内条件无关，和 (iii) 即使不存在体内遗传学筛选，也可以进行筛选。业已证实了所述方法在体外筛选特定的新型 RNA 和 DNA 序列的能力，所述新型 RNA 和 DNA 序列具有非常特异的蛋白结合功能(例如，参见 Tuerk 和 Gold 科学 249: 505 (1990)；Irvine 等，分子生物学杂志 222: 739(1991); Oliphant 等，分子细胞生物学 9: 2944(1989); Blackwell 等，科学 250: 1104 (1990)；Pollock 和 Treisman, 核酸研究 18: 6197(1990); Thiesen 和 Bach, 核酸研究 18: 3203(1990); Bartel 等，细胞 57: 529 (1991); Stormo 和 Yoshioka, Proc. Natl. Acad. Sci. USA 88: 5699 (1991)；和 Bock 等，自然 355: 564 (1992))，小分子结合功能 (Ellington 和 Szostak, 自然 346: 818 (1990)；25 Ellington 和 Szostak, 自然 355: 850(1992))，和催化功能 (Green 等，自然 347: 406 (1990)；Robertson 和 Joyce, 自然, 34: 467 (1990)；Beaudry 和 Joyce, 科学 257: 635 (1992)；Bartel 和 Szostak, 科学 261: 1411 (1993)；Lorsch 和 Szostak, 自然 371: 31-36 (1994)；Cuenoud 和 Szostak, 自然 375: 611-614 (1995)；30 Chapman 和 Szostak, 化学和生物学 2: 325-333 (1995) 和 Lohse 和 Szostak, 自然 381: 442-444 (1996))。用于筛选和扩增蛋白的类似方法尚未得到证实。

发明概述

本发明的目的是，将所述体外筛选和体外进化的原理应用于蛋白。本发明有利于从由部分或完全随机的氨基酸序列组成的大库中
5 提取具有所需特性的蛋白。另外，本发明通过将所述 mRNA 编码序列共价连接到所述蛋白分子上解决了回收和扩增所需蛋白序列信息的问题。

一般，本发明的方法包括体外或原位转录/翻译方案，该方法能制备与其自身的 mRNA 3' 端共价连接的蛋白，即 RNA-蛋白融合体。这
10 一目的是通过合成和体外或原位翻译在其 3' 末端结合有肽受体的 mRNA 分子而实现的。一种优选的肽受体是嘌呤霉素，一种加在生长肽链 C-末端并终止翻译的核苷酸类似物。在一种优选设计中，在所述信使和肽受体之间包括一个 DNA 序列，将该序列设计成导致核糖体停留在开放读框的末端，以便在所述肽基-tRNA 键水解之前提供用于
15 所述肽受体（例如，嘌呤霉素）接受新生肽链的额外时间。

如果需要，可以对所得到的 RNA-蛋白融合体进行反复地筛选和扩增，因为可以通过逆转录和扩增（例如，通过 PCR 扩增以及任何其它扩增技术，包括基于 RNA 的扩增技术，如 3SR 或 TSA）回收所述蛋白序列信息。然后可以对扩增的核酸进行转录、修饰、以及体外或原位翻译，以产生用于下一轮筛选的 mRNA-蛋白融合体。所述进行多个轮次的筛选和扩增的能力可以富集和提取极其罕见的分子，例如，从由 10^{15} 个成员组成的库中筛选出一个需要的分子。这样又可以提取新的或改进的蛋白，该蛋白实际上能特异识别所有目标，或能催化所需要的化学反应。

因此，第一方面，本发明的特征是一种用于筛选需要的蛋白的方法，该方法包括以下步骤：(a) 提供一群候选 RNA 分子，每一个分子包括一个翻译起始序列和一个可操作地连接于一个候选蛋白编码序列上的起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；(b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和(c) 筛选所需要的 RNA-蛋白融合体，从而筛选所需要的蛋白。

因此，在一个相关的方面，本发明的特征是一种用于筛选编码一

一种需要的蛋白的 DNA 分子的方法，该方法包括以下步骤：(a) 提供一群候选 RNA 分子，每一个分子包括一个翻译起始序列和一个可操作地连接于一个候选蛋白编码序列上的起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；(b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；(c) 筛选所需要的 RNA-蛋白融合体；和 (d) 由所述融合体的 RNA 部分制备一种编码所需要的蛋白的 DNA 分子。

因此，在另一个相关方面，本发明的特征是一种用于筛选相对参考蛋白而言具有改变了的功能的蛋白的方法，该方法包括以下步骤：(a) 由一群 DNA 模板产生一群候选 RNA 分子，所述候选 DNA 模板各有一个不同于所述参考蛋白编码序列的候选蛋白编码序列，所述 RNA 分子各自包括一个翻译起始序列和一个可操作地连接于所述候选蛋白编码序列上的起始密码子，而且各自在其 3' 末端可操作地连接一个肽受体；(b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和 (c) 筛选具有改变了的功能的 RNA-蛋白融合体，从而筛选具有改变了的功能的蛋白。

在另一个相关方面，本发明的特征是一种用于筛选编码一种相对参考蛋白而言具有改变了的功能的蛋白的 DNA 分子的方法，该方法包括以下步骤：(a) 由一群候选 DNA 模板产生一群候选 RNA 分子，所述候选 DNA 模板各自具有不同于所述参考蛋白编码序列的候选蛋白编码序列，所述 RNA 分子各自包括一个翻译起始序列和一个可操作地连接于所述候选蛋白编码序列上的起始密码子，而且各自在其 3' 末端可操作地连接一个肽受体；(b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；(c) 筛选具有改变了的功能的 RNA-蛋白融合体；和 (d) 由所述融合体的 RNA 部分制备一种编码具有改变了的功能的蛋白的 DNA 分子。

在另一个相关方面，本发明的特征是一种用于筛选需要的 RNA 的方法，该方法包括以下步骤：(a) 提供一群候选 RNA 分子，其中的每一个 RNA 分子包括一个翻译起始序列和一个可操作地连接于所述候选蛋白编码序列上的起始密码子，而且在所述每一个候选蛋白编码序列的 3' 末端可操作地连接一个肽受体；(b) 体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；和 (c) 筛

选所需要的 RNA-蛋白融合体，从而筛选所需要的 RNA。

在上述方法的优选实施方案中，所述肽受体是嘌呤霉素；所述每一个候选 RNA 分子还包括间歇序列（pause sequence）或还包括一个共价结合于所述 RNA 的 3' 末端的 DNA 或 DNA 类似物序列；所述候选 RNA 分子群体至少包括 10^9 ，优选至少 10^{10} ，更优选至少 10^{11} 、 10^{12} 或 10^{13} ，以及最优选至少 10^{14} 个不同的 RNA 分子；所述体外翻译是在用真核细胞或其部分制备的裂解物中进行的（而且，可以是，例如，在网织红细胞裂解物或小麦胚裂解物中进行的）；所述体外翻译反应是在由原核细胞（例如，大肠杆菌）或其部分制备的提取物中进行的；所述筛选步骤包括将所需要的蛋白结合在一种固定化的结合配偶体上；所述筛选步骤包括测定所述需要的蛋白的功能活性；所述 DNA 分子被扩增；所述方法还包括重复上述筛选方法的步骤；所述方法还包括由所述 DNA 分子转录 RNA 分子，并重复步骤（a）-（d）；在所述体外翻译步骤之后，该方法还包括一个在有 $50\text{--}100\text{mM Mg}^{2+}$ 的条件下进行的温育步骤；而且，所述 RNA-蛋白融合体还包括一个靠近所述肽受体的核酸或核酸类似物序列，它能提高其柔性。

在另一个相关方面，本发明的特征是通过本发明任一种方法筛选的 RNA-蛋白融合体；一个通过酰胺键共价连接于一个氨基酸序列上的核糖核酸，所述氨基酸序列是由所述核糖核酸编码的；以及一个包括一个翻译起始序列和一个可操作地连接于候选蛋白编码序列上的起始密码子的核糖核酸，所述核糖核酸通过所述候选蛋白编码序列的 3' 末端可操作地连接于一种肽受体（例如嘌呤霉素）上。

第二方面，本发明的特征是一种用于通过富集一种序列库筛选需要的蛋白或需要的 RNA 的方法。该方法包括以下步骤：（a）提供一群候选 RNA 分子，其中的每一个 RNA 分子包括一个翻译起始序列和一个可操作地连接于一个候选蛋白编码序列上的起始密码子，而且各自通过所述候选蛋白编码序列的 3' 末端可操作地连接于一个肽受体上；（b）体外或原位翻译所述候选蛋白编码序列，以产生一群候选 RNA-蛋白融合体；（c）让所述 RNA-蛋白融合体群体与一种对所述 RNA-蛋白融合体的 RNA 部分或蛋白部分专一的结合配偶体接触，接触条件为基本上能将所述结合配偶体-RNA-蛋白融合体的复合体与所述群体的未结合的成员分离；（d）从所述复合体上释放出结合的 RNA-蛋

白融合体；和 (e) 让来自步骤 (d) 的 RNA-蛋白融合体群体与一种对所需要的 RNA-蛋白融合体的蛋白部分具有专一性的结合配偶体接触，接触条件为基本上能将所述结合配偶体-RNA-蛋白融合体的复合体与所述群体的未结合的成员分离，从而筛选需要的蛋白和需要的 5 RNA。

在优选实施方案中，所述方法还包括重复步骤 (a)-(e)。另外，在重复进行以上步骤时，可以任何顺序使用相同或不同的结合配偶体，以便选择性地富集需要的 RNA-蛋白融合体。在另一种优选实施方案中，步骤 (d) 涉及使用对所述需要的融合体的蛋白部分具有专 10 一性的结合配偶体（例如，单克隆抗体）。该步骤优选在对所述融合体的 RNA 部分进行逆转录之后进行，以制备一种能编码所需要的蛋白的 DNA。如果需要，可以提取该 DNA 和/或进行 PCR 扩增。可将这种富集技术用于筛选需要的蛋白或用于筛选相对参考蛋白而言具有改变了的功能的蛋白。

在富集方法的其它优选实施方案中，所述肽受体是嘌呤霉素；所述每一个候选 RNA 分子还包括间歇序列或还包括一个共价结合于所述 RNA 的 3' 末端的 DNA 或 DNA 类似物序列；所述候选 RNA 分子群体至少包括 10^9 ，优选至少 10^{10} ，更优选至少 10^{11} 、 10^{12} 或 10^{13} ，以及最 20 优选至少 10^{14} 个不同的 RNA 分子；所述体外翻译是在用真核细胞或其部分制备的裂解物中进行的（而且，可以是，例如，在网织红细胞裂解物或小麦胚裂解物中进行的）；所述体外翻译反应是在由原核细胞（例如，大肠杆菌）或其部分制备的提取物中进行的；扩增所述 DNA 分子；将所述结合配偶体的至少一种固定在一种固体支持物上；在所述体外翻译步骤之后，所述方法还包括一个在有 $50\text{--}100\text{mM Mg}^{2+}$ 的条件下进行的温育步骤；而且，所述 RNA-蛋白融合体还包括一个靠近所述肽受体的核酸或核酸类似物序列，它能提高其柔性。 25

在一个相关方面，本发明的特征是用于实施本文所披露的任一种筛选方法的试剂盒。

在第三方面，也是最后一个方面，本发明的特征是一种小嵌片， 30 该嵌片包括一系列固定化的单链核酸，所述核酸能与 RNA-蛋白融合体杂交。优选地，所述 RNA-蛋白融合体的蛋白成分是由所述 RNA 编码的。

在本文中，“群体”是指一个以上的分子（例如，一个以上的 RNA、DNA、或 RNA-蛋白融合体分子）。因为本发明的方法有利于在必要时由大量的候选分子开始筛选，本发明的“群体”优选是指 10⁹ 以上的分子，更优选 10¹¹、10¹² 或 10¹³ 以上的分子，最优选 10¹³ 以上的分子。

“筛选”是指将一种分子从一个群体中的其它分子中基本上分离出来。在本文中，在经过一个筛选步骤之后，该“筛选”步骤能使所需要的分子相对一个群体中不需要的分子富集 2 倍，优选 30 倍，更优选 100 倍，最优选 1000 倍。正如本文所述的，筛选步骤可以重复进行任意的次数，而且，在一种特定的方法中可以进行不同类型筛选步骤的组合。

“蛋白”是指所有通过一个或几个肽键连接的两个或两个以上天然存在的或修饰过的氨基酸。在本文中，“蛋白”和“肽”可以互换使用。

“RNA”是指两个或两个以上共价结合的天然存在的或修饰过的核糖核苷酸的序列。该术语所包括的修饰 RNA 的一种例子是硫代磷酸 RNA.

“翻译起始序列”是指能够提供功能性核糖体进入位点的所有序列。在细菌系统中，该序列有时被称为 SD 序列。

“起始密码子”是指指示一种蛋白编码序列开始的三个碱基。一般，这些碱基是 AUG (或 ATG)；不过，任何能以这种方式使用的其它碱基三联体均可取代。

与一种肽受体的“共价结合”是指所述肽受体通过一个共价键直接地或者通过另一个共价结合的序列（例如，相应于一个间歇位点的 DNA）间接地连接于“蛋白编码序列”上。

“肽受体”是指能通过核糖体肽基转移酶功能的催化活性添加到一种生长肽链的 C-末端的所有分子。通常，这种分子包括 (i) 一个核苷酸或核苷酸-样部分（例如，腺苷或腺苷类似物（在 N-6 氨基位置上二-甲基化是可以接受的）），(ii) 一个氨基酸或氨基酸-样部分（例如，20 种 D-或 L-氨基酸的任一种或其所有的氨基酸类似物（例如，由 Ellman 等在酶学方法 202: 301, 1991 中所披露的 O-甲

基酪氨酸或任一种类似物），和（iii）两者之间的连键，（例如，位于3'位置或不太理想的是位于2'位置的酯键、酰胺键、或酮键）；优选地，所述连键不会明显影响具有天然核糖核苷酸构象的环的折叠。肽受体还可以具有一个亲核体，该亲核体可以是（但不限于）氨基，羟基或巯基，另外，肽受体可以由核苷酸模拟物、氨基酸模拟物，或组合的核苷酸和/或氨基酸结构的模拟物组成。

肽受体位于一种蛋白编码序列的“3'末端”是指该肽受体分子位于所述蛋白编码序列的最终的密码子之后。该术语包括，但不限于精确位于所述蛋白编码序列的3'末端的肽受体分子，以及通过间插编码或非编码序列（例如，相当于间歇位点的序列）与所述最终的密码子分离的肽受体分子。该术语还包括这样的结构，其中，编码或非编码序列位于所述肽受体分子之后（即位于所述肽受体分子的3'末端）。另外，该术语包括，但不限于共价结合于（直接地或通过间插核酸序列间接地）所述蛋白编码序列上的肽受体分子，以及通过某种非共价方式，例如，通过利用第二种核酸序列杂交的方式与所述蛋白编码序列结合的肽受体分子，所述第二种核酸序列结合于所述蛋白编码序列的3'末端或接近该末端，而且，它本身结合于一种肽受体分子。

“改变了的功能”是指一种分子的功能的任何定性或定量变化。“间歇序列”是指能导致核糖体降低或停止其翻译速率的核酸序

20 列。

本文所说的“结合配偶体”是指对一种需要的RNA-蛋白融合体的一部分具有特异的、共价的或非共价的亲和力的所有分子。结合配偶体的例子包括，但不限于抗原/抗体对、蛋白/抑制剂对、受体/配体对（例如，细胞表面受体/配体对，如激素受体/肽激素对）、酶/底物对（例如，激酶/底物对）、凝集素/碳水化合物对、寡聚蛋白或杂聚蛋白凝聚体、DNA结合蛋白/DNA结合位点对、RNA/蛋白对、核酸双链体、异源双链体、或连接的链，以及能够与RNA-蛋白融合体的任意部分形成一个或几个共价或非共价键（例如，二硫键）的所有分子。结合配偶体包括，但不限于图2中所示的筛选基序中的任一种。

“固体支持物”是指，但不限于可以直接或间接地（例如，通过诸如其它抗体或蛋白A的其它结合配偶体中间物）结合一种亲和性复

合体、或者可以埋入亲和性复合体（例如，通过一种受体或通道）的任何柱（或柱材料）、玻璃珠、试管、微量滴定皿、固体颗粒（例如，琼脂糖或琼脂糖凝胶）、小嵌片（例如，硅、硅-玻璃、或金嵌片）、或膜（例如，脂质体或小泡）。

5 本发明具有多种突出的优点，首先，它是这一类型的用于筛选和扩增蛋白的第一个例子。该技术克服了由于需要回收相应于需要的、提取的蛋白的核苷酸序列而产生的障碍（因为只有核酸能够复制）。具体地讲，很多现有方法都可以通过体内步骤从部分或全部随机化的库中提取蛋白。这种类型的方法包括单克隆抗体技术（Millstein,
10 美国科学 243: 66(1980) 和 Schultz 等, J. Chem. Engng. News 68:26
(1990)），噬菌体展示（Smith, 科学 228: 1315(9185); Parmley
和 Smith, 基因 73: 305(1988); 和 McCafferty 等, 自然 348: 552
(1990)），肽-乳糖阻抑物融合体(Cull 等, Proc. Natl. Acad. Sci.
USA 89:1865 (1992)），以及经典的遗传学筛选。与本发明的方法
15 不同，上述每一种方法依赖于蛋白和核酸之间的拓扑联系，因此，可以保留所述蛋白的信息，并能以可读的核酸形式回收。

另外，本发明相对缓慢的翻译方法（Tuerk 和 Gold, 科学 249:
505(1990); Irvine 等, 分子生物学杂志 222: 739(1991); Korman
等, Proc. Natl. Acad. Sci. USA 79: 1844-1848(1992); Mattheakis
20 等, Proc. Natl. Acad. Sci. USA 91:9022-9026(1992); Mattheakis
等, 酶学方法 267: 195(1996); 和 Hanes 和 Pluackthun, Proc. Natl.
Acad. Sci. USA 94: 4937 (1997)）而言具有优点，所述方法是用
于筛选仍然与核糖体及其 mRNA 配合的新生蛋白链的某种特性。与所
述缓慢的翻译技术不同的是，本发明方法不依赖于保持 mRNA: 核糖
25 体：新生链三元复合体的完整性，所述复合体是非常脆弱的，因此，
被局限于那些技术上可行的类型的筛选。

本发明还具有优于由 Drenner 和 Lerner(Proc. Natl. Acad. Sci.
USA 89: 5381-5383 (1992)) 提出的分支合成方法，在其中，制备
DNA-肽融合体，并通过一轮筛选从理论上讲回收遗传信息。与所述分
30 支合成方法不同的是，本发明方法不需要由一种融合体的 DNA 部分再
生肽（该再生过程在分支合成方法中通常是通过一轮化学合成而实现
的）。因此，本发明方法可以使用成群的候选分子进行重复地筛选。

另外，与所述分支合成技术不同的是，该技术通常局限于筛选很短的序列，而本发明方法可用于筛选相当长度的蛋白分子。

另一个优点是，本发明的筛选及直接进化技术可以利用很大而且复杂的候选序列文库。相反，依赖于体内步骤的现有蛋白筛选方法通常局限于具有有限的复杂性的较小的文库。这一优点在筛选功能性蛋白序列，例如，从 10^{13} 个可能的序列中筛选一个仅有 10 个氨基酸长的肽时显得特别重要。在经典的遗传学方法中，lac 阻抑物融合方法，和噬菌体展示方法，最大复杂性通常是在 10^{13} 个成员以下的范围内。

大型文库还具有直接进化利用的优点，可以更深入地调查任何特定起始序列周围的序列空间。

本发明技术与现有方法的不同之处还在于其筛选步骤是与环境无关的。在很多其它筛选方法中，表达蛋白所存在的环境能显著影响所产生的文库的性质。例如，一种表达蛋白不能在一种特定系统中适当表达，或者不能适当展示（例如，展示在噬菌体颗粒的表面）。另外，

一种蛋白的表达可能实际干扰筛选循环中的一个或几个关键步骤，例如，噬菌体的生活力或感染性，或 lac 阻抑物的结合。以上问题会导致功能性分子的丧失或对可以使用的筛选方法的性质的局限。

最后，本发明的优点还在于它能够对可以测定的蛋白的所有成分进行控制。在某些技术中（例如，抗体筛选），对起始库的性质有很小的控制或无控制。在另一种技术中（例如，lac 融合和噬菌体展示），

候选库必须在融合蛋白的环境中表达。相反，RNA-蛋白融合体结构能够控制可用于筛选的候选库的性质。另外，所述候选库的大小可能与 RNA 或 DNA 库一样大（大约 10^{15} 个成员），仅受所进行的体外翻译反应的大小的限制。而且，候选库的制备完全取决于实验设计；可以在提取时或在一种需要的融合蛋白环境中筛选随机裂片，而且大部分

（如果不是全部的话）可能的序列能够在 RNA-蛋白融合体候选库中表达。

通过以下详细说明和权利要求书可以了解本发明的其它特征和优点。

30

详细说明

首先对附图进行简要说明。

附图的简要说明

图 1A-1C 是生产 RNA-蛋白融合体时有关步骤的示意图。图 1A 表示用于制备融合体的 RNA 部分的样品 DNA 结构。图 1B 表示 RNA/嘌呤霉素缀合物的制备。图 1C 表示 RNA-蛋白融合体的制备。

图 2 是本发明的一般化筛选方法的示意图。

图 3 是用于合成含有 3' 嘌呤霉素的最低翻译模板的合成方法的示意图。步骤 (A) 表示将保护基团添加到嘌呤霉素上的活性官能团 (5'-OH 和 NH₂) 上；在经过修饰以后，上述基团受到适当的保护，以便用于亚磷酰胺基寡核苷酸合成。用标准方法通过 2' OH 基团将受到保护的嘌呤霉素连接在可控孔度玻璃 (CPG) 的氨基上，以便通过其 3' OH 结合 DNA (Gait, 寡核苷酸合成, 实用方法, 实用方法系列 (IRL 出版社, 牛津, 1984))。在步骤 (B) 中，用标准 RNA 和 DNA 化学方法 (Millipore, Bedford, MA) 合成含有 43 个核苷酸的最低翻译模板 (称为“43-P”)，用 NH₄OH 和 TBAF 去保护，并进行凝胶纯化。所述模板在其 5' 末端含有 13 个碱基的 RNA，随后是 29 个碱基的 DNA，该 DNA 通过其 5' OH 与 3' 嘌呤霉素连接。所述 RNA 序列包括 (i) 一个互补于 16S rRNA 的 5 个碱基的 SD 共有序列 (Stormo 等, 核酸研究 10: 2971-2996(1982); Shine 和 Dalgarno Proc. Natl. Acad. Sci. USA 71: 1342-1346(1974)；和 Steitz 和 Jakes, Proc. Natl. Acad. Sci. USA 72: 4734-4738(1975))，(ii) 一个有 5 个碱基的间隔区，和 (iii) 一个 AUG 起始密码子。所述 DNA 序列是 dA₂₇dCdCP，其中，“P”是嘌呤霉素。

图 4 是用于制备受保护的 CPG-结合的嘌呤霉素的优选方法的示意图。

图 5 是表示将甲硫氨酸插入本发明模板中的可行方案的示意图。如反应 (A) 所示，所述模板与核糖体结合，以便形成 70S 的起始复合体。Fmet tRNA 结合于所述 P 位点，并且与所述模板碱基配对。位于所述模板的 3' 末端的嘌呤霉素以分子内方式进入 A 位点，并通过肽基转移酶中心与 N-甲酰基甲硫氨酸形成酰胺键，从而对所述 tRNA 脱酰化。用苯酚/氯仿萃取所述反应物，得到甲硫氨酸共价连接的模板。反应 (B) 示出的是不希望有的模板与含有嘌呤霉素的寡核苷酸

之间的分子间反应。与以前一样，所述最低模板能刺激含有 fmet tRNA 的 70S 核糖体与所述 P 位点的结合。随后进入一个反式的第二模板，以形成共价结合的甲硫氨酸。

图 6A-6H 是 ^{35}S 甲硫氨酸 (^{35}S met) 掺入翻译模板的照片。图 6A 表示该反应对镁 (Mg^{2+}) 的依赖性。图 6B 表示该产物的碱基稳定性；在该图中所示出的迁移率的变化相当于失去了 43-P 的 5' RNA 序列(也被称作“Met 模板”)，以生成 DNA-嘌呤霉素部分，被称为 30-P。在碱处理以后所保留的标记与在 ^{35}S 甲硫氨酸和所述模板的 3' 嘌呤霉素之间肽键的形成一致。图 6C 表示在有肽基转移酶抑制剂的条件下对产物形成的抑制。图 6D 表示 ^{35}S 甲硫氨酸掺入一个模板编码序列的依赖性。图 6E 表示 ^{35}S 甲硫氨酸掺入对 DNA 模板长度的依赖性。图 6F 表示用模板 43-P 和 25-P 形成的顺式与反式产物。图 6G 表示用模板 43-P 和 13-P 形成的顺式与反式产物。图 6H 表示用模板 43-P 和 30-P 在网织红细胞裂解物系统中形成的顺式与反式产物。

图 7A-7C 是用于检测肽融合体形成和筛选的结构的示意图。图 7A 表示 LP77 (“连接-产物”，“77”个核苷酸长)(又被称为“短 myc 模板”) (SEQ ID NO: 1)，该序列含有 c-myc 单克隆抗体表位标记 EQKLISEEDL (SEQ ID NO: 2) (Evan 等, 分子细胞生物学 5: 3610-3616 (1985))，其旁侧为 5' 起始密码子和 3' 接头。所述 5' 区含有与 43-P 相同的细菌 SD 序列。对所述编码序列进行优化，以便在细菌系统中表达。具体地讲，43-P 的 5' UTRs 和含有 SD 的 LP77 互补于 16SrRNA 的 5 个碱基 (Steitz 和 Jakes, Proc. Natl. Acad. Sci. USA 72: 4734-4738 (1975))，并以类似于核糖体蛋白序列形式间隔 (Stormo 等, 核酸研究 10: 2971-2996 (1982))。图 7B 表示 LP154 (连接产物，154 个核苷酸长)(又被称为“长的 myc 模板”) (SEQ ID NO: 3)。该序列含有产生用于提取所述 c-myc 抗体的肽的密码。其 5' 末端含有截短形式的 TMV 上游序列(被命名为“TE”)。该 5' UTR 含有一个源于 TMV 5' UTR 的 22 个核苷酸的序列，包括两个 ACAAAUUAAC 直接重复 (Gallie 等, 核酸研究 16: 883 (1988))。图 7C 表示 1 号库 (SEQ ID NO: 4)，用于肽筛选的一种典型序列。所述模板包括源于最初的 myc 肽的最后 7 个氨基酸，用作该模板 PCR 扩增所需要的 3' 恒定区。已知该序列不是所述抗体结合表位的一部分。

图 8 是验证利用模板 43-P, LP77, 和 LP154 以及网织红细胞 (“Retic”) 和小麦胚 (“Wheat”) 翻译系统合成 RNA-蛋白融合体的照片。该图的左半部表示 35 S 甲硫氨酸掺入 3 种模板中的每一种中的情况。该图的右半部表示用 RNaseA 处理所述 3 种模板中的每一种以除去 RNA 编码区后所得到的产物；示出了 35 S 甲硫氨酸标记的 RNA-蛋白融合体。各个 DNA 部分与寡聚 30-P 相同。因此，迁移率的差别与所述编码区的长度呈正比，与各种情况下不同长度的蛋白的存在一致。

图 9 是验证由 LP154 合成的 RNA-蛋白融合体对蛋白酶的敏感性的照片，通过变性聚丙烯酰胺凝胶电泳进行分析。泳道 1 含有 32 P 标记的 30-P。泳道 2-4, 5-7 和 8-10 含有 35 S 标记的翻译模板，所述模板是从网织红细胞裂解反应物中回收的，所述反应物分别未经处理，用 RNaseA 处理或用 RNaseA 和蛋白酶 K 处理。

图 10 是表示使用体外翻译的 33 个氨基酸的 myc-表位蛋白进行的免疫沉淀反应的结果的照片。泳道 1 和 2 分别表示 myc-表位蛋白和 β -珠蛋白模板的翻译产物。泳道 3-5 表示分别使用 c-myc 单克隆抗体和 PBS, DB, 和 PBSTD 洗涤缓冲液进行的所述 myc-表位肽的免疫沉淀的结果。泳道 6-8 表示相同的免疫沉淀反应，不过使用的是 β -珠蛋白翻译产物。

图 11 是验证源于体外翻译反应的 RNA-蛋白融合体的免疫沉淀的照片。标明了用于所述反应的模板的皮摩尔数。泳道 1-4 表示 RNA124 (融合体 LP154 的 RNA 部分)，泳道 5-7 表示 RNA-蛋白融合体 LP154。用 c-myc 单克隆抗体和蛋白 G 琼脂糖凝胶进行免疫沉淀，然后用 RNaseA 和 T4 多核苷酸激酶处理样品，再加样到变性尿素聚丙烯酰胺凝胶上，以显示所述融合体。在泳道 1-4 中，其样品要么不含有任何样品，或者仅含有所述长的 myc 模板 (RNA124) 的 RNA 部分，未发现融合体。在泳道 5-7 中，相应于所述融合体的带清晰可见。示出了 32 P 标记的 30-P 的位置，并且在该图的顶部示出了所投入的模板量。

图 12 是表示由一种体外翻译反应所获得的融合材料的量的曲线。在磷显像板上对图 11 中的泳道 5-7 所示的融合体带和 30-P 带(以类似的方式在 dT₂₅ 上提取，未示出) 的强度进行定量，并作为投放的 LP154 浓度的函数作图。所回收的修饰过的 30-P (左侧 y 轴) 与投入

的模板 (x 轴) 呈线性正比, 而接头-肽融合体 (右侧 y 轴) 是恒定的。通过以上分析, 可以计算出每毫升翻译反应样品产生了大约 10^{12} 融合体。

图 13 是硫代丙基琼脂糖凝胶和 dT_{25} 琼脂糖的示意图, 以及这些 5 底物与本发明的 RNA-蛋白融合体的相互作用能力的示意图。

图 14 是表示顺序提取本发明融合体的结果的照片。泳道 1 含有 ^{32}P 标记过的 30-P。泳道 2 和 3 表示从翻译反应物中提取的、并用 RNaseA 处理过的 LP154。在泳道 2 中, LP154 是用硫代丙基琼脂糖凝胶和 dT_{25} 琼脂糖顺序提取的 LP154。泳道表示仅用 dT_{25} 琼脂糖提取。10 以上结果表明, 所述产物含有游离的巯基, 很可能是 myc 表位编码序列上的倒数第二个半胱氨酸。

图 15A 和 15B 是表示用 β -珠蛋白模板形成融合产物的照片。通过 STS-麦黄酮-PAGE (聚丙烯酰胺凝胶电泳) 进行分析。图 15A 表示无模板 (泳道 1)、syn- β -珠蛋白模板 (泳道 2-4)、或 LP- β -珠蛋白模板 (泳道 5-7) 的 ^{35}S 掺入情况。图 15B (泳道标记如图 15A) 表示通过寡核苷酸亲和层析提取的 ^{35}S -标记的材料。在没有 15 30-P 尾的情况下 (泳道 2-4) 未提取到材料。

图 16A-16C 是表示通过体外筛选富集 myc dsDNA 与库 dsDNA 的示意图和照片。图 16A 是所述筛选方案的示意图。在体外翻译 myc 和库模板的 4 种混合物, 并在 dT_{25} 琼脂糖和 TP 琼脂糖凝胶上分离, 以便从未修饰过的模板中纯化所述模板融合体。然后对 mRNA-肽融合体进行逆转录, 以抑制存在于所述模板中的任何二级或三级结构。在亲和筛选之前 (图 16B) 和之后 (图 16C) 取每一种混合物的等分试样, 在有标记引物的情况下通过 PCR 扩增, 并用仅能裂解 myc DNA 20 25 的限制酶消化。所投入的模板混合物是纯的 myc (泳道 1)、或 1:20、1:200、1:2000 的 myc: 库 (泳道 2-4)。未筛选过的材料因为 myc 模板的优势翻译和逆转录作用而偏离其投入的比例。在筛选步骤中模板的富集作用是通过筛选之前和之后库: myc 比例的变化而计算出来的。

图 17 是表示 myc RNA 模板翻译的照片。使用了下列接头: 泳道 30 1-4, dA₂₇dCdCP; 泳道 5-8, dA₂₇rCrCP; 和 泳道 9-12, dA₂₁C₉C₉C₉dAdCdCP。在每一个泳道中, RNA 模板的浓度为 600nM, 并将

³⁵S -Met 用于标记。反应条件如下：泳道 1、5 和 9, 30°C 1 小时；泳道 2、6 和 10, 30°C 2 小时；泳道 3、7 和 11, 30°C 1 小时，-20°C 16 小时；和泳道 4、8 和 12, 30°C 1 小时，-20°C 16 小时，具有 50mM Mg²⁺。在该图中，“A”表示游离肽，而“B”表示 mRNA-肽融合体。

图 18 是表示用 ³²P 标记的 myc RNA 模板的翻译的照片。所使用的接头是 dA₂₁C₉C₉C₉dAdCdCP。翻译是在 30°C 下进行 90 分钟，而温育是在不添加 Mg²⁺ 的条件下在 -20°C 下进行 2 天。mRNA 模板的浓度为 400nM (泳道 3), 200nM (泳道 4)、100nM (泳道 5) 和 100nM (泳道 6)。泳道 1 表示用 ³⁵S -Met 标记的 mRNA-肽融合体。泳道 2 表示用 ³²P 标记的 mRNA。在泳道 6 上，所述反应是在由 0.5mM cap 类似物的条件下进行的。

图 19 是表示用获自 Ambion (泳道 1)、Novagen (泳道 2)、Amersham (泳道 3) 的裂解物进行的 myc RNA 模板翻译的照片。所使用的接头是 dA₂₇dCdCP。模板的浓度为 600nM。用 ³⁵S -Met 标记。翻译是在 30°C 下进行 1 小时，在有 50mM Mg²⁺ 的条件下在 -20°C 下温育过夜。

本文所披露的是一种利用融合体筛选具有需要的功能的蛋白的方法，其中，所述蛋白与其自身的信使 RNAs 共价连接。这种 RNA-蛋白融合体是通过在其 3' 末端含有肽受体的 mRNA 库的体外或原位翻译而合成的 (图 1B)。在一种优选实施方案中，在通读所述信使的开放读框之后，核糖体在到达设计的间歇位点之后停留，并由所述受体部分占据核糖体的 A 位点，接受来自 P 位点的肽基-tRNA 的新生肽链，以产生 RNA-蛋白融合体 (图 1C)。所述蛋白和 RNA 之间的共价连接 (以所述 mRNA 的 3' 末端和它所编码的蛋白的 C-末端之间的酰胺键的形式连接)使得所述蛋白上的遗传信息可以在通过所述 RNA 的逆转录进行的筛选之后得到回收和扩增 (例如，通过 PCR)。一旦产生所述融合体，就可以根据该 mRNA-蛋白融合体的特性进行筛选或富集，或者用所述 mRNA 模板进行逆转录，同时该模板与所述蛋白结合，以避免所述单链 RNA 对筛选的一切影响。当使用 mRNA-蛋白结构时，可以检验筛选的融合体，以确定哪一部分 (蛋白、mRNA、或两者) 具有所需要的功能。

在一种优选实施方案中，嘌呤霉素 (它类似于酪氨酸酰胺昔) 起着

受体的作用，以便将生长肽连接在其 mRNA 上。嘌呤霉素是一种通过中止肽延伸而起作用的抗生素。作为氨基-tRNA 的模拟物，它通过结合所述 A 位点、接纳生长肽链、和脱落核糖体 ($K_d=10^{-4} M$) 而起着蛋白合成的通用抑制剂的作用 (Traut 和 Monro, 分子生物学杂志 5 10: 63 (1964); Smith 等, 分子生物学杂志 13: 617 (1965))。嘌呤霉素的一个最诱人的特征是，它能与生长肽链生成稳定的酰胺键，从而可以进行更稳定的融合，这种融合比能形成不稳定的酯键的潜在受体稳定。具体地讲，所述肽基-嘌呤霉素分子含有一种位于肽和嘌呤霉素的 O-甲基酪氨酸部分之间的稳定的酰胺键。所述 O-甲基酪氨酸又通过一个稳定的酰胺键与嘌呤霉素的修饰过的腺昔部分的 10 3'-氨基结合。

受体的其它可行筛选包括位于所述 mRNA 的 3' 末端的 tRNA-样结构，以及以类似于嘌呤霉素的方式起作用的其它化合物。这些化合物包括，但不限于具有一个与腺嘌呤或腺嘌呤-样化合物连接的氨基酸的所有化合物，如氨基酸核苷酸、苯丙氨酰-腺昔 (A-Phe)，酪氨酰腺昔 (A-Tyr) 和丙氨酰腺昔 (A-Ala)，以及酰胺-连接的结构，如苯丙氨酰 3' 脱氧 3' 氨基腺昔，丙氨酰 3' 脱氧 3' 氨基腺昔，和酪氨酰 3' 脱氧 3' 氨基腺昔；在上述任一种化合物中，可以使用任何天然存在的 L-氨基酸或其类似物。另外，还可将组合的 tRNA-样-3' 结构-20 嘌呤霉素缀合物用于本发明中。

图 2 中示出了本发明的一种优选筛选方案。该筛选方案所涉及的步骤通常是按如下方式进行的。

步骤 1：制备 DNA 模板。作为制备本发明的 RNA-蛋白融合体的一个步骤，合成该融合体的 RNA 部分。这一目的可以通过直接化学 RNA 合成而实现，更常见的是，通过转录合适的双链 DNA 模板而实现。25

所述 DNA 模板可以通过任何标准技术（包括重组 DNA 技术、化学合成的任一种技术或两种技术）制备。原则上讲，能产生一种或几种含有一种已知的、随机的、随机化的或诱变的序列的模板的方法均可用于这一目的。在一种具体方法中，在转录前合成并扩增（例如，通过 PCR）一种寡核苷酸（例如，含有随机碱基）。化学合成还可用于产生随机框，然后将该框插入已知蛋白编码序列的中央（例如，参见 Ausubel 等所著的“现代分子生物学方法”的第 8.2 章，John 30

Wiley&Sons 和 Greene 出版公司, 1994)。后一种方法能在所述蛋白的感兴趣的特定位点附近产生高密度的突变。

DNA 模板序列的完全随机化的替代方案是部分随机化, 而且, 用这种方法合成的一种库通常被称作“掺杂库”。例如, 在 RNA 序列上 5 进行的该技术的一个例子由 Ekland 等披露 (核酸研究 23: 3231 (1995))。部分随机化可以通过偏离所述合成反应而以化学方法完成, 使得每一种碱基添加反应混合物含有过量的一种碱基和少量的每一种其它的碱基; 通过小心控制所述碱基的浓度, 可以用该方法获得理想的突变频率。还可以利用易错 PCR 技术制备部分随机化的库, 例 10 如, 由 Beaudry 和 Joyce(科学 257: 635(1992)) 和 Bartel 和 Szostek (科学 261: 1411 (1993)) 所披露的。

还有用于制备 DNA 结构的多种方法, 该方法由一个已知序列开始, 产生一种突变的 DNA 库。所述方法的例子披露于 Ausubel 等的著述 (同上, 第 8 章) 和 Sambrook 等的著述中 (分子克隆: 实验室手册, 第 15 章, 冷泉港出版社, 纽约, 第 2 版 (1989))。还可以通过披露于 Stemmer 的著述中的 (自然 370: 389 (1994)) 的“改组” 15 技术制备随机序列。

为了优化本发明的筛选方案, 还可以改变一个模板的 5' 和 3' 末端的序列和结构。优选地, 这一目的是通过两次独立的筛选实现的, 20 每一次筛选包括将随机区插入所述模板的接近合适的末端的部位, 然后进行筛选。所述筛选可以起到以下作用: (i) 将制备的融合体的量最大化 (并因此使得文库的复杂性最大化) 或 (ii) 提供优化的翻译序列。另外, 所述方法具有通用性, 可以与诱变 PCR 结合, 以便优化翻译模板的编码和非编码区。

步骤 2: 制备 RNA。如上文所述, 可以用标准寡核苷酸合成技术, 25 通过化学方法合成 RNA-蛋白融合体的 RNA 部分。另外, 特别是在使用较长的 RNA 序列时, 所述 RNA 部分是通过 DNA 模板的体外转录而制备的。在一种优选方案中, 将 T7 聚合酶用于酶促制备所述 RNA 链。适用于上述目的的其它 RNA 聚合酶包括, 但不限于 SP6、T3 和大肠杆菌 RNA 聚合酶 (例如, 披露于 Ausubel 等的著述 (同上, 第 3 章))。 30 另外, 所合成的 RNA 可以是全部或部分修饰的 RNA。在一个具体例子中, 可以使用修饰的核糖核苷酸和标准技术产生硫代磷酸 RNA (例如,

通过 T7 转录)。上述修饰的 RNA 具有对核酸酶稳定的优点。

步骤 3: 将嘌呤霉素连接到所述模板上。接着, 将嘌呤霉素(或任何其它合适的肽受体)共价连接在所述模板序列上。这一步骤可以是通过利用 T4 RNA 连接酶将所述嘌呤霉素直接连接在所述 RNA 序列上, 或者优选通过利用 T4 DNA 连接酶或任何其它能够将两个核苷酸序列连接在一起的酶将所述嘌呤霉素与 DNA “裂片”连接的方式而实现(参见图 1B)(例如, 同样参见 Ausubel 等, 第 3 章, 第 14 和 15 节)。还可将 tRNA 合成酶用于将嘌呤霉素-样化合物连接在 RNA 上。例如, 苯丙氨酸合成酶能将苯丙氨酸连接在含有 3' 氨基的苯丙氨酰-tRNA 分子上, 产生具有嘌呤霉素-样 3' 末端的 RNA 分子(Fraser 和 Rich, Proc. Natl. Acad. Sci. USA 70: 2671 (1973))。可以使用的其它肽受体包括, 但不限于具有连接于腺嘌呤或腺嘌呤-样化合物上的氨基酸的所有化合物, 如氨基酸核苷酸、苯丙氨酰-腺苷(A-Phe)、酪氨酰-腺苷(A-Tyr)、和丙氨酰-腺苷(A-Ala)、以及腺苷-连接的结构, 如苯丙氨酰 3' 脱氧 3' 氨基腺苷、丙氨酰 3' 脱氧 3' 氨基腺苷、和酪氨酰 3' 脱氧 3' 氨基腺苷; 在上述任一种化合物中, 可以使用任何天然存在的 L-氨基酸或其类似物。例如, 在 Krayevsky 和 Kukhanova 的著述中(核酸研究和分子生物学进展 23: 1 (1979))披露了多种肽受体。

步骤 4: RNA-蛋白融合体的制备和回收。为了制备 RNA-蛋白融合体, 任何体外或原位翻译系统均可使用。如下文所述, 优选真核系统, 两种特别优选的系统包括小麦胚和网织红细胞裂解物系统。不过, 原则上讲, 任何能够生成 RNA-蛋白融合体, 而且不会明显降解该融合体的 RNA 部分的翻译系统均可用于本发明中。另外, 为了减少上述任何系统中的 RNA 降解, 可将降解-抑制反义寡核苷酸加入所述翻译反应混合物中; 所述寡核苷酸能与所述分子中引起降解作用的 RNA 部分的序列特异杂交并覆盖该序列(例如, 参见 Hanes 和 Pluckthun, Proc. Natl. Acad. Sci. USA 94: 4937 (1997))。

如上文所述, 可将多种真核翻译系统中的任一种用于本发明中。所述系统包括, 但不限于源于酵母、腹水、肿瘤细胞(Leibowitz 等, 酶学方法 194: 536 (1991))、和非洲爪蟾卵母细胞的裂解物。源于细菌系统的体外翻译系统包括, 但不限于披露于以下文献中的系

统: Zubay (遗传学年度综述 7: 267 (1973)) ; Chen 和 Zubay (酶学方法 101: 44(1983)) ; 和 Ellman(酶学方法 202: 301(1991))。

另外, 翻译反应可以原位进行。在一种特定实施例中, 翻译可以通过用标准技术将 mRNA 注入非洲爪蟾卵中而实现。

5 一旦产生了 RNA-蛋白融合体, 就可以通过任何标准的蛋白或 RNA 纯化技术从所述翻译反应混合物中回收该融合体。通常, 使用蛋白纯化技术。例如, 如下文所示, 可以通过使用合适的层析试剂, 如 dT₂₅ 琼脂糖或硫代丙基琼脂糖凝胶实现融合体的纯化。不过, 纯化还可以涉及基于所述融合体的 RNA 部分的纯化; 例如, 所述纯化技术披露于
10 Ausubel 等的著述中 (同上, 第 4 章)。

15 步骤 5: 筛选需要的 RNA-蛋白融合体。需要的 RNA-蛋白融合体的筛选可以通过任何可用于从一群候选融合体中选择性地分离或提取需要的融合体的方法实现。提取技术的例子包括, 但不限于, 与诸如结合配偶体的选择性结合, 所述配偶体直接或间接固定在柱、珠、膜、或其它固体支持物上。并用一种对所述融合体的蛋白部分特异的
20 抗体进行免疫沉淀。所述技术的第一种利用了一种固定筛选基序, 该基序可以由任何类型的分子组成, 其可以发生结合。在图 2 中示出了一系列可以筛选的基序分子。筛选还可以基于连接在一种亲和标记物 (例如, 底物-生物素) 上的底物分子的利用, 所述标记物能与候选分子起反应, 或者基于与融合分子的任何其它类型的相互作用。另外,
25 还可以按照类似于由 Bartel 和 Szostak 所披露的用于提取 RNA 酶 (同上) 的方法根据其催化活性筛选蛋白; 按照特定的技术, 根据其将靶分子连接在它本身上面的能力筛选需要的分子, 然后根据靶分子的存在提取功能性分子。本发明可以利用相同的方法或任何其它功能性筛选的筛选方法, 筛选新型或改进的催化蛋白。

另外, 如本文所述, 可以通过在一种候选分子库中富集融合体来筛选所需要的 RNA-蛋白融合体 (或其 DNA 拷贝)。为了进行这样一种任选的富集, 将一群候选 RNA-蛋白融合体与一种结合配偶体 (例如, 上述结合配偶体之一) 接触, 所述结合配偶体对所述融合体的 RNA 部分或蛋白部分专一, 接触条件为能基本上将该结合配偶体-融合体从样品中未结合的成员里分离出来。这一步骤可以重复, 而且该技术优选包括至少两个连续的富集步骤, 一个步骤是用一种对 RNA 部

分专一的结合配偶体筛选融合体，另一个步骤是用对蛋白部分专一的结合配偶体筛选融合体。另外，如果重复针对融合体的相同部分（例如，蛋白部分）的富集步骤，优选使用不同的结合配偶体。在本文所披露的一种具体例子中，首先通过使用对融合体的 RNA 部分专一的结合配偶体从一群分子中富集所需要的融合体，然后在两个连续的步骤中使用两个不同的结合配偶体，这两种结合配偶体都对所述融合体的蛋白部分具有专一性。另外，还可以通过任何标准分离技术从样品成分中分离，这些技术包括，但不限于柱亲和层析、离心、或免疫沉淀。

另外，还可以通过多种方法实现从一种富集（或筛选）复合物中洗脱 RNA-蛋白融合体。例如，如本文所述，可以利用一个变性或非特异性化学洗脱步骤分离需要的 RNA-蛋白融合体。这样一个步骤有利于将复合的成分彼此分离或者与结合的固体支持物分离，这种分离是通过破坏所述成分之间和/或所述成分与固体支持物之间的非共价键而以相对非特异的方式实现的。如本文所述，一种典型的变性或非特异化学洗脱试剂是 4% HOAc/H₂O。其它典型的变性或非特异化学洗脱试剂包括鸟嘌呤、尿素、高盐、洗涤剂、或任何一般能除去非共价加成物的其它方法。另外，可以使用一种特殊的化学洗脱方法，其中，利用一种化合物引起融合分子的特异释放。在一种具体实例中，如果所需要的融合蛋白的连接臂含有一个或几个二硫键，可以通过添加诸如 DTT 导致二硫键减弱并释出结合的靶分子，而将结合的融合 aptamers 洗脱。

另外，洗脱还可以通过特异破坏亲和性复合体而实现；这种技术能通过添加过量的该复合物的一种成分而选择性地分离该复合物的成分。例如，在 ATP-结合筛选中，洗脱是通过将过量的 ATP 加入温育混合物中而实现的。最后，还可以进行一个酶促洗脱步骤。在该方法中，结合分子本身或外源添加的蛋白酶（或其它合适的水解酶）能裂解并释放出所述靶或酶。在一种具体例子中，蛋白酶靶位点可以包含在任一种复合成分中，并通过添加蛋白酶洗脱结合分子。另外，在催化筛选中，可将洗脱作为一个筛选步骤，以便分离能够将其本身从固体支持物上释出（例如，裂解）的分子。

步骤 6: 利用逆转录酶制备所述 RNA 序列的 DNA 拷贝。如果需要，通过使用任何标准技术（例如，使用 Superscript 逆转录酶）对 RNA

序列进行逆转录可以方便地获得特定 RNA 融合序列的 DNA 拷贝。这一步骤可以在筛选或富集步骤之前（例如，如图 16 所示）或之后进行。另外，所述逆转录过程还可以在将所述融合体从体外或原位翻译混合物中提取出来之前进行。

5 然后，作为部分或完整长度的双链序列扩增所述 DNA 模板。优选地，在该步骤中，产生完整长度的 DNA 模板，利用合适的寡核苷酸和 PCR 扩增。

下面将采用具体实例的形式详细说明上述步骤以及用于完成这些步骤的试剂和技术。提供这些实例是为了说明本发明，不应视为对
10 本发明的限定。

制备 RNA-蛋白融合体的模板

如图 1A 和 2 所示，本发明的筛选方案优选利用包括若干设计元件的双链 DNA 模板。所述因子的第一个是启动子，它与所需要的 RNA
15 聚合酶一起使用，以便合成 mRNA。如图 1A 所示，而且如这里所披露的，尽管能够指导由线性双链 DNA 进行的合成的任何启动子均可使用，但优选 T7 启动子。

图 1A 所示模板的第二个因子被称为 5' 非翻译区（或 5' UTR），
20 它相当于翻译起始位点的 RNA 上游，图 1A 中示出了一种优选的 5' UTR
(称为“TE”)，它是烟草花叶病毒 5' 非翻译区的缺失突变型，具体地讲，相当于 TMV 翻译起点的 5' 碱基；该 UTR 的序列如下：
rGrGrGrArCrArArUrUrArCrUrArUrUrArCrArArUrUrArCrA (其头三个 G 核苷酸被插入以增强转录) (SEQ ID NO: 5)。任何其它合适的
25 5' UTR 均可使用 (例如，参见 Kozak, 微生物学综述 47: 1 (1983))。

图 1A 中所示的第三个因子是翻译起始位点。一般，它是一个 AUG
密码子。不过，存在着将除 AUG 之外的密码子用于天然存在的编码序
30 列中的例子，这些密码子也可被用于本发明的筛选方案中。

图 1A 的第四个因子是编码所述蛋白序列的蛋白的开放读框 (称作 ORF)。该开放读框可编码任何天然存在的、随机的、随机化的、
诱变的、或完全合成的蛋白序列。

图 1A 所示的第五个因子是 3' 恒定区。该序列有利于所述库序列的 PCR 扩增，以及含有嘌呤霉素的寡核苷酸与 mRNA 的连接。如果需

要，该区还可以包括一个间歇位点，该位点是一个能引起核糖体停留，并因此使得受体部分(例如，嘌呤霉素)具有接受来自肽基-tRNA的新生肽链的额外时间；该间歇位点将在下面作更详细的讨论。

为了完成本发明的方法，首先利用高度简化的含有1-2个密码子的mRNA模板制备RNA-蛋白融合体。实施上述方法是出于两种考虑。首先，这种大小的模板容易通过化学合成制备。其次，小的开放读框使得所述反应的重要特征，包括连接的效率、末端异质性、模板依赖性、和翻译精确性能够方便地加以测定。

结构设计。将一种基本结构用于制备试验RNA-蛋白融合体。所述分子由含有用于启动翻译的SD序列的mRNA组成，该分子包括源于核糖体蛋白L1的SD序列的三个碱基的缺失，而且它互补于16S rRNA的5个碱基(即rGrGrArGrGrArCrGrArA)(SEQ ID NO: 6)(Stormo等，核酸研究10: 2971-2996(1982); Shine和Dalgarno, Proc. Natl. Acad. Sci. USA 71: 1342-1346(1974)；和Steitz和Jakes, Proc. Natl. Acad. Sci. USA 72: 4734-4738(1975))，(ii)一个AUG起始密码子，(iii)一个用作间歇位点的DNA接头(即5'-(dA)₂₇)，(iv)dCdC-3'，和(v)一个3'嘌呤霉素(P)。选择poly dA序列是因为已知它在A位点作为tRNA的模板的能力较差(Morgan等，分子生物学杂志26: 477-497(1967)；Ricker和Kaji，核酸研究19: 6573-6578(1991))，并将其设计成良好的间歇位点。选择所述寡聚dA接头的长度，使其跨过解码位点和所述核糖体的肽基转移中心之间的大约60-70埃的距离。所述dCdCP模拟tRNA的CCA末端，设计它是为了有利于嘌呤霉素与核糖体的A位点的结合。

最低模板43-P的化学合成。为了合成结构43-P(如图3所示)，首先将嘌呤霉素连接在一种固体支持物上，使其适用于标准亚磷酸胺寡核苷酸合成化学。在图3中示意性地示出了这种寡聚物的合成方案，并将在下文作更详细的说明。为了将嘌呤霉素连接在可控孔度玻璃(CPG)固体支持物上，用三氟乙酰基保护氨基，如在DNA合成仪380型的Applied Biosystem用户简报#49(9188)中所披露的。然后，用标准的DMT-C1方法对5'OH进行保护(Gait, 寡核苷酸合成，实用方法，实用方法系列(IRL出版社，牛津，1984))，并通过2'OH连接在氨己基CPG上，其连接方式与将3'OH用于连接脱氧核苷的方式

式相同（参见图 3 和 Gait, 同上, p. 47）。5' DMT-CPG-连接的保护嘌呤霉素适合于用亚磷酰胺单体进行链的延伸。所述寡聚物的合成沿 3' → 5' 方向以以下顺序进行：(i) 3' 嘌呤霉素, (ii) pdCpdC, (iii) 大约 27 个单位的 dA 作为接头, (iv) AUG, 和 (v) SD 序列。43-P
5 结构的序列如下文所示。

CPG 嘌呤霉素的合成。保护的 CPG 嘌呤霉素的合成遵循上文所述的用于脱氧核苷合成的一般途径 (Gait, 寡核苷酸合成, 实用方法, 实用方法系列 (IRL 出版社, 牛津, 1984))。主要差别包括筛选合适的 N 封闭基团, 结合在固体支持物的 2' OH, 以及与该固体支持物的连接反应。在后一种情况下, 反应是在极低浓度的活化核苷酸中进行的, 因为该材料显著优于所述固体支持物。就所述稀释反应条件而言, 所得到的产量 (大约 20 μmol/g 支持物) 是十分理想的。
10

合成 N-三氟乙酰嘌呤霉素。首先将 267mg (0.490mmol) 嘌呤霉素*HCl 溶于水中, 加入 pH11 的碳酸缓冲液, 并萃取 (3X) 到氯仿中, 以便将其转化成游离碱的形式。将有机相蒸发干燥, 并称重 (242mg, 0.513mmol)。然后将所述游离碱溶解在 11ml 无水吡啶和 11 ml 无水乙腈中, 并搅拌添加 139 μl (2.0 mmol) 三乙胺 (TEA) 和 139 μl (1.0 mmol) 三氟乙酸酐 (TFAA), 然后将 TFAA 加入 20 μl 等分试样的浑浊溶液中, 直到不再有原始材料, 通过薄层层析 (tlc) (93: 7, 氯仿/MeOH) (共 280 μl) 测定。让所述反应进行 1 小时。此时, 通过薄层层析发现了两条带, 这两条带的迁移率均高于原始材料的。用 NH₄OH 和水调节所述反应, 将其产物还原成一条带。硅层析 (93: 7 氯仿/MeOH) 得到 293 mg (0.515 mmol) 的产物 N-TFA-Pur。该反应的产物示意性地示于图 4 中。
15
20

合成 N-三氟乙酰 5'-DMT 嘌呤霉素。等分来自上述反应的产物, 并与无水吡啶共蒸发 2 次, 以除去水。准备多个试管, 以便实验多种反应条件。在一种小规模反应中, 将 27.4mg (48.2mmol) N-TFA-Pur 溶解在 480 μl 含有 0.05 当量 DMAP 和 1.4 当量 TEA 的吡啶中。将 20.6mg 的三苯甲基氯 (60 μmol) 加入该混合物中, 并搅拌让反应进行到结束。通过向该溶液中添加等体积的水 (大约 500 μl) 中止该反应。因为该反应似乎是成功的, 进行大规模的反应。具体地讲, 将 262mg (0.467mmol) N-TFA-Pur 溶解在 2.4ml 吡啶中, 然后添加 1.4
25
30

当量的 TEA, 0.05 当量的 DMAP, 和 1.2 当量的三苯甲基氯。大约 2 小时之后, 再添加 50mg(0.3 当量)二甲氧基三苯甲基*Cl(DMT*Cl), 让反应再继续进行 20 分钟, 通过添加 3ml 的水中止反应, CH₃CN 共蒸发 3 次。在 2mm 直径的 100ml 硅(干燥)柱上通过 95: 5 氯仿/MeOH 5 纯化所述反应物。由于纯化不彻底, 在第二个相同的柱上用 97.5: 2.5 的氯仿/MeOH 进一步纯化。总产量为 325mg 或 0.373mmol(或产率为 72%)。该反应的产物示意性地示于图 4 中。

合成 N-三氟乙酰, 5'-DMT, 2'琥珀酰嘌呤霉素。在小规模反应中, 将上面合成的 32mg (37 mmol) 的产物与 1.2 当量的 DMAP 混合溶解在 350 μl 的吡啶中。将溶解在 44 μl 的无水 CH₃CN 中的 1.2 当量的琥珀酸酐加入该溶液中, 并搅拌过夜。薄层层析发现仅存很少的原始材料。在大规模的反应中, 将 292mg (336 mmol) 的上述产物与 1.2 当量的 DMAP 混合于 3ml 的吡啶中。将 403 μl 的 1M 溶于无水 CH₃CN 中的琥珀酸酐加入该混合物中, 并将该混合物搅拌过夜。薄层层析再次发现了仅存很少的原始材料。合并以上两种反应物, 再添加 0.2 当量的 DMAP 和琥珀酸盐。该产物与甲苯共蒸发 1 次, 并在高真空中干燥成黄色泡末。加入 CH₂Cl₂ (20ml), 用 15ml 的 10% 的冰镇柠檬酸萃取 2 次, 然后再用纯水萃取 2 次。干燥该产物, 重新溶解在 2ml CH₂Cl₂ 中, 并通过搅拌添加 50ml 的己烷沉淀。然后对该产物进行涡旋搅拌, 10 并在临床离心机上以 600rpm 的速度离心 10 分钟。排出大部分洗脱液, 并干燥其余的产物, 首先在低真空下干燥, 然后在干燥器中以高真空干燥。该反应的产量大约为 260 μmol, 分级产率为大约 70%。

合成 N-三氟乙酰 5'-DMT, 2'琥珀酰, CPG 嘌呤霉素。随后将来自上一步骤的产物溶解在 1ml 的二恶烷中, 随后用 0.2 ml 二恶烷/0.2 ml 吡啶溶解。将 40mg 对-硝基苯酚和 140mg 的二烷己基碳二亚胺 (DCC) 加入该溶液中, 并让该反应进行 2 小时。通过离心除去由该反应产生的不可溶的环己基尿素, 将该产物溶液加入悬浮在 22ml 无水 DMF 中的 5g 氨己基可控孔度玻璃 (CPG) 中, 并搅拌过夜。然后用 DMF、甲醇、和乙醚洗涤该树脂, 并干燥。分析证实, 所得到的树脂 20 每克含有 22.6 μmol 三苯甲基游基, 处于这种类型的支持物的可接受的范围内。然后, 通过与 15ml 吡啶、1ml 乙酸酐、和 60mg 的 DMAP 一起温育 30 分钟封闭该支持物。所得到的柱材料能产生一种负的(无

色)水合茚三酮实验,与封闭之前获得的结果相反,在封闭之前所获得的结果中,所述材料能产生一种深兰色反应。该反应的产物示意性地示于图4中。

合成 mRNA-嘌呤霉素缀合物。如上文所述,可将嘌呤霉素固定的寡聚物用于由两种方法中的任一种制备能起着翻译模板作用的 mRNA-嘌呤霉素缀合物。对于极短的开放读框来说,通常通过化学方法用 RNA 或 DNA 单体延伸嘌呤霉素寡聚物,以产生完全合成的模板。当需要较长的开放读框时,通常利用 DNA 裂片和 T4 DNA 连接酶将 RNA 或 DNA 寡聚物连接在 mRNA 的 3'末端,如 Moore 和 Sharp 所披露的(科学 256: 992 (1992))。

RNA-蛋白融合体的体外翻译和测定

用如下的细菌和真核体外翻译系统体外翻译上面制备的模板。

最低模板的体外翻译。将 43-P 和相关的 RNA-嘌呤霉素缀合物加入几个不同的体外翻译系统中,所述系统包括: (i) 源于大肠杆菌 MRE600 的 S30 系统(Zubay, 遗传学年度综述 7: 267(1973); Collins, 基因 6: 29 (1979); Chen 和 Zubay, 酶学方法 101: 44 (1983); Pratt, 转录和翻译: 实践方法, B. D. Hammes, S. J. Higgins 著 (IRL 出版社, 牛津, 1984) pp. 179-209; 和 Ellman 等, 酶学方法 202: 20 301 (1991)), 通过 Ellman 等所披露的方法制备(酶学方法 202: 301 (1991)); (ii) 源于相同菌株的核糖体部分,通过 Kudlicki 等所述的方法制备(分析化学 206: 389 (1992)); 和 (iii) 源于大肠杆菌 BL21 的 S30 系统,通过 Lesley 等所述方法制备(生物化学杂志 266: 2632 (1991))。在每一种情况下,所使用的预混合物是 Lesley 等的混合物(生物化学杂志 266: 2632 (1991)),温育持续 30 分钟。

测定所述融合体的性质。首先用源于大肠杆菌 S30 的翻译提取物检验 43-P 模板。图 5(反应 “A”)证实了希望的分子内(顺式)反应,其中,43-P 结合核糖体,并起着 fMet 的模板和受体的作用。首先测定³⁵S-甲硫氨酸的掺入及其在所述模板中的位置,结果示于图 6A 和 6B 中。用苯酚/氯仿萃取所述体外翻译反应混合物,并通过 SDS-PAGE 分析该产物,然后,一个³⁵S 标记的带以与 43-P 模板相同的迁

移率出现。所合成的该材料的量取决于 Mg^{2+} 的浓度 (图 6A)。最佳 Mg^{2+} 浓度为 9–18 mM, 这一浓度类似于在该系统中进行翻译的最佳浓度 (Zubay, 遗传学年度综述 7: 267 (1973); Collins, 基因 6: 29 (1979); Chen 和 Zubay, 酶学方法 101: 44 (1983); Pratt, 转录和翻译: 实践方法, B. D. Hammes, S. J. Higgins 著 (IRL 出版社, 牛津, 1984) pp. 179–209; Ellman 等, 酶学方法 202: 301 (1991)); Kudlicki 等 (分析化学 206: 389 (1992)); 和 Lesley 等 (生物化学杂志 266: 2632 (1991))。另外, 结合的标记物在用 NH_4OH 处理时是稳定的 (图 6B), 表明该标记物位于该分子的 3' 半边 (碱稳定性 DNA 部分), 并通过碱稳定性键连接, 这正是嘌呤霉素和 fMet 之间的酰胺键所出现的情形。

核糖体和模板依赖性。为了证实上面所观察到的发生在核糖体上的反应, 检验核糖体的肽基转移酶功能的特异抑制剂的作用 (图 6C), 并检验改变编码甲硫氨酸的序列的作用 (图 6D)。图 6C 清楚地表明所述反应受到肽基转移酶抑制剂维及尼霉素、谷氏菌素、和氯霉素的有效抑制 (Monro 和 Vazquez, 分子生物学杂志 28: 161–165 (1967); 和 (Vazquez 和 Monro, 生物化学和生理学学报 142: 155–173 (1967)))。图 6D 表明, 所述模板中发生的一个碱基由 A 到 C 的变化, 可破坏 ^{35}S 甲硫氨酸在 9 mM Mg^{2+} 下的掺入, 并在 18 mM 下大大减弱其掺入 (这一结果与高浓度的 Mg^{2+} 容许信息的错读的事实一致)。上述实验表明, 所述反应以模板依赖形式发生在核糖体上。

接头长度。还试验了所述反应用于接头长度的依赖性 (图 6E)。将原始模板设计成使得所述接头跨越从解码位点 (由该模板的 AUG 占据) 到受体位点 (由嘌呤霉素部分占据) 的距离, 该距离的长度大约与 tRNA 上的反密码子环和受体臂之间的距离相同, 或者大约为 60–70 埃。试验过的第一个接头为 30 个核苷酸长, 基于每个碱基最少 3.4 埃 (≥ 102 埃)。在 30–21 个核苷酸的长度范围内 ($n=27–18$; 长度 $\geq 102–71$ 埃), 在所述反应的效率上很少出现变化。因此, 接头长度可以改变。尽管 21–30 个核苷酸的接头代表优选的长度, 短于 80 个核苷酸, 优选短于 45 个核苷酸的接头也可用于本发明中。

分子内与分子间反应。最后, 我们检验了所述反应是否如希望的那样以分子内方式发生 (图 5, 反应 “A”) 或以分子间方式发生 (图

5, 反应“B”）。这一试验是这样进行的：将具有 3' 嘴吟霉素，但没有核糖体结合序列的寡核苷酸（模板 25-P, 13-P, 和 30-P）加入含有 43-P 模板的翻译反应物中（图 6F、6G、和 6H）。如果该反应是通过分子间机制发生，则较短的寡聚物也会被标记。如图 6F-H 所示，
 5 掺入 3 种较短寡聚物中的 ^{35}S 甲硫氨酸很少，这表明该反应主要是以分子内方式进行。下面示出了序列 25-P (SEQ ID NO: 10)、13-P (SEQ ID NO: 9)、30-P (SEQ ID NO: 8)。

网织红细胞裂解物。图 6H 表明，除了上面所使用的大肠杆菌裂解物之外，还可在将 ^{35}S -甲硫氨酸掺入 43-P 之前，利用兔网织红细胞裂解物（参见下文）进行体外翻译。该反应主要是以希望的分子内机制进行。
 10

含有 c-myc 肽标记的融合体的合成和试验

还制备了代表性的融合体，该融合体在其蛋白部分含有 c-myc
 15 单克隆抗体 9E10 的表位标记 (Evan 等, 分子细胞生物学 5: 3610
 (1985))。

模板设计。设计了三种原始表位标记模板（即 LP77、LP154 和 1 号库），如图 7A-C 所示。前两个模板含有 c-myc 表位标记序列 EQKLISEEDL (SEQ ID NO: 2)，而第三个模板被设计成用于合成随机筛选库。LP77 编码 12 个氨基酸的序列，将其密码子优化用于细菌翻译。LP154 及其衍生物含有 33 个氨基酸的 mRNA 序列，其中，将其密码子优化用于真核翻译。其所编码的氨基酸序列 MAEEQKLISEEDLLRKRREQQLKHKLEQLRNSCA (SEQ ID NO: 7)，相当于用于提取 9E10 抗体的原始肽。1 号库含有 NNG/C 的 27 个密码子（用于产生随机肽），其后是一个相当于 myc 肽的最后 7 个氨基酸的序列（这 7 个氨基酸不是 myc 表位序列的一部分）。上述序列如下文所示。
 20
 25

网织红细胞与小麦胚体外翻译系统。在兔网织红细胞和小麦胚提取物 (Promega, Boehringer Mannheim) 翻译系统中试验 43-P、LP77、和 LP154 模板（图 8）。翻译是在 30°C 的条件下进行 60 分钟，模板是用 dT₂₅ 琼脂糖在 4°C 下提取的。用 15mM NaOH、1mM EDTA 将模板从琼脂糖上洗脱，用乙酸钠/乙酸缓冲液中和，立即用乙醇沉淀 (2.5-3 体积)，洗涤（用 100% 的乙醇），并在真空离心蒸发浓缩器上干燥。
 30

图 8 表示在小麦胚和网织红细胞系统中 ^{35}S 甲硫氨酸掺入到所有三种模板中。在源于网织红细胞系统的融合反应物中观察到较少的模板降解作用，因此，该系统被优先用于制备 RNA-蛋白融合体。另外，一般真核系统优于细菌系统。因为真核细胞倾向于含有较低水平的核酸酶，在这种细胞中 mRNA 的寿命通常比在细菌细胞中的寿命长 10-100 倍。在使用一种特定大肠杆菌翻译系统的试验中，使用编码 c-myc 表位的模板未观察到融合体的产生；通过在不同的位点标记该模板证实了这可能是由于模板的 RNA 和 DNA 部分的降解所致。

为了检验所述融合体的肽部分，用 RNase 处理样品，以除去编码序列。在该处理之后，43-P 产物大体上以与 ^{32}P 标记的 30-P 寡聚物相同的迁移率运动，这一结果与很小的肽（也许仅有甲硫氨酸）添加到 30-P 上一致。对于 LP77 来说，除去编码序列所产生的产物具有低于 30-P 寡聚物的迁移率，这一结果与发现有 12 个氨基酸的肽添加到嘌呤霉素上一致。最后，对于 LP154 来说，除去编码序列产生一种具有更低迁移率的产物，这一结果与有 33 个氨基酸序列连接到 30-P 寡聚物上一致。由于加样误差，在 RNase-处理过的 LP154 网织红细胞泳道上未发现寡聚物。在图 9 中，证实该产物的迁移率与在小麦胚提取物中所产生的产物的迁移率相同。总之，以上结果表明能抗 RNase 的产物被添加到 30-P 寡聚物的末端，该产物的大小与所述编码序列的长度成正比，而且，该产物的大小相当一致。另外，尽管两种系统都能产生类似形式的产物，但由于网织红细胞系统具有较高的模板稳定性，该系统是优选的。

对 RNase A 和蛋白酶 K 的敏感性。 在图 9 中用 LP154 融合体试验了对 RNase A 和蛋白酶 K 的敏感性。如泳道 2-4 所示，在 LP154 模板上证实了 ^{35}S 甲硫氨酸的掺入。当用 RNase A 处理该产物时，该融合体的迁移率降低，但明显高于 ^{32}P 标记的 30-P 寡核苷酸的迁移率，这一结果与 33 个氨基酸的肽添加到其 3' 末端一致。当同样用蛋白酶 K 处理该材料时， ^{35}S 的信号完全消失，这一结果同样与发现该标记物存在于位于 30-P 片段的 3' 末端的一个肽上的事实一致。在使用 43-P 和 LP77 融合体的等同实验中获得了类似结果。

为了证实 ^{35}S Met 对所述模板的标记是由翻译造成的，更具体地讲，是由核糖体的肽基转移酶活性造成的，试验了各种抑制剂对标记

反应的影响。真核肽基转移酶的特定抑制剂茴香霉素、谷氏菌素、和稀疏霉素 (Vazquez, 蛋白合成的抑制剂 (Springer-Verlag, 纽约), pp. 312 (1979)), 以及转运抑制剂环己酰亚胺和吐根碱 (Springer-Verlag, 纽约), pp. 312 (1979)) 均能降低利用长的
5 myc 模板和网织红细胞裂解物翻译提取物形成 RNA-肽融合体大约 95%.

10 免疫沉淀实验。在设计用于说明免疫沉淀一种 mRNA-肽融合体效力的实验中, 试图免疫沉淀通过体外翻译产生的游离 c-myc 肽。图 10 表示在 SDS-PAGE 肽凝胶上测定的上述实验的结果。泳道 1 和 2 表明源于翻译反应的标记材料含有 RNA124 (LP154 的 RNA 部分) 或 β -珠蛋白 mRNA。泳道 3-8 表示在几个不同的缓冲条件下 (如下文所述) 使用 c-myc 单克隆抗体 9E10 的上述反应样品的免疫沉淀。泳道 3-5 表示源于 RNA124 的肽能有效进行免疫沉淀, 最好的结果出现在第四泳道, 其中, 提取了总 TCA 可沉淀部分的大约 83%。泳道 6-8 显示很少的 β -珠蛋白, 表明纯化程度>100 倍。以上结果表明由 RNA124 以及 LP154 编码的肽可以通过该免疫沉淀方法定量提取。
15

20 融合体的免疫沉淀。随后, 我们用 LP154 翻译反应物和 c-myc 单克隆抗体 9E10 试验了对嵌合 RNA-肽产物进行免疫沉淀的能力 (图 11), 通过免疫沉淀提取源于网织红细胞反应的翻译产物 (如本文所披露的), 并用 1 μ g RNaseA 在室温下处理 30 分钟, 以除去编码序列。由此产生的 5' OH 通过 T4 多核苷酸激酶用 32 P 标记, 并通过变性 PAGE 测定。图 11 表明, 提取到的产物具有类似于用 RNase 处理 LP154 融合体所产生的 c-myc 表位与 30-P 的融合体的迁移率 (参见上文), 不过, 当仅翻译所述模板的 RNA 部分 (RNA124) 时, 不能产生相应的产物。在图 12 中, 测定了所提取的融合蛋白的量, 并相对未修饰的 30-P 的量 (在该图中未示出) 作图。未修饰的接头与接头-myc 融合体的比例的定量表明, 0.2-0.7% 的输入信息被转化成了融合产物。在存在较高的核糖体/模板比例的情况下, 有较大比例的输入 RNA 转化成融合产物; 在试验过的输入 mRNA 浓度范围内, 每毫升的翻译提取物大约产生 0.8-1.0 $\times 10^{12}$ 融合分子。
25
30

另外, 我们的结果表明, 连接在所述类型 RNA 上的肽是由其 mRNA 编码的, 即新生肽未转移到某些其它 mRNA 的嘌呤霉素上。当接头

(30-P)与长的 myc 模板以高达 20: 1 的比例在翻译提取物中共温育时，未发现交叉转移的迹象，而且游离接头的存在也不会明显降低所产生的长 myc 融合体的量。类似的，短模板 43-P 和长模板 LP154 的共翻译仅产生所述模板单独翻译时出现的融合产物，而未发现具有中间迁移率的产物，这种产物是预计中的短模板与长 myc 肽的融合体。以上结果表明，融合体的形成主要出现在与同一个核糖体结合的新生肽和 mRNA 之间。

顺序提取。为了进一步证实体外翻译的 LP154 模板产物的性质，我们在两种不同类型的层析介质上检验了上述产物的特征。硫代丙基 (TP) 琼脂糖凝胶可以分离含有游离半胱氨酸的产物（例如，LP154 产物，该产物具有一个靠近其 C 末端的半胱氨酸残基）（图 13）。类似地，dT₂₅ 琼脂糖可以分离含有 poly dA 序列的模板（例如，30-P）（图 13）。图 14 表明依次在 TP 琼脂糖凝胶上和在 dT₂₅ 琼脂糖上分离所产生的产物与仅在 dT₂₅ 琼脂糖上分离的产物相同。所述体外翻译产物含有 poly -A 尾和游离巯基的事实有力地说明，该翻译产物是需要的 RNA-肽融合体。

上述结果与合成 mRNA-肽融合体并从体外翻译提取物中回收完整的融合体的能力吻合。所合成的融合体的肽部分似乎具有需要的序列，这一点通过免疫沉淀和利用合适的层析技术进行提取得到了证实。根据上述结果，上述反应是分子内的，而且以模板依赖形式发生。最后，即使模板的修饰低于 1%，该系统也有利于基于大约 10¹³ 个分子的候选复杂性的筛选。

C-Myc 表位回收筛选。为了筛选其它的 c-myc 表位，制备一个含有随机化区域（参见图 7C 和后述）的大的翻译模板文库（例如，10¹⁵ 个成员）。将该文库用于制备大约 10¹²-10¹³ 融合体（如本文所披露的），用抗- c-myc 抗体处理所述融合体（例如，通过免疫沉淀或使用固定在柱上或其它固体支持物上的抗体），以便通过反复的体外筛选富集 c-myc 编码模板。

融合体生成模型。在不固于具体理论的前提下，我们提出了一种融合体形成机制的模型，其中，翻译正常开始，并且延伸进展到开放读框的末端。当核糖体到达该模板的 DNA 部分时，翻译停止。此时，所述复合体面临两种结果：解离新生肽，或将新生肽转移到位于该模

板 3'-末端的嘌呤霉素上。转移反应的效率很可能受到多种因素的控制，这些因素影响停止翻译的复合体的稳定性，并使得 3'-嘌呤霉素残基进入肽基转移酶中心的 A 位点。在转移反应之后，mRNA-肽融合体很可能仍然与核糖体配合，因为已知的释放因子不能水解 RNA 和肽结构域之间的稳定的酰胺键。

延伸的经典模式 (Watson, Bull. Soc. Chim. Biol. 46: 1399 (1964)) 和中间状态模式 (Moazed 和 Noller, 自然 342: 142 (1989)) 需要 A 位点空出，以便嘌呤霉素进入肽基转移酶中心。为了让嘌呤霉素进入空的 A 位点，所述接头必须是环绕核糖体外面的环或者直接由 10 解码位点通过 A 位点到达肽基转移酶中心。本文所披露的资料不能清楚地区分这些不同模式，因为试验过的最短的接头 (21 个核苷酸) 仍然长到足于环绕核糖体的外面。在核糖体结构的某些模式中 (Frank 等, 自然 376: 441 (1995)), mRNA 穿过沿所述解码位点的任一侧分布的通道，在这种情况下，有必要不让接头穿过上述通道，以便嘌呤霉素能通过 A 位点到达肽基转移酶中心。

连接在所述接头上的肽的同质性和长度，证实了新生肽向嘌呤霉素的转移似乎比延伸过程慢。如果在延伸期间嘌呤霉素能与氨酰基 tRNAs 有效竞争，出现在融合产物中的接头-肽融合体在大小方面将会不一致。另外，Met-融合体和未修饰的接头之间在凝胶迁移率方面的相似性表明，核糖体似乎没有读入接头区。 dA_{3n} 应当编码 (赖氨酸)_n，这肯定能降低接头的迁移率。未穿过的 mRNA 的较慢的速率可以解释融合体形成的速率慢于翻译速率。初步结果表明，在低温下延长翻译后温育之后，所产生的融合体量明显增加，这也许是因为增加了用于将新生肽转移到嘌呤霉素的时间。

25

详细材料和方法

下面披露了与体外翻译和试验包括具有 myc 表位标记的融合体在内的 RNA-蛋白融合体相关的详细材料和方法。

序列。上面将多种寡核苷酸用于制备 RNA-蛋白融合体。这些寡 30 核苷酸具有如下序列。

名称	序列
30-P	5'AAA AAA AAA AAA AAA AAA AAA AAA AAA CCP (SEQ ID NO:8)
13-P	5'AAA AAA AAA ACC P (SEQ ID NO: 9)
25-P	5'CGC GGT TTT TAT TTT TTT TCC P (SEQ ID NO: 10)
43-P	5'rGrGrA rGrGrA rCrGrA rArArU rGAA AAA AAA AAA AAA AAA AAA AAA AAA ACC P (SEQ ID NO: 11)
43-P [CUG]	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA AAA AAA AAA AAA ACC P (SEQ ID NO: 12)
40-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA AAA AAA AAA AAA ACC P (SEQ ID NO: 13)
37-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA AAA AAA AAA ACC P (SEQ ID NO: 14)
34-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA AAA AAA ACC P (SEQ ID NO: 15)
31-P	5'rGrGrA rGrGrA rCrGrA rArCrU rGAA AAA AAA AAA AAA ACC P (SEQ ID NO: 16)
LP77	5'rGrGrG rArGrG rArCrG rArArA rUrGrG rArArC rArGrA rArArC rUrGrA rUrCrU rCrUrG rArArG rArArG rArCrC rUrGrA rArC AAA AAA AAA AAA AAA AAA AAA AAA AAA CCP (SEQ ID NO: 1)

LP154 5'rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rA
 rArUrG rGrCrU rGrArA rGrArA rCrArG rArArA rCrUrG rArUrC rUrCrU rGrArA
 rGrArA rGrArC rCrUrG rCrUrG rCrGrU rArArA rCrGrU rCrGrU rGrArA rCrArG
 rCrUrG rArArA rCrArC rArArA rCrUrG rGrArA rCrArG rCrUrG rCrGrU rArArC
 rUrCrU rUrGrC rGrCrU AAA AAA AAA AAA AAA AAA AAA CCP
 (SEQ ID NO: 3)

LP160 5' 5'rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rA
 rArUrG rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS
 rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS
 rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rNrNrS rCrArG rCrUrG
 rCrGrU rArArC rUrCrU rUrGrC rGrCrU AAA AAA AAA AAA AAA AAA AAA CCP (SEQ ID NO: 17)

所有寡核苷酸是沿 5' → 3' 方向排列的。核糖核苷酸碱基通过在该核苷酸符号前面加上小写“r”表示的；P 是嘌呤霉素；rN 表示相同量的 rA、rG、rC 和 rU；rS 表示相同量的 rG 和 rC；而且，所有其它碱基符号都表示 DNA 寡核苷酸。

化合物。盐酸嘌呤霉素、长链烷基胺可控孔度玻璃、谷氏菌素、氯霉素、维及尼霉素、DMAP、二甲基三苯甲基氯、以及乙酸酐获自 Sigma Chemical (St. Louis, MO)。吡啶、二甲基甲酰胺、甲苯、琥珀酸酐、和对硝基苯酚获自 Fluka Chemical (Ronkonkoma, NY)。
 10 β - 珠蛋白 mRNA 获自 Novagen (Madison, WI)。TMV RNA 获自 Boehringer Mannheim (Indianapolis, IN)。

酶。蛋白酶 K 获自 Promega (Madison, WI)。无 DNase 的 RNase 酶通过 Sambrook 等的方法(同上)生产或者从 Boehringer Mannheim 公司购买。T7 聚合酶是通过由 Zawadzki 和 Gross 改进的(核酸研究 19: 1948 (1991))、由 Grodberg 和 Dunn 公开的方法(细菌学杂志 170: 1245 (1988))制备的。T4 DNA 连接酶获自 New England Biolabs (Beverly, MA)。

放射性标记物掺入的定量。对于放射性凝胶带来说，通过在一台

Betagen 603 印迹分析仪 (Betagen, Waltham, MA) 上定量或使用磷成像板 (分子动力学, Sunnyvale, CA) 测定每一条带中所存在的放射性标记物 (^{35}S 或 ^{32}P) 的量。对于液体和固体样品而言，通过闪烁计数 (Beckman, Columbia, MD) 测定存在的放射性标记物 ^{35}S 或 ^{32}P 5 的量。

凝胶成像。通过放射自显影 (使用柯达 XAR 胶片) 或使用磷成像板 (分子动力学) 获得凝胶图象。

CPG 嘧啶霉素的合成。合成 CPG-嘌呤霉素的详细方案如上文所述。

10 酶促反应。一般，利用大肠杆菌提取物制备激酶、转录、PCR、和翻译反应的核酸的方法是相同的。每一种制备方法均由用等体积的 1: 1 的苯酚-氯仿萃取开始，然后离心并分离水相。分别以 300mM 和 1mM 的最终浓度加入乙酸钠 (pH5.2) 和亚精胺，添加 3 倍体积的 100% 15 乙醇，并在 -70°C 下温育 20 分钟以使所述样品沉淀。以大于 12,000g 的速度离心样品，除去上清液，用过量的、温度为 0°C 的 95% 乙醇洗涤沉淀。然后在真空条件下干燥所得到的沉淀，并重新悬浮。

20 寡核苷酸。合成的所有 DNA 和 RNA 都是在一台 Millipore Expedite 合成仪上使用由生产商 (Milligen, Bedford, MA) 提供的各自的标准化学法合成的。含有 3' 嘧啶霉素的寡核苷酸是使用装有 30-50mg 固体支持物的 CPG 嘧啶霉素柱合成的 (大约 20 μmol 嘧啶霉素/克)。含有 3' 生物素的寡核苷酸是使用购自 Glen Research (Sterling VA) 的 1 μmol bioteg CPG 柱合成的。含有 5' 生物素的 25 寡核苷酸是通过添加 bioteg 亚磷酸酰胺 (Glen Research) 作为 5' 碱基而合成的。待连接到 RNA 分子 3' 末端的寡核苷酸要么在去保护之前通过化学方法将其 5' 末端磷酸化 (使用购自 Glen Research 的化学磷酸化试剂)，要么在去保护之后使用 ATP 和 T4 多核苷酸激酶 (New England Biolabs) 酶促磷酸化。通过添加 25% 的 NH_4OH 并在 55°C 下温育 12 小时将仅含有 DNA (以及 3' 嘧啶霉素或 3' 生物素) 的样品去保护。通过将乙醇 (25% (v/v)) 添加到 NH_4OH 溶液中并在 55°C 下温 30 育 12 小时将含有 RNA 单体 (例如，43-P) 的样品去保护。在室温下，用溶解在 THF (Sigma) 中的 1M TBAF 将 2' OH 去保护 48 小时。用 NAP-25 Sephadex 柱 (Pharmacia, Piscataway, NJ) 除去 TBAF。

然后，通过使用变性 PAGE 纯化去保护的 DNA 和 RNA 样品，然后浸泡或使用 Elutrap (Schleicher 和 Schuell, Keene, NH) 从凝胶中电洗脱，并使用上述的 NAP-25 Sephadex 柱或乙醇沉淀方法脱盐。

Myc DNA 构建。构建两种含有 c-myc 表位标记的 DNA 模板。第一种模板是通过以下两种寡核苷酸的组合制成的：64.27 (5'-GTT CAG GTC TTC TTG AGA GAT CAG TTT CTG TTC CAT TTC GTC CTC CCT ATA GTG AGT CGT ATT A-3') (SEQ ID NO: 18) 和 18.109 (5'-TAA TAC GAC TCA CTA TAG-3') (SEQ ID NO: 19)。用该模板转录产生的 RNA47.1 能编码肽 MEQKLISEEDLN (SEQ ID NO: 20)。将 RNA47.1 连接到 30-P 上得到图 7A 所示的 LP77。

首先，以单链寡核苷酸的形式制备第二种模板，该模板长 99 个碱基，名称为 RWR99.6，序列为 5'-AGC GCA AGA GTT ACG CAG CTG TTC CAG TTT GTG TTT CAG CTG TTC ACG ACG TTT ACG CAG CAG GTC TTC TTC AGA GAT CAG TTT CTG TTC AGC CAT-3' (SEQ ID NO: 21)。通过 PCR 构建含有该序列的双链转录模板，PCR 是使用寡聚物 RWR21.103 (5'-AGC GCA AGA GTT ACG CAG CTG-3') (SEQ ID NO: 22) 和 RWR63.26 (5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG GCT GAA GAA CAG AAA CTG-3') (SEQ ID NO: 23) 按照公开的方案进行的 (Ausubel 等，同上，第 15 章)。用该模板转录产生的 RNA 被称为 RNA124，它编码肽 MAEEQKLISEEDLLRKRREQLKHGLEQLRNSCA (SEQ ID NO: 24)。该肽含有用于制备（与载体蛋白结合）单克隆抗体 9E10 的序列（癌基因科学技术公报）。RNA124 的长度为 124 个核苷酸，将 RNA124 与 30-P 连接产生如图 7B 所示的 LP154。RNA124 的序列如下 (SEQ ID NO: 32)：

5'-rGrGrG rArCrA rArUrU rArCrU rArUrU rUrArC rArArU rUrArC rArArUrG
 rGrCrU rGrArA rGrArA rCrArG rArArA rCrUrG rArUrC rUrCrU rGrArA rGrArA
 rGrArC rCrUrG rCrUrG rCrGrU rArArA rCrGrU rCrGrU rGrArA rCrArG rCrUrG
 rArArA rCrArC rArArA rCrUrG rGrArA rCrArG rCrUrG rCrGrU rArArC rUrCrU
 rUrGrC rGrCrU-3'

随机化库的构建。以单链寡核苷酸的形式构建随机化库，其长度为 130 个碱基，并命名为 RWR130.1。从 3' 末端开始，其序列为 3'-CCCTGTTAATGATAAATGTTAACCGA-5' (SEQ ID NO: 25)。N 表示随机位点，该序列是通过标准合成方法产生的。S 表示 dG 和 dC 碱基的等量混合物。PCR 是使用寡核苷酸 42.108 (5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA) (SEQ ID NO: 26) 和 21.103 (5'-AGC GCA AGA GTT ACG CAG CTG) (SEQ ID NO: 27) 进行的。该模板转录所产生的 RNA 被命名为库 130.1。将库 130.1 与 30-P 连接得到图 7C 所示的 1 号库 (也被称作 LP160)。

按照公开的方法 (Ausubel 等, 同上) 进行 7 轮 PCR，仅有以下例外：(i) RWR130.1 的起始浓度为 30 纳摩尔，(ii) 每一种引物的使用浓度为 1.5 μM，(iii) 每一种碱基的 dNTP 浓度为 400 μM，和(iv) Taq 聚合酶 (Boehringer Mannheim) 的使用浓度为每 100 μl 五个单位。在非变性 PAGE 上纯化双链产物，并通过电洗脱提取。通过 260nm 波长下的 UV 吸收值和与已知的标准物进行溴化乙啶荧光比较测定 DNA 的量。

RNA 的酶促合成。由双链 PCR DNA 进行的转录反应和寡核苷酸的合成是按照以前披露的方法进行的 (Milligan 和 Uhlenbeck, 酶学方法 180: 51 (1989))。按照上述方法通过变性 PAGE 纯化完整长度的 RNA，电洗脱，并脱盐。估算库 RNA 浓度时使用的消光系数为 13000. D. / μM; RNA124, 1250 O.D. / μM; RNA47.1, 4800. D. / μM。由双链库 DNA 转录产生大约 90 纳摩尔的库 RNA。

RNA 嘧呤霉素缀合物的酶促合成。通过类似于由 Moore 和 Sharp (科学 250: 992 (1992)) 所披露的方法，用被称为 19.35 (5'-TTT TTT TTT TAG CGC AAG A) (SEQ ID NO: 28) 的 DNA 裂片将 myc 和库 mRNA 序列连接到含有嘌呤霉素的寡核苷酸上。反应物由摩尔比为 0.8: 0.9: 1.0 的 mRNA、裂片、和嘌呤霉素寡核苷酸 (30-P, dA₂₇dCdCP) 组成，并且，每皮摩尔库 mRNA 含有 1-2.5 单位的 DNA 连接酶。反应在室温下进行 1 小时。为了构建库 RNA 融合体，mRNA 的浓度为大约 6.6 μM。在连接之后，按照上述酶促反应的方法制备 RNA-嘌呤霉素缀合物。重新悬浮沉淀，在变性 PAGE 上纯化全长融合体，并通过上

述电洗脱提取。使用 1650 O.D. / μM 的消光系数估算库 RNA 浓度，并用 1600 O.D. / μM 的消光系数估算 myc 模板的浓度。这样，制备了 2.5 纳摩尔的缀合物。

5 制备 dT_{25} 链霉抗生物素琼脂糖。在室温下，在 TE(10mM Tris-HCl pH8.2, 1 μM EDTA) 中以 1-10 μM 的浓度将含有 3' 生物素的 dT_{25} (在 bioteg 亚磷酰胺柱 (Glen Research) 上合成) 与链霉抗生物素琼脂糖浆体 (体积为 50% 的琼脂糖, Pierce, Rockford, IL) 一起温育 1 小时，并洗涤。然后通过生物素- dT_{25} 从溶液中的消失和/或通过用已知量的互补寡核苷酸滴定所述树脂，通过光学方法估计所述琼脂糖的结合能力。
10

15 用大肠杆菌产生的提取物和核糖体进行的翻译反应。一般，翻译反应是用购买的试剂盒 (例如，用作线性模板的大肠杆菌 S30 提取物, Promega, Madison, WI) 进行的。不过，也可使用大肠杆菌 MRE600 (获自 ATCC, Rockville, MD) 按照公开方法 (例如，Ellman 等，酶学方法 202: 301 (1991)) 制备 S30 提取物，并通过 Kudlicki 等所披露的方法 (分析生物化学 206: 389 (1992)) 制备核糖体部分。标准反应是在以 20-40 μCi 的 ^{35}S 甲硫氨酸为标记的 50 μl 体积中进行的。该反应混合物包括 30% 的提取物 v/v, 9-18 mM MgCl₂, 40% 的无甲硫氨酸预混合物 (Promega) v/v 和 5 μM 模板 (例如，43-P)。为了进行共温育试验，以 5 μM 的浓度添加寡聚物 13-P 和 25-P。对于使用核糖体的试验来说，将 3 μl 核糖体溶液加入每一个反应中，取代所述裂解物。所有反应物均在 37°C 下温育 30 分钟。在酶促反应下按上述方法纯化模板。
20

25 小麦胚翻译反应。图 8 所示的翻译反应是使用购买的缺少甲硫氨酸的试剂盒 (Promega)，按照生产商的推荐进行的。对于 43-P 来说，模板浓度为 4 μM ，而对于 LP77 和 LP154 来说，模板浓度为 0.8 μM 。反应是在 25°C 下进行，在总体积为 25 μl 的反应物中含有 30 μCi 的 ^{35}S 甲硫氨酸。
30

网织红细胞翻译反应。用购买的试剂盒 (Novagen, Madison, WI) 或使用按照公开方案 (Jackson 和 Hunt, 酶学方法 96: 50 (1983)) 制备的提取物进行翻译反应。富含网织红细胞的血液获自 Pel-Freez Biologicals (Rogers, AK)。在以上两种情况下，反应条件为推荐

用于 Red Nova 裂解物 (Novagen) 的方法。反应物包括 100mM KCl, 0.5mM 乙酸镁, 2mM DTT, 20mM HEPES pH7.6, 8mM 磷酸肌酸, 25 μM 的每一种氨基酸 (甲硫氨酸例外, 如果使用 ^{35}S Met 的话), 和 40% v/v 的裂解物。在 30°C 下温育 1 小时。模板浓度取决于试验, 但通常在 5 50nM-1 μM 的范围内, 43-P 除外 (图 6H), 它的浓度是 4 μM。

为了制备随机化库, 以大约 0.1 μM 的模板浓度 (1.25 纳摩尔的模板) 进行 10ml 的翻译反应。另外, 在该反应中包括 ^{32}P 标记的模板, 以便确定在纯化和筛选过程的每一个步骤中所存在的材料的量。在 30 °C 下翻译 1 小时, 然后将反应物在冰上冷却 30-60 分钟。

10 用 dT_{25} 链霉抗生物素琼脂糖提取融合体。温育之后, 将翻译反应物用提取缓冲液 (1.0M NaCl, 0.1M Tris-HCl pH8.2, 10mM EDTA, 1mM DTT) 稀释大约 150 倍, 该缓冲液含有摩尔数超过 10 倍以上的 dT_{25} -生物素-链霉抗生物素琼脂糖, 其 dT_{25} 浓度为大约 10 μM (浆体的体积等于或大于裂解物的体积), 并在 4°C 下搅拌温育 1 小时。然后通过过滤 (Millipore ultrafree MC 滤器) 或离心从该混合物中除去琼脂糖, 并用冷却的提取缓冲液洗涤 2-4 次。然后用 50-100 μl 的 15mM NaCl, 1 mM EDTA 反复洗涤, 将所述模板从 dT_{25} 链霉抗生物素琼脂糖上释出。将洗脱液立即中和在 3M 乙酸钠 pH5.2, 10mM 亚精胺中, 并用乙醇沉淀。对于所述库反应来说, 回收到的总的放射性表明, 大约回收到了 50-70% 的投入模板。

20 用硫代丙基琼脂糖凝胶分离融合体。如图 13 所示, 可以使用硫代丙基琼脂糖凝胶 6B (Pharmacia) 纯化含有半胱氨酸的融合体。在本文所披露的实验中, 分离是直接用翻译反应物进行的或者在初步分离所述融合体之后 (例如, 用链霉抗生物素琼脂糖) 进行的。例如, 直接纯化时使用比例为 1: 10 (v/v) 的裂解物和琼脂糖凝胶。对于所述库来说, 将 0.5ml 的琼脂糖凝胶浆体用于从 5ml 反应混合物中分离所有的融合材料。将样品稀释到溶解在 1XTE8.2 (10mM Tris-C1, 1 mM EDTA, pH8.2) 中的硫代丙基琼脂糖凝胶的 50: 50 (v/v) 浆体中, 所述缓冲液含有无 RNase 的 DNase (Boehringer Mannheim), 并在 4°C 下旋转温育 1-2 小时, 以便充分反应。除去多余的液体, 用含有 20mM DTT 的提取缓冲液反复洗涤琼脂糖凝胶, 并通过离心或过滤回收。用溶解在 10mM Tris-C1 pH8.2, 1 mM EDTA 的 25-30 mM 二

硫苏糖醇 (DTT) 溶液将融合体从琼脂糖凝胶上洗脱。然后通过上述在高真空下蒸发和乙醇沉淀的组合方法浓缩该融合体。对于库反应来说，所回收到的总的放射活性表明，大约 1% 的模板被转化成融合体。

对某些用途来说，将 dT₂₅ 加入该洗出物中，并在 4°C 下旋转温育 5 小时。用冰镇的提取缓冲液漂洗琼脂糖 3 次，通过过滤分离，并按照上述方法洗脱结合材料。加入载体 tRNA，并用乙醇沉淀融合产物。将样品重新悬浮在含有无 RNaseA 的 DNase 的 TE pH8.2 中，以除去所述模板的 RNA 部分。

免疫沉淀反应。从翻译反应物中分离肽的免疫沉淀（图 10）是这样进行的：将 4 μl 网织红细胞翻译反应物、2 μl 正常小鼠血清、和 20 μl 蛋白 G+A 琼脂糖（Calbiochem，La Jolla, CA）与 200 μl 的 PBS（58 mM 磷酸氢二钠、17 mM 磷酸二氢钠、68 mM 氯化钠）、稀释缓冲液（10 mM Tris-Cl pH8.2, 140 mM 氯化钠, 1% v/v Triton X-100）、或 PBSTDS（PBS+1% Triton X-100, 0.5% 脱氧胆酸, 0.1% SDS）混合。然后在 4°C 下转动样品 1 小时，接着以 2500 rpm 的速度离心 15 分钟。除去洗脱液，加入 10 μl c-myc 单克隆抗体 9E10（Calbiochem，La Jolla, CA）和 15 μl 蛋白 G+A 琼脂糖，并在 4°C 下转动 2 小时。然后用 1 ml 体积的 PBS、稀释缓冲液、或 PBSTDS 洗涤样品 2 次，将 40 μl 的凝胶加样缓冲液（Calbiochem 产品简报）加入所述混合物中，并将 20 μl 混合物加样到变性 PAGE 上，如 Schagger 和 von Jagow 所披露的（分析生物化学 166: 368(1987)）。

将 8 μl 网织红细胞翻译反应物与 300 μl 稀释缓冲液（10 mM Tris-Cl pH8.2, 140 mM 氯化钠, 1% v/v Triton X-100），15 μl 蛋白 G 琼脂糖凝胶（Sigma），10 μl (1 μg) c-myc 抗体 9E10（Calbiochem）混合，然后在 4°C 下转动数小时。分离之后洗涤样品，用无 RNase A 的 DNase 处理，用多核苷酸激酶和 ³²P γ ATP 标记，并通过变性尿素 PAGE 分离（图 11）。

融合体库的逆转录。按照 Superscript II 的生产商的推荐进行逆转录反应，所不同的是，模板、水和引物仅在 70°C 下温育 2 分钟（Gibco BRL Grand Island, NY）。为了监测延伸反应，将 50 μCi α ³²P dCTP 加入某些反应物中。在其它反应中，用 5' ³²P-标记的引物监测逆转录，所述引物是用 ³²P α ATP (New England Nuclear, Boston,

MA) 和 T4 多核昔酸激酶 (New England Biolabs, Beverly, MA) 制备的。

制备蛋白 G 和抗体琼脂糖凝胶。用稀释缓冲液 (10 mM Tris-HCl pH8.2, 140mM NaCl, 0.025%的 NaN₃, 1% v/v Triton X-100) 洗涤 5 2 份 50 μl 的蛋白 G 琼脂糖凝胶浆体 (体积百分比为 50%的固体) (Sigma)，并通过离心分离。保存第一份样品，以便用作筛选基质之前预装柱。将第二份样品重新悬浮在稀释缓冲液里，然后加入 40 μg c-myc AB-1 单克隆抗体 (Oncogene Science)，在 4°C 下旋转温育该反应物过夜。在微量离心机中以 1500–2500 rpm 的速度离心 10 分钟，纯化抗体琼脂糖凝胶，并用稀释缓冲液洗涤 1–2 次。

筛选。提取融合体并合成互补链，然后将整个逆转录反应物直接用于筛选过程。本文披露了两种方法。第一轮，将逆转录反应物直接加入按上述方法制备的抗体琼脂糖凝胶中，并温育 2 小时。在随后的轮次中，将所述反应物与洗涤过的蛋白 G 琼脂糖凝胶温育大约 2 小时，然后再加入抗体柱上，以减少与蛋白 G 而不是固定化抗体相互作用的结合体的数量。

为了从所述基质上洗脱所述库，可以采用几种方法。第一种方法是用 4% 的乙酸洗涤所述筛选基质。该方法能将肽从所述基质上释出。另外，可以用更严格的洗涤（例如使用尿素或其它变性剂）取代所述乙酸方法或者与该方法一起使用。

筛选的融合体的 PCR。用上文所述的用于构建所述库的标准方法，通过 PCR 扩增筛选的分子。

β -珠蛋白融合体的合成和测试

25 为了合成 β -珠蛋白融合体结构，按照生产商推荐的方法通过用 200 皮摩尔引物 18.155 (5'GTG GTA TTT GTG AGC CAG) (SEQ ID NO: 29) 和 Superscript 逆转录酶 (Gibco BRL) 进行逆转录，由 2.5 μg 珠蛋白 mRNA 制备 β -珠蛋白 cDNA。引物的序列互补于 β -珠蛋白的 5' 终止密码子的 18 个核昔酸。为了增加一个 T7 启动子，取出 20 30 μl 逆转录反应物，并用引物 18.155 和 40.54 (5'TAA TAC GAC TCA CTA TAG GGA CAC TTG CTT TTG ACA CAA C) (SEQ ID NO: 30) 进行 6 个轮次的 PCR。然后按照 Milligan 和 Uhlenbeck 披露的方法（酶学

方法 180: 51 (1989)), 通过 T7 失控转录制备“合成- β -珠蛋白” mRNA, 按照本文所披露的方法对该 RNA 进行凝胶纯化, 电洗脱, 并脱盐。然后按照 Moore 和 Sharp 的方法 (科学 256: 992 (1992)) 使用引物 20.262 (5' TTT TTT TTT T GTG GTA TTT G) (SEQ ID NO: 5 31) 作为裂片, 通过将合成- β -珠蛋白结构连接到 30-P 上产生“LP- β -珠蛋白”。然后对连接反应的产物进行上文所述的凝胶纯化, 电洗脱, 以及脱盐。通过其在 260nm 波长下的吸收值测定最终产物的浓度。

然后如表 1 所示各自在 25 μ l 的总体积中体外翻译上述 β -珠蛋白模板。添加 25mM 的 Mg²⁺母液。所有反应均在 30°C 下温育 1 小时, 10 并放在-20°C 下过夜。然后用 6 μ l 的裂解物和平均负背景测定可由 dT₂₅ 沉淀的 CPM's 两次。

表 1

用 β -珠蛋白模板进行的翻译反应

15	反应	样品	Mg ²⁺ (mM)	³⁵ S-Met (μ l)	TCA CPM (2 μ l)	dT ₂₅ CPM (6 μ l)
1		---	1.0	2.0 (20 μ Ci)	3312	0
2		2.5 μ g	0.5	2.0 (20 μ Ci)	33860	36
	合成-β-珠蛋白					
20	3	2.5 μ g	1.0	2.0 (20 μ Ci)	22470	82
	合成-β-珠蛋白					
4		2.5 μ g	2.0	2.0 (20 μ Ci)	15696	86
	合成-β-珠蛋白					
25	5	2.5 μ g	0.5	2.0 (20 μ Ci)	32712	218
	LP-β-珠蛋白					
6		2.5 μ g	1.0	2.0 (20 μ Ci)	24226	402
	LP-β-珠蛋白					
7		2.5 μ g	2.0	2.0 (20 μ Ci)	15074	270
	LP-β-珠蛋白					

30

为了制备凝胶分析所使用的样品, 将 6 μ l 的每一种翻译反应物与 1000 μ l 的提取缓冲液 (1M NaCl 100mM Tris-C1 pH8.2. 10mM

EDTA, 0.1mM DTT), 1 μ l RNase A(无 DNase, Boehringer Mannheim), 和 20 μ l 的 20 μ M 的 dT_{25} 链霉抗生物素琼脂糖混合。在 4°C 下旋转温育样品 1 小时。除去过量的提取缓冲液，将该样品加入 Millipore MC 滤器中，以除去所有残留的提取缓冲液。然后用 50 μ l 的水洗涤样品 5 次，用 50 μ l 的 15mM NaOH, 1mM EDTA 洗涤两次。用 100 μ l TE pH6.8 (10mM Tris-Cl, 1mM EDTA) 中和样品 (300 μ l)，加入 1 μ l 的 1mg/ml RNaseA (同上)，在 37°C 下温育样品，然后加入 10 μ l 的 2XSDS 加样缓冲液 (125mM Tris-Cl pH6.8, 2% SDS, 2% β -巯基乙醇, 20% 甘油, 0.001% 溴酚兰)，将该样品冻干，并重新悬浮在 20 μ l 水和 10 1% β -巯基乙醇中。然后按照 Schagger 和 von Jagow 所披露的方法 (分析生物化学 166: 368 (1987)) 将样品加样到肽分离凝胶上，并通过放射自显影观察。

以上实验的结果示于图 15A 和 15B 中。如图 15A 所示， 35 S-甲硫氨酸结合到了合成- β -珠蛋白和 LP- β -珠蛋白融合体的蛋白部分中。该蛋白是不一致的，不过一条强的带表明了预期的 β -珠蛋白 mRNA 的迁移率。另外，如图 15B 所示，在 dT_{25} 分离和 RNaseA 消化之后，在合成- β -珠蛋白泳道上不再有 35 S-标记的材料 (图 15B, 泳道 2-4)。相反，在 LP- β -珠蛋白泳道上，观察到了大小一致的 35 S-标记的产物。

上面的结果表明，只有当模板含有 3' 嘧啶霉素时才能通过寡核苷酸亲和层析提取到融合产物。这一结果得到了闪烁计数 (参见表 1) 的证实。所获得的材料预计含有与 β -珠蛋白的某一部分融合的 30-P 接头。通过凝胶分析判断该融合产物的大小似乎十分均匀。不过，由于该产物具有非常类似于天然 β -珠蛋白的迁移率 (如 15A 和 15B, 对照泳道)，难于确定该融合产物的蛋白部分的确切长度。

进一步优化 RNA-蛋白融合体的形成

业已发现某些因素能进一步提高 RNA-肽融合体形成的效率。融合体的形成，即新生肽链由其 tRNA 向位于 mRNA 3' 末端的嘌呤霉素的转移是一种缓慢的反应，该反应是继最初的、较快的开放读框的翻译之后进行的，所述翻译过程产生新生肽。通过翻译后在高 Mg²⁺ 条件下 (优选地，在 50-100mM 范围内) 温育和/或在 RNA 和嘌呤霉素部分

之间使用更有柔性的接头可明显加强融合体形成的程度。另外，在低温下（优选-20°C）长时间（12-48 小时）温育，也能导致融合体产量的提高，与在 30°C 下温育相比，有较少的 mRNA 降解。通过综合以上因素，可将高达 40% 的投入的 mRNA 转化成 mRNA-肽融合产物，如下文所述。

mRNA-嘌呤霉素缀合物的合成。在这些优化实验中，在有互补 DNA 裂片的条件下用噬菌体 T4 DNA 连接酶将含有嘌呤霉素的接头寡核苷酸连接到 mRNA 的 3' 末端，基本上如上文所述，因为 T4 DNA 连接酶倾向于在连接接点附近的精确碱基配对，T7、T3、或 SP6 RNA 聚合酶的失控转录产物在其 3' 末端通常是一致的（核酸研究 15: 8783 (1987)），仅有含有正确的 3'-末端核苷酸的 RNA 能有效地连接。当使用标准 DNA 裂片时，大约有 40% 的失控转录产物连接到嘌呤霉素寡聚物上。使用过量的 RNA 可以提高连接产物的量，但使用过量的嘌呤霉素核苷酸不能提高连接产物的量。并不囿于特定的理论，连接的限制因素似乎是完全互补于所述 DNA 裂片的相应部分的 RNA 的量。

为了连接其 3' 末端具有一个额外的非模板核苷酸的转录物（被称为“N+1 产物”），使用标准 DNA 裂片和在连接结合点处具有一个额外的随机碱基的新的 DNA 裂片的混合物。在有上述混合 DNA 裂片的情况下，可将一种代表性 myc RNA 模板（即 RNA124）的连接效率提高到高于 70%。

除了上述改进的 DNA 裂片方法之外，通过考虑以下 3 个因素可进一步优化 mRNA-嘌呤霉素缀合物形成的效率。首先，优选设计或使用这样的 mRNAs：它缺少具有任何明显的稳定二级结构的 3' 末端，这种结构会干扰与裂片寡核苷酸的退火。另外，由于高浓度的盐有时会导致连接反应的失败，该方法优选包括一个用 NAP-25 柱对所述寡核苷酸进行彻底脱盐的步骤。最后，由于连接反应比较迅速，而且，在室温下通常在 40 分钟之内结束，一般不会使用较长的温育时间，而且，较长的温育时间通常会导致 RNA 的不必要降解。

使用上述条件按如下方法合成了 mRNA-嘌呤霉素缀合物。用标准 DNA 裂片（例如，5'-TTTTTTTTTAGCGCAAGA）或在连接点处含有随机碱基（N）的裂片（例如，5'-TTTTTTTTNAGCGCAAGA）将 myc RNA 序列（RNA124）连接到含有嘌呤霉素的寡核苷酸上。反应物包括

mRNA、DNA 裂片、和嘌呤霉素寡核苷酸，其摩尔比为 1.0: 1.5-2.0: 1.0。首先将上述成分的混合物在 94°C 下加热 1 分钟，然后在冰上冷却 15 分钟。连接反应是在室温下，在 50mM Tris-HCl (pH7.5)，10mM MgCl₂、10mM DTT、1mM ATP、25 μg/ml BSA、15 μM 嘌呤霉素寡聚物、5 15 μM mRNA、22.5-30 μM DNA 裂片、浓度为 1U/μl 的 RNase 蛋白抑制剂 (Promega)、和每皮摩尔嘌呤霉素寡聚物 1.6 单位的 T4 DNA 连接酶中进行的。在温育之后，将 EDTA 以 30mM 的终浓度加入，并用苯酚/氯仿萃取反应混合物。通过变性 PAGE 纯化完整长度的缀合物，并通过电洗脱提取。

10 通用网织红细胞翻译条件。除了改善 mRNA-嘌呤霉素的缀合物的合成之外，还要对翻译反应作如下进一步的优化。反应是在由不同商业渠道获得的兔网织红细胞裂解物中进行的 (Novagen, Madison, WI; Amersham, Arlington Heights, IL; Boehringer Mannheim, Indianapolis, IN; Ambion, Austin, TX; 和 Promega, Madison, WI)。典型的反应混合物 (25 μl 终体积) 包括 20mM HEPES pH7.6, 2mM DTT, 8mM 磷酸肌酸, 100mM KCl, 0.75mM 乙酸镁, 1mMATP, 0.2mM GTP, 25 μM 的每一种氨基酸 (如果使用 ³⁵S -Met 的话, 0.7 μM 的甲硫氨酸)，浓度为 1U/μl 的 RNase 蛋白抑制剂，和 60% (v/v) 裂解物。模板的终浓度在 50-800nM 的范围内。对于每一种温育来说，将除裂解物之外的所有成分在冰上仔细混合，并在使用之前将冰冻的裂解物解冻。在添加裂解物之后，通过温和的移液重复混合该反应混合物，并在 30°C 下温育以便开始翻译。对于不同的 mRNAs 来说，Mg²⁺ 和 K⁺ 的最佳浓度分别在 0.25-2mM 和 75-200 mM 的范围内变化，并且优选通过预备实验确定。特别地，对于翻译较差的 mRNAs 来说，有时也要优化氯高铁血红素、磷酸肌酸、tRNA、和氨基酸的浓度。对于融合反应来说，KCl 通常要优于乙酸钾，不过，氯化钾和乙酸钾的混合物有时能产生更好的效果。

30 在 30°C 下翻译 30-90 分钟，然后将该反应物放在冰上冷却 40 分钟，并加入 Mg²⁺。将在这一步骤中加入的 Mg²⁺ 终浓度相对不同的 mRNA 模板优化，不过，通常在 50-100 mM 范围内 (对于混合模板库来说优选使用 50 mM)。将所得到的混合物在 -20°C 下温育 16-48 小时。为了观察标记过的融合产物，将 2 μl 反应混合物与 4 μl 加样缓冲液混

合，并将该混合物在 75°C 下加热 3 分钟。然后将所得到的混合物加样到 6% 甘氨酸 SDS-聚丙烯酰胺凝胶上 (^{32}P -标记的模板) 或加样到 8% 麦黄酮 SDS-聚丙烯酰胺凝胶上 (^{35}S -Met-标记的模板)。除了该方法之外，还可以用 dT_{25} 链霉抗生素琼脂糖或硫代丙基琼脂糖凝胶 5 (或同时用这两种材料) 分离所述融合产物，基本上如本文所披露的。

为了除去 RNA-接头-嘌呤霉素-肽缀合物的 RNA 部分，以便随后通过 SDS-PAGE 进行分析，在翻译后温育之后添加适量的 EDTA，并用 microcon-10 (或 microcon-30) 柱将所述反应混合物脱盐。将所得 10 到的混合物 (大约共 25 μl) 与 18 μl RNaseH 缓冲液 (30mM Tris-HCl, pH7.8, 30mM 硫酸铵, 8mM 氯化镁, 1.5 β -巯基乙醇，和适量的互补 DNA 裂片) 混合，并将该混合物在 4°C 下温育 45 分钟。然后加入 RNaseH，并在 37°C 下消化 20 分钟。

15 嘌呤霉素寡聚物的质量。对于有效制备融合产物来说，嘌呤霉素寡核苷酸的质量同样是重要的。5'-DMT, 2'-琥珀酰, N-三氟乙酰嘌呤霉素与 CPG 的连接不如标准寡核苷酸的连接有效。因此，要仔细监测连接反应，以避免 CPG 与太低浓度的偶联嘌呤霉素的形成，而且将 CPG 上的未反应的氨基完全淬火，以避免随后合成缺少 3'-末端嘌呤霉素的寡核苷酸。避免使用含有极细目数的颗粒的 CPG 同样很重要，20 因为这种颗粒会导致在随后的自动寡核苷酸合成步骤中出现堵塞阀的问题。

另外，优选在大规模使用之前试验合成的嘌呤霉素寡聚物，以确保嘌呤霉素在 3'-末端的存在。在我们的实验中，如果嘌呤霉素被主要在其 3'-末端含有氨基的脱氧腺苷所取代，则监测不到融合体。为了监测 3' 羟基的存在 (即缺少 3'-末端嘌呤霉素的不希望的寡核苷酸的合成)，可以首先用放射性标记嘌呤霉素寡聚物 (例如，通过 5'-磷酸化)，然后将其用作引物，由末端脱氧核苷酸基转移酶进行延伸。在有 3'-末端嘌呤霉素部分存在的情况下，不会出现延伸产物。

30 翻译和翻译后温育的时间过程。翻译反应比较迅速，而且，通常在 30°C 下在 25 分钟内完成。不过，所述融合反应比较慢。当在 30 °C 下使用标准接头 ($\text{dA}_{27}\text{dCdCP}$) 时，再用 45 分钟融合体的合成会达到最高水平。翻译后温育可以在低温度下进行，例如在室温、0°C、

或-20℃下进行。在-20℃下发现 mRNA 模板的降解较少，在-20℃下温育两天之后，获得了最佳的融合效果。

Mg²⁺浓度的作用。 在翻译后温育中，高浓度的 Mg²⁺可大大促进融合体的形成。例如，对于上述 mycRNA 模板来说，在-20℃下进行的 5 16 小时的温育中，在有 50mM Mg²⁺的条件下，用标准接头 (dA₂₇dCdCP) 可提高融合体形成 3-4 倍 (图 17, 比较泳道 3 和 4)。类似地，当所述反应是在室温下进行 30-45 分钟时，在有 50-100mM Mg²⁺的浓度下进行的翻译后温育中也会出现有效的融合体形成。

接头长度和序列。 还实验了接头长度对融合反应的影响。在 10 21-30 个核苷酸 (n=18-27) 的范围内，观察到了融合反应效率的较小的变化 (如上文所述)。接头越短 (例如，13 个核苷酸长)，效率越低。另外，尽管特定的较大长度的接头 (即，45 个核苷酸和 54 个核苷酸) 有时也会导致较低的效率，但仍然有可能将较长的接头用于优化融合反应的效率。

15 就接头序列而言，用核糖核苷酸残基取代靠近 3' 末端的脱氧核糖核苷酸不会显著改变其融合效率。不过，位于接头 3' 末端 dCdCP (或 rCrCP) 序列对于融合体形成来说很重要。用 dUdUP 取代 dCdCP 明显较低融合体形成的效率。

接头柔性。 还实验了接头柔性对融合反应的影响。在这些实验 20 中，如果所述接头的刚性因为与位于其 3' 末端的互补核苷酸退火而提高，则测得的融合效率低。类似地，当使用柔性更大的接头 (例如，dA₂₇C₉C₉C₉ dAdCdCP，其中，C₉ 表示 HO (CH₂CH₂O)₃PO₂)，可明显提高其融合效率。与标准接头 (dA₂₇dCdCP)，采用柔性更大的接头 (dA₂₇C₉C₉C₉ dAdCdCP) 可以提高 RNA124 的融合效率 4 倍以上 (图 17, 比较泳道 1 和 9)。另外，与具有标准接头的模板相反，这种模板在缺少高浓度 25 Mg²⁺的条件下其翻译后融合进行的很差 (图 17, 比较泳道 3 和 4)，具有柔性接头的模板不需要高浓度 Mg²⁺以便在-20℃下进行的延长的翻译后温育中产生高产量的融合产物 (图 17, 比较泳道 11 和 12)。因此，如果不添加高浓度的 Mg²⁺的话，这种接头是非常有用的。另外，柔性接头也能在有高浓度 Mg²⁺的条件下产生最佳的融合产率。

融合效率的定量。 融合效率可以用翻译肽转化成融合产物的比率

或投入的模板转化成融合产物的比率表示。为了测定翻译肽转化成融合产物的比率，使用³⁵S-Met标记的翻译肽。在这些实验中，当使用dA₂₇dCdCP或dA₂₇rCrCP接头时，在30°C下进行1小时的翻译温育之后，有大约3.5%的翻译肽与其mRNA融合。在-20°C下进行温育过夜之后，这一数值可提高到12%。当翻译后温育是在有高浓度的Mg²⁺的条件下进行时，50%以上的翻译肽与所述模板融合。

对于具有柔性接头的模板来说，在30°C下翻译温育1小时之后，大约25%的翻译肽与模板融合。在-20°C下进行温育过夜之后，这一数值可提高到超过50%，而且如果翻译后温育是在有50mM Mg²⁺的条件下进行，可提高到超过75%。

为了测定所投入的模板转化成融合产物的百分比，用³²P-标记的mRNA-接头模板进行翻译。当使用柔性接头，而且翻译后温育是在不添加Mg²⁺的条件下在-20°C下进行的时，当投入的RNA模板的浓度分别为800、400、200、100、和50nM时，所投入的模板的大约20%、15 40%、40%、35%、和20%被转化成mRNA-肽融合体（图18）。当翻译后温育是在有50mM Mg²⁺的条件下进行时，获得了类似结果。用获自Novagen、Amersham、或Ambion的裂解物获得了最佳结果（图19）。

如果mRNA模板较长的话，通过SDS-PAGE测定的mRNAs和mRNA-肽融合体之间的迁移率差别很小。在这种情况下，可以用³²P标记模板的接头的5'末端。然后在翻译/温育之后，在有互补DNA裂片的条件下，用RNaseH消化长的RNA部分，并通过定量未修饰过的接头与接头-肽融合体的比率测定融合效率。与RNaseA消化相比，这种酶的消化能产生3'-P和5'-OH，该方法所具有的优点是，位于接头5'末端的³²P不会被除去。

在翻译后温育中进行的分子内与分子间融合。除了上述实验之外，我们还试验了在有Mg²⁺的条件下在-20°C下进行的融合反应是分子内性质的还是分子间性质的。在上述翻译和翻译后温育条件下，将游离接头（dA₂₇dCdCP或dA₂₁C₉C₉C₉，dAdCdCP，其中C₉是-O(CH₂CH₂O)₃PO₂）与含有一个DNA接头但在其3'末端没有嘌呤霉素的模板共温育。30 在这些实验中，没有可检测量的（即低于正常水平的2%）³⁵S-Met掺入接头-肽产物中，表明翻译后融合主要是发生在新生肽与结合于相同核糖体上的mRNA之间。

优化结果。如上文所述，通过使用柔性接头和/或在有高浓度 Mg²⁺ 的条件下进行翻译后温育，融合效率可提高到大约为投入 mRNA 的 40%。以上结果表明，每 1 毫升的体外翻译反应混合物可以产生多达 10¹⁴ 个 mRNA-肽融合体分子，产生用于体外筛选实验中的具有极高复 5 杂性的 mRNA-肽融合体库。

RNA-蛋白融合体的筛选富集

我们业已证实了通过基于其所编码的肽将一种特定的 RNA-肽融合体从随机序列融合体的复合库中富集起来而将 RNA-肽融合体用于 10 筛选和进化实验的可行性。具体地讲，我们制备了一系列混合物，其中，将少量的已知序列（在这种情况下是长 myc 模板 LP154）与某种量的随机序列库（即 LP160）混合。翻译该混合物，并通过本文所披露的寡核苷酸和二硫亲和层析筛选 RNA-肽融合产物。用抗-myc 单克隆抗体选择性地免疫沉淀 myc-模板融合体（图 16A）。为了测定在该 15 筛选步骤中获得的富集程度，在有放射性标记的引物的条件下通过 PCR 扩增来自免疫沉淀之前和之后的 cDNA/mRNA-肽融合体混合物的等分试样。用能裂解 myc 模板序列不能裂解所述库的限制性内切酶消化扩增的 DNA（图 16B 和 16C）。对裂解和非裂解 DNA 的比例的定量表明，通过免疫沉淀使 myc 序列相对随机文库富集了 20-40 倍。

20 以上实验是按如下方法进行的。

翻译反应。翻译反应的进行大体上如上文所述。具体地讲，是按照生产商的说明（Novagen）在 30°C 下进行反应 1 小时，并在 -20°C 下冷冻过夜。制备了两种形式的六个样品，一种含有³⁵S 甲硫氨酸，而另一种含有以 52 μM 的终浓度添加的冷却甲硫氨酸。反应物 1-6 含有表 2 所示量的模板。表 2 中的所有数字表示每 25 μl 反应混合物中的模板的皮摩尔数。

表 2
用于掺杂筛选中的模板比例

反应	LP154	LP160
30 1	---	---
2	5	---
3	1	20

4	0.1	20
5	0.01	20
6	---	20

dT₂₅ 链霉抗生物素琼脂糖的制备。用 TE8.2 (10mM Tris-C1 pH8.2, 1mM EDTA) 洗涤链霉抗生物素琼脂糖 (Pierce) 3 次，并以 1:1 (v/v) 浆液的比例重新悬浮于 TE8.2 中。然后加入用 Bioteq CPG (Glen Research) 合成的 3' 生物素酰 T₂₅ 至需要的最终浓度（通常为 10 或 20 μM），并在搅拌条件下温育 1 小时。然后用 TE8.2 洗涤 dT₂₅ 链霉抗生物素琼脂糖 3 次，并在 4°C 下保存待用。

从翻译反应物中纯化模板。为了从翻译反应物中纯化模板，取出每一种反应物 25 μl，并加入 7.5ml 的提取缓冲液 (1M NaCl, 100mM Tris-C1 pH8.2, 10mM EDTA, 0.1mM DTT) 和 125 μl 的 20 μM dT₂₅ 链霉抗生物素琼脂糖中。在 4°C 下旋转温育该溶液 1 小时。对所述试管进行离心，并除去洗脱剂。添加 1ml 的提取缓冲液，重新悬浮该浆体，并将该混合物转移到 1.5ml 的微量离心管中。然后用 1ml 的冰镇提取缓冲液的等分试样洗涤所述样品 4 次。将来自相同反应的热的和冷的样品合并在 Millipore MC 过滤装置中，并通过用 2 体积的 100 μl 水, 0.1 mM DTT 和 2 体积的 15 mM NaOH, 1 mM EDTA 从 dT₂₅ 琼脂糖上洗脱。

向该洗脱液中添加 40 μl 洗涤过的 50% 硫代丙基琼脂糖凝胶 (Pharmacia) 浆体，并在 4°C 下旋转温育 1 小时。然后用 1ml 体积的 TE8.2 洗涤所述样品 3 次，除去洗脱液。将 1 μl 的 1MDTT 加入固体中（总体积大约为 20-30 μl），并温育该样品若干小时，取出，并用 20 μl 的水洗涤 4 次（总体积 90 μl）。通过 UV 吸收值判断洗脱液中含有 2.5mM 的硫代毗啶酮。通过加入 6 μl 3M 乙酸钠 pH5.2, 10 mM 亚精胺, 1 μl 糖原 (10mg/ml, Boehringer Mannheim)，和 170 μl 100% 的乙醇，并在 -70°C 下温育 30 分钟对 50 μl 的上述样品进行乙醇沉淀，并在微量离心机中以 13,000 rpm 的速度离心 30 分钟。

逆转录酶反应。按如下方法对乙醇沉淀的和硫代毗啶酮洗脱的样品进行逆转录反应。对于乙醇沉淀的样品来说，向 30 μl 的再悬浮模板中加水至 48 μl，并让 200 皮摩尔的引物 21.103 (SEQ ID NO: 22) 在 70°C 下退火 5 分钟，并在冰上冷却。向该样品中加入 16 μl 的第

一链缓冲液 (250mM Tris-Cl pH8.3, 375mM KC1, 和 15mM MgCl₂; 购自 Gibco BRL Grand Island, NY), 8 μl 100mM DTT 和 4 μl 10mM NTP, 并在 4°C 下平衡, 加入 4 μl Superscript II 逆转录酶 (Gibco BRL Grand Island, NY)。将水 (13 μl) 加入 TP 琼脂糖凝胶洗脱液 (35 μl) 中, 并按上述方法进行反应。温育 1 小时, 然后合并相同数目的样品 (总体积 160 μl)。保存每一种未筛选样品的 10 μl 样品, 用于 PCR, 并保存 150 μl 的样品用于免疫沉淀。

免疫沉淀。为了进行免疫沉淀, 将 170 μl 的逆转录反应物加入 1ml 的稀释缓冲液 (10mM Tris-Cl, pH8.2, 140mM NaCl, 1% v/v Triton X-100) 和 20 μl 蛋白 G/A 缓合物 (Calbiochem, La Jolla CA) 中, 并在 4°C 下旋转温育 1 小时进行预澄清。除去洗脱液, 加入 20 μl G/A 缓合物和 20 μl 单克隆抗体 (2 μg, 12 皮摩尔), 并将该样品在 4°C 下旋转温育 2 小时。通过以 2500rpm 的速度离心 5 分钟对该缓合物进行沉淀, 除去洗脱液, 并用 1ml 等分试样的冰镇稀释缓冲液洗涤该缓合物 3 次。然后用 1ml 的冰镇 10 mM Tris-Cl pH8.2, 100mM NaCl 洗涤该样品。用 3 倍体积的冷冻 4% 乙酸除去结合的片段, 并将该样品冻干。

筛选过的和未筛选过的样品的 PCR。通过将 20 μl 的浓缩 NH₄OH 添加到 10 μl 的未筛选的材料和全部筛选材料中并分别在 55°C 和 70°C 下温育 5 分钟进行 PCR 反应, 并在 90°C 温育以分解存在于该样品中的所有 RNA。并用真空离心蒸发浓缩器蒸发干燥该样品。向每一种样品中添加 200 μl 的 PCR 混合物 (1 μM 引物 21.103 和 42.108, 200 μM 溶解在 PCR 缓冲液 + Mg²⁺ 中的 dNTP (Boehringer Mannheim), 和 2 μl Taq 聚合酶 (Boehringer Mannheim))。对未筛选过的样品 2 进行 16 个轮次的 PCR, 而对所有的其它样品进行 19 个轮次的 PCR。

然后按照表 3 所示, 在有 ³²P-标记的引物 21.103 的条件下对样品进行扩增, 并用 Wizard 直接 PCR 纯化试剂盒 (Promega) 分别纯化 2 次, 以便除去所有引物和较短的片段。

表 3

筛选过的和未筛选过的 PCR 样品的扩增

样品	类型	体积	轮次
1	未筛选过的	20 μl	5
5 2	未筛选过的	5 μl	4
3	未筛选过的	20 μl	5
4	未筛选过的	20 μl	5
5	未筛选过的	20 μl	5
6	未筛选过的	20 μl	5
10 1	筛选过的	20 μl	5
2	筛选过的	5 μl	4
3	筛选过的	20 μl	5
4	筛选过的	20 μl	7
5	筛选过的	20 μl	7
15 6	筛选过的	20 μl	7

限制性消化。将由上述每一种 PCR 反应制备的 ^{32}P 标记的 DNA 以相同的量（以样品的 cpm 计）加入表 4 所示的限制消化反应物中。每一个反应的总体积为 25 μl。将 0.5 μl 的 AlwNI（5 个单位，New England Biolabs）加入每一种反应物中。在 37°C 温育样品 1 小时，通过在 65°C 下温育 20 分钟使酶失活。然后将该样品与 10 μl 的变性加样缓冲液（1ml 超纯甲酰胺（USB），20 μl 0.5M EDTA 和 20 μl 1M NaOH）混合，加热到 90°C 保持 1 分钟，冷却，并加样到含有 8M 尿素的 12% 的变性聚丙烯酰胺凝胶上。电泳以后，用 10%（v/v）乙酸，10%（v/v）甲醇，水固定所述凝胶。

表 4

限制性消化条件 w/AlwNI

样品	类型	加入反应中的 DNA 的体积	总体积
1	未筛选过的	20 μl	25 μl
2	未筛选过的	4 μl	25 μl
30 3	未筛选过的	20 μl	25 μl
4	未筛选过的	20 μl	25 μl
5	未筛选过的	4 μl	25 μl

6	未筛选过的	20 μl	25 μl
1	筛选过的	20 μl	25 μl
2	筛选过的	8 μl	25 μl
3	筛选过的	12 μl	25 μl
5 4	筛选过的	12 μl	25 μl
5	筛选过的	20 μl	25 μl
6	筛选过的	20 μl	25 μl

消化物的定量。用磷成像仪 (Molecular Dynamics) 对存在于样品中的 myc 与库 DNA 的量进行定量。存在于每一条带中的材料量通过凝胶带周围的相同的矩形的总体积确定。存在于每一条带中的总的 cpm 通过所述体积减去背景来计算。使用了背景的三个数值：(1) 在凝胶上进行计数的部位外面的相同方形的平均数；(2) 可能出现 myc 带的（在凝胶上的这一位置上不会出现带）未筛选的库的泳道中的 cpm 数；和 (3) 可以再现未筛选的泳道之间的 10 倍模板增量的近似值的归一化值。图 16B 和 16C 中的泳道 2、3、和 4 表示靶序列与库序列的富集。由于所述信号与该样品的噪声比的优化，在泳道 3 中可证明的富集（未筛选过的/筛选过的）产生了最大值（用方法 1-3 分别为 17、43 和 27 倍）。以上结果归纳于表 5 中。

表 5

Myc 模板与库的富集

方法	泳道 2 (20)	泳道 3 (200)	泳道 4 (2000)
1	7.0	16.6	5.7
2	10.4	43	39
3	8.7	27	10.2

25

在第二组实验中，用 Wizard 直接 PCR 纯化试剂盒将这些相同的 PCR 产物纯化一次，通过上面的方法 (2) 定量消化物。在这些实验中，获得了类似结果；测出相应于上面的泳道 2、3 和 4 的样品分别富集了 10.7、38 和 12 倍。

30

蛋白筛选系统的应用

本发明的筛选系统在很多领域具有商业用途，其中，蛋白质技术

被用于解决治疗、诊断或工业问题。这种筛选方法可用于改进或改变现有蛋白，并用于提取具有需要的功能的新的蛋白。所述蛋白可以是天然存在的序列，可以是天然存在的序列的改变形式，或者是部分或完全合成的序列。

5 新型结合试剂的提取。 在一种具体用途中，本文所披露的 RNA-蛋白融合技术可用于提取具有特定结合（例如，配体结合）特征的蛋白。具有高度特异的结合相互作用的蛋白可被用作非抗体识别剂，使得 RNA-蛋白融合技术超过传统的单克隆抗体技术。通过该方法提取的抗体-型试剂可用于使用传统抗体的任何领域，包括诊断和治疗目的。
10

15 人抗体的改进。 本发明还可以用于改进人或人源化抗体，以便治疗多种疾病中的任一种。在这种用途中，形成抗体文库，并在体外筛选，消除了对诸如细胞融合或噬菌体展示技术的需要。在一种重要用途中，本发明可用于改善单链抗体文库（Ward 等，自然 341: 544
15 (1989)；和 Goulot 等，分子生物学杂志 213: 617 (1990)）。为了这一用途，其可变区可以由人源构建（以便降低受体不利的免疫反应的可能性）或者可以含有完全随机化的框（以使该文库的复杂性最大）。为了筛选改进的抗体分子，检验候选分子库对一种靶分子（例如，如图 2 所示的固定化抗原）的结合能力。随着筛选从一轮进行到
20 下一轮，对结合步骤施加越来越高的严格条件。为了提高严格性，改变诸如洗涤步骤的次数、过量竞争剂的浓度、缓冲条件、结合反应的时间长度、以及固定化基质筛选的条件。
25

可将单链抗体直接用于治疗或者间接用于标准抗体的设计。这种抗体具有多种潜在用途，包括提取抗-自身免疫抗体，免疫抑制，以及开发诸如 Aids 的病毒性疾病的疫苗。

30 新型催化剂的提取。 本发明还可用于筛选新型催化蛋白。以前，业已将体外筛选和进化用于提取新型催化 RNAs 和 DNAs，而在本发明中，将其用于提取新型蛋白酶。在该方法的一种具体实例中，可以通过筛选对催化剂转变状态的化学类似物的结合，间接提取催化剂。在另一种具体实例中，可以通过筛选与一种底物（例如，使用与一种亲和标记连接的底物）形成的共价键或通过裂解（例如，通过筛选裂解特殊键并因此从固体支持物上释放出一个文库中的催化性成分的能

力) 进行直接提取。

这种提取新型催化剂的方法与催化性抗体技术相比至少具有两个重要优点(参见 Schultz 等的综述, J. Chem. Engng. News 68: 26 (1990))。首先, 在催化抗体技术中起始库通常被局限于免疫球蛋白折叠; 相反, RNA-蛋白融合体的起始文库可以是完全随机的或者可以包括, 但不限于已知酶结构或蛋白支架的变体。另外, 催化抗体的提取通常取决于首先筛选对转变态反应类似物的结合, 随后是对活性抗体的费劲的筛选; 同样, 相反, 可以用 RNA-蛋白融合体文库方法直接筛选催化剂, 正如早先用 RNA 文库所证实的。在提取蛋白酶的另一种方法中, 可以组合使用转变-态-类似物和直接筛选方法。

通过该方法获得的酶具有很高的价值。例如, 目前就存在着需要新型、有效的工业催化剂的压力, 这种催化剂可以改进化学过程。本发明的一个主要优点是, 所述筛选可以在任意的条件下进行, 而且举例来说, 并不局限于体内条件。因此, 本发明有利于提取新型酶或现有酶的改进的变体, 这种酶能进行高度特异的转化(因此降低不需要的副产物的形成), 同时可以在预定的环境下起作用, 在高温、高压、或高溶质浓度环境下工作。

体外相互作用阱。 RNA-蛋白融合技术还可用于筛选 cDNA 文库和根据蛋白-蛋白相互作用克隆新的基因。通过该方法, 可以从需要的来源制备 cDNA 文库(例如, 通过 Ausubel 等的方法, 同上, 第 5 章)。对每一种候选 cDNAs 分子来说, 连接(例如, 使用上文所披露的用于制备 LP77、LP154 和 LP160 的技术)一个肽受体(例如, 嘧啶霉素尾)。然后按照本文所披露的方法制备 RNA-蛋白融合体, 并按照上述方法测定该融合体(或该融合体的改进形式)与特定分子的相互作用的能力。如果需要, 可以通过以下方法在上述过程中避免终止密码子和 3' UTR 区: (i) 添加抑制剂 tRNA, 以便能通读其终止区, (ii) 通过免疫沉淀除去翻译反应物中的释放因子, (iii) 组合(i)和(ii), 或(iv) 除去所述 DNA 序列上的终止密码子和 3' UTR。

由于相互作用步骤是在体外进行的, 这样可以利用非特异竞争剂、温度、和离子条件仔细控制反应严格性。具有非可水解类似物(例如 ATP 与 ATPgS)的正常小分子的改变, 提供了筛选相同分子不同构象子的区别特征。由于筛选的结合配偶体的 RNA 序列是共价连接的,

该方法可用于许多蛋白的克隆和功能性鉴定，并因此便于提取。另外，该技术可用于鉴定大约 50-100, 000 个人基因的功能和相互作用，所述基因的序列目前正通过人类基因组工程进行测定。

5

将 RNA-蛋白融合体用于小嵌片方案

“DNA 嵌片”包括固定化寡核苷酸或 cDNA 或基因组 DNA 的克隆片段的空间特定排列，并可具有诸如快速测序和转录仿型的用途。通过让 RNA-蛋白融合体的混合物（例如，由细胞 DNA 或 RNA 库制备）与这样的 DNA 嵌片退火，可以产生一种“蛋白展示嵌片”，其中，相应于一个固定化序列的每一个点能够与所述 RNA-蛋白融合体库中的其相应的 RNA 序列退火。通过该方法，相应的蛋白因为其与自身 mRNA 的连接被以空间限定的方式固定，而含有 DNA 序列组的嵌片展示相应的蛋白组。另外，如果所述融合体文库是由 cDNAs 或基因组 DNAs 的较小片段产生的，可以展示这些蛋白的肽片段。

15

这种级别的蛋白和肽的展示具有很多用途。例如，它是一种用于鉴定以前未知的蛋白-蛋白相互作用的有效工具。在一种具体方案中，对一种探针蛋白进行可检测的标记（例如，用荧光染料标记），并将标记过的蛋白与蛋白展示嵌片一起培养。通过该方法，能够确定能与探针蛋白结合的蛋白，所述测定是根据所述嵌片上由于结合了所述探针而被标记的斑点的位置而作出的。另一种用途是快速测定通过修饰酶的作用（例如，蛋白激酶，乙酰转移酶，和甲基转移酶）而被化学修饰过的蛋白。通过与感兴趣的酶和放射性标记过的底物一起培养所述蛋白展示嵌片，然后进行洗涤和放射自显影，可以方便地确定作为所述修饰过的酶的底物的蛋白的位置和性质。另外，该方法在有序展示小的肽的方面的应用，可以进一步定位所述修饰位点。

20

蛋白展示技术可以用固定在任何合适的固体支持物上的一系列核酸（包括 RNA，但优选 DNA）进行。典型的固体支持物可以由诸如玻璃（例如，玻璃板）、硅或硅玻璃（例如小嵌片）、或金（例如，金板）的材料制成。用于将核酸连接在上述固体表面的特定部位上的方法，例如，光能无机营养方法在本领域中是众所周知的，并可将其用于制备本发明使用的固体支持物（如 DNA 嵌片）。用于这一目的的典型方法包括，但不限于 Schena 等，科学 270: 467-470 (1995)；

Kozal 等, 自然医学 2: 753-759 (1996); Cheng 等, 核酸研究 24:
380-385 (1996); Lipshutz 等, 生物技术 19: 442-447 (1995);
Pease 等, Proc. Natl. Acad. Sci. USA 91: 5022-5026 (1994);
Fodor 等, 自然 364: 555-556 (1993); Pirrung 等, 美国专利
5 US5,143,845; 和 Fodor 等, WO92/10092.

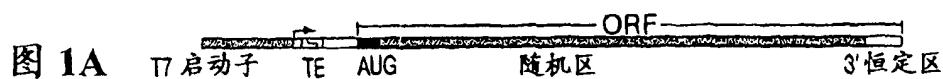


图 1B

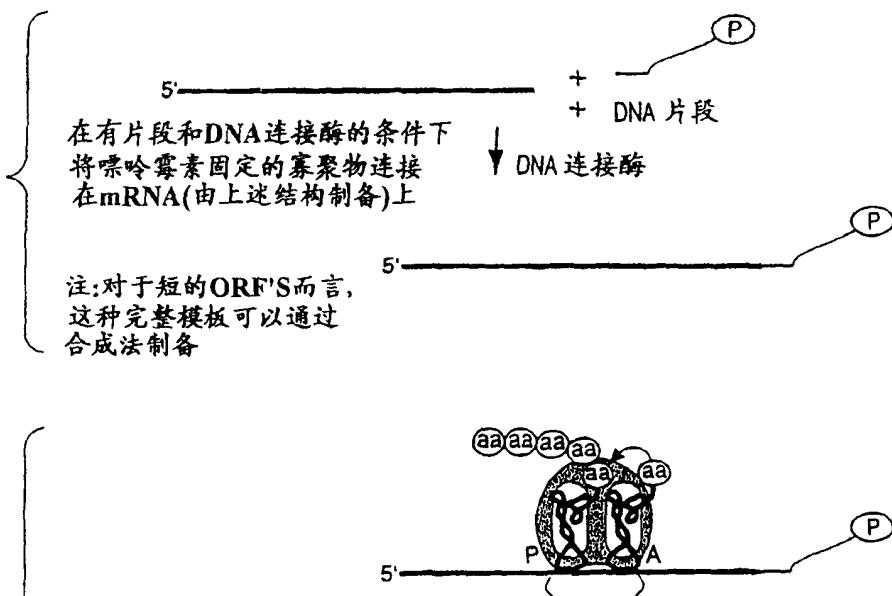
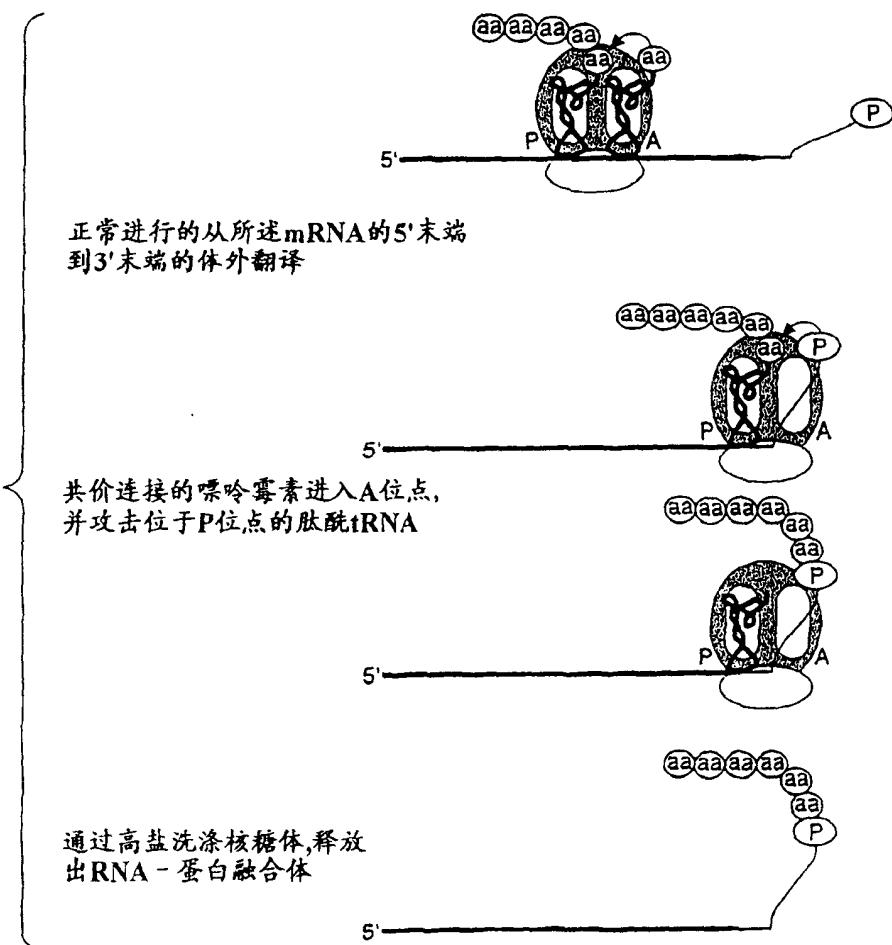


图 1C



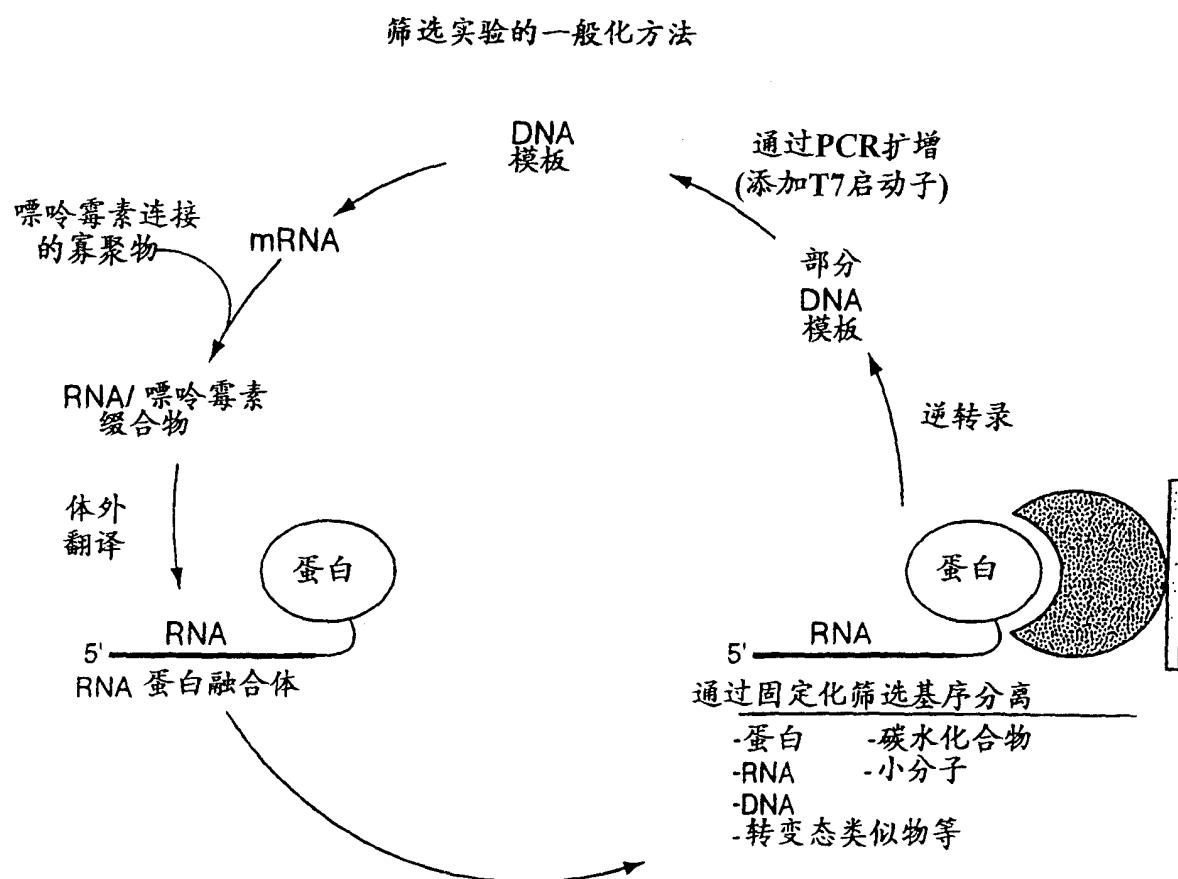


图 2

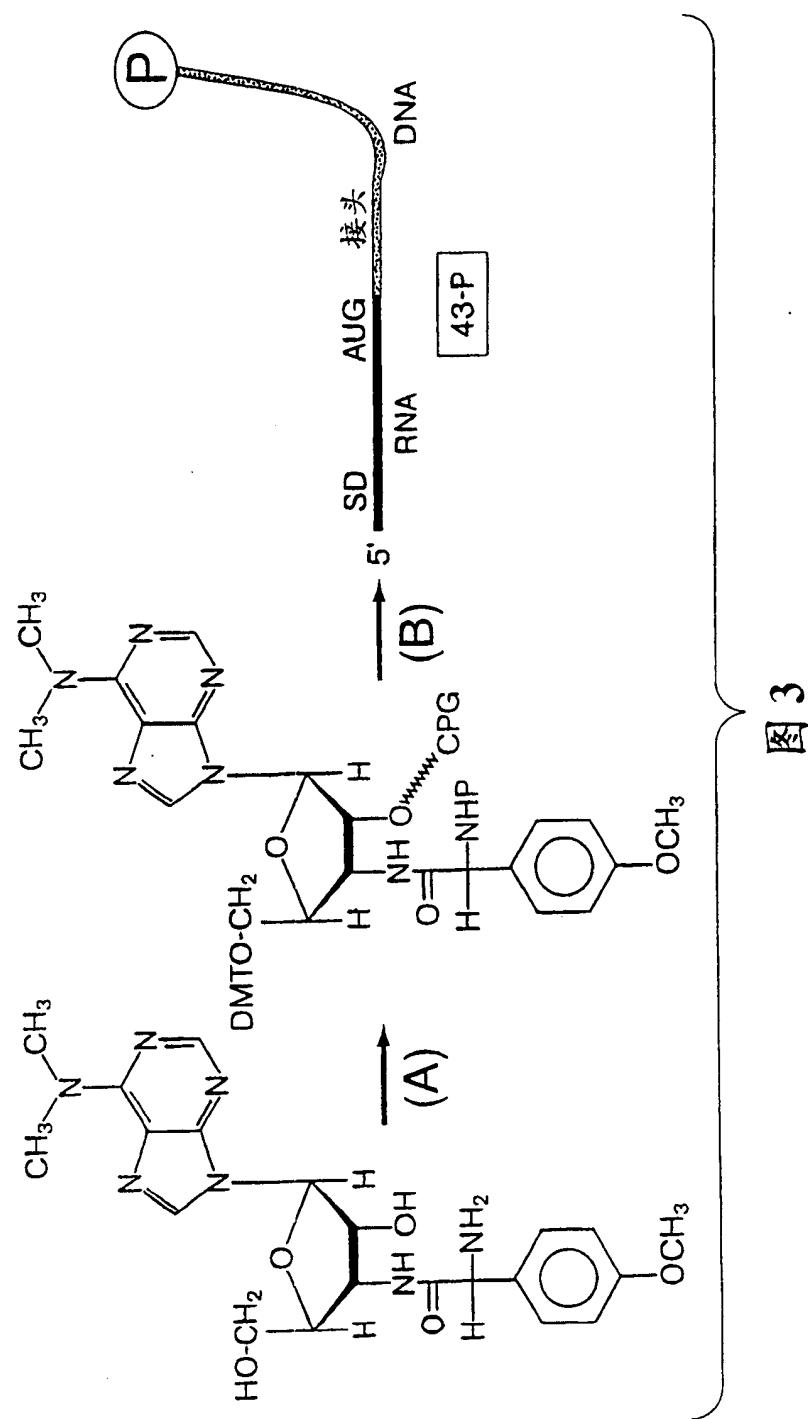
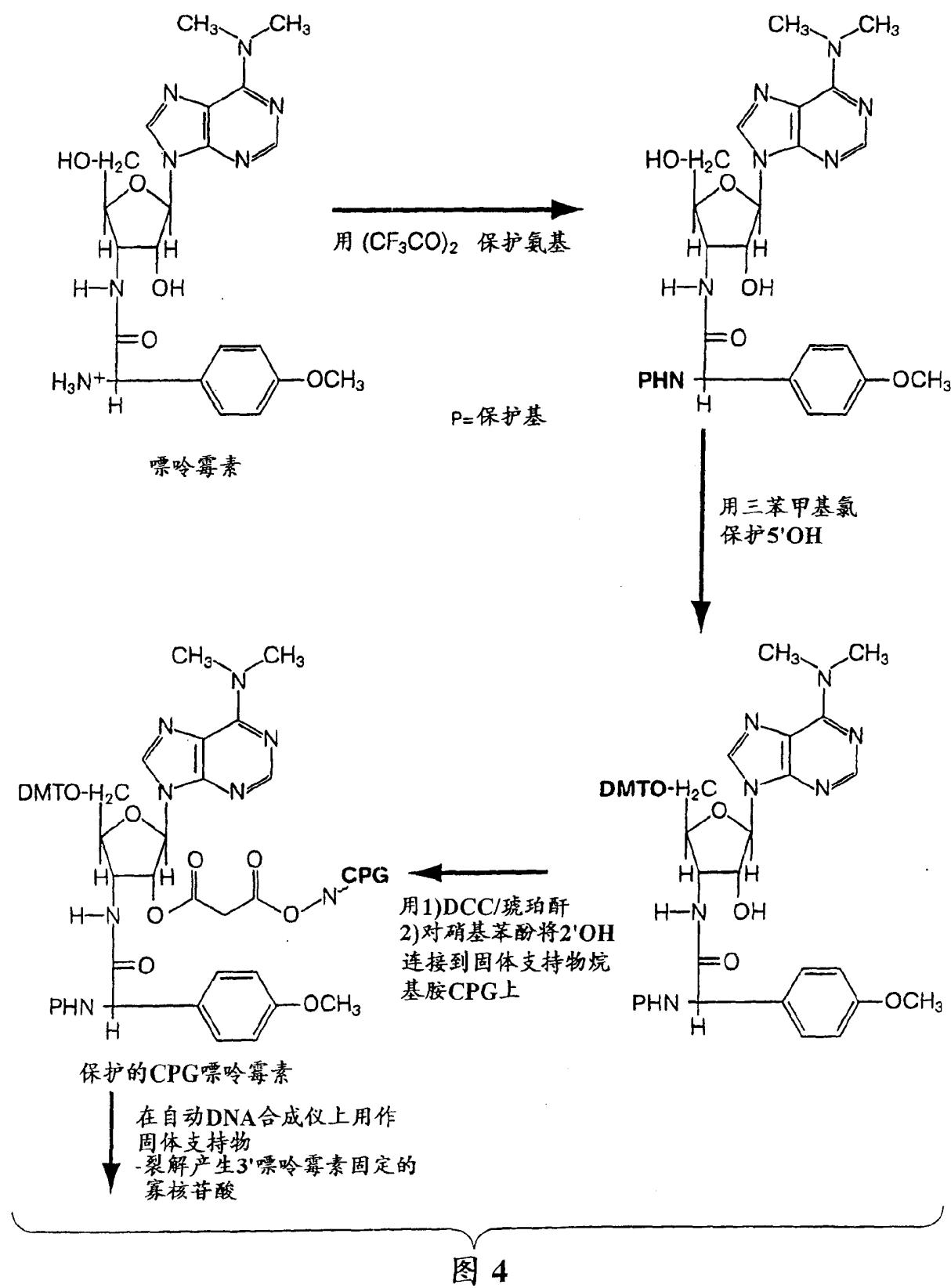


图 3



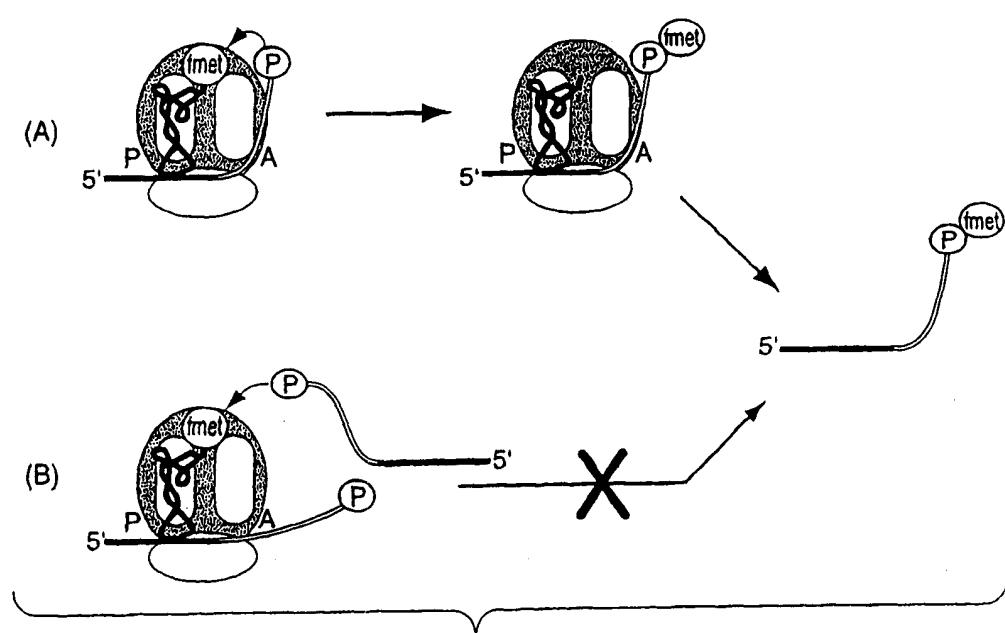


图 5

泳道 1 2 3 4 5 6
 Mg^{+2} [mM] 32P 6 9 12 15 18

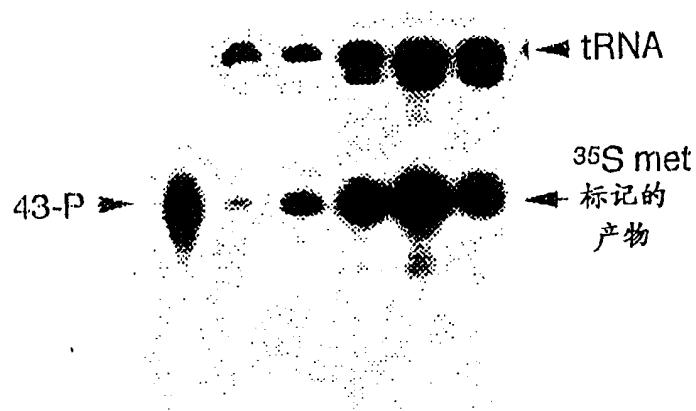


图 6A

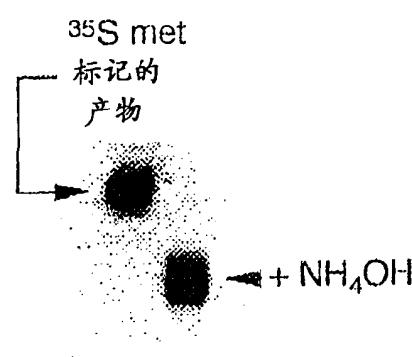


图 6B

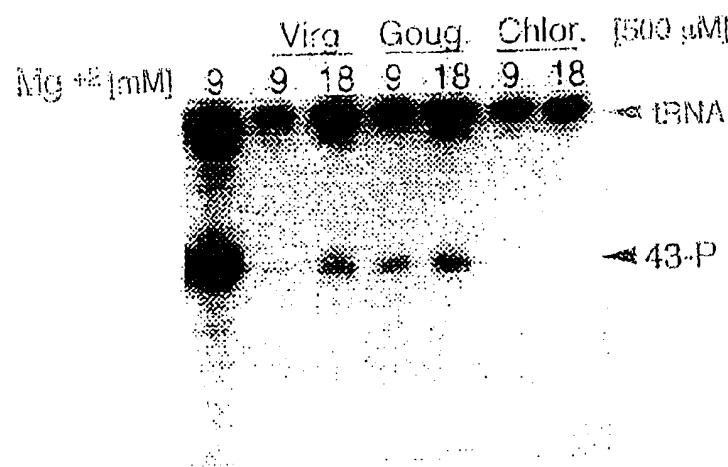


图 6C

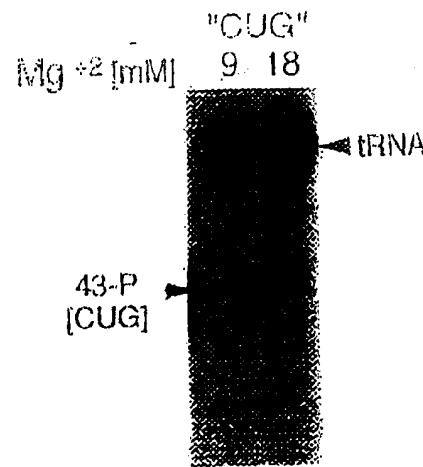


图 6D

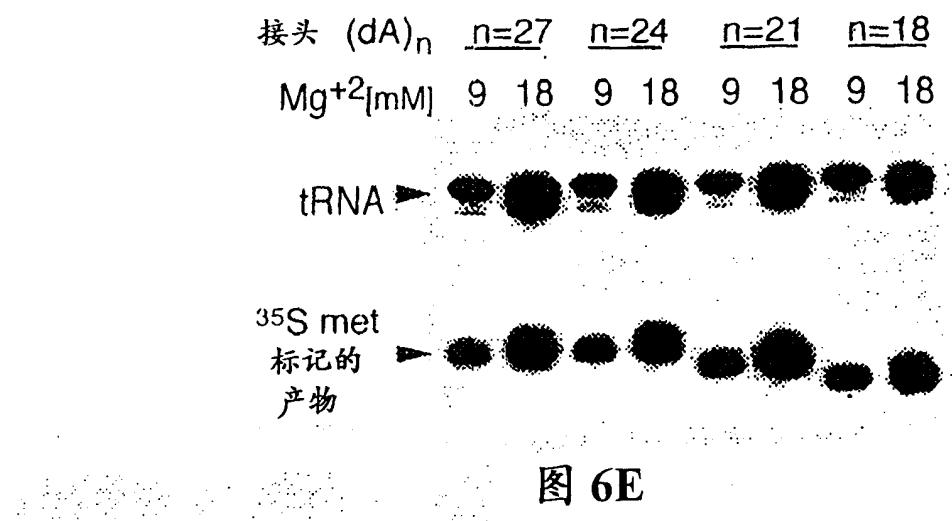
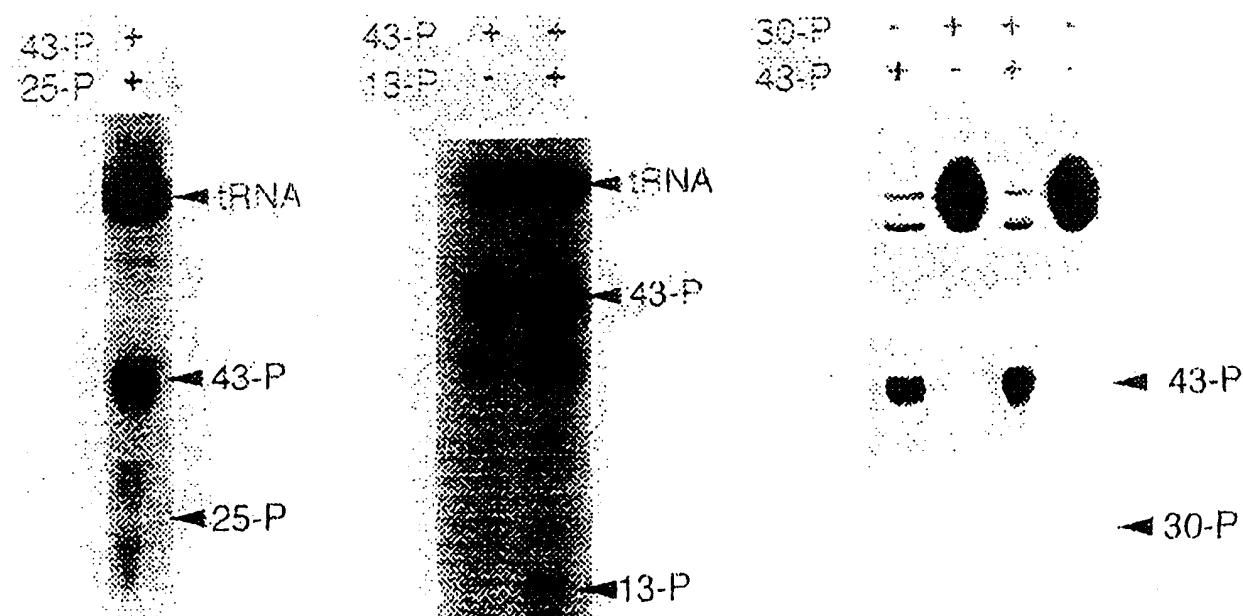


图 6E



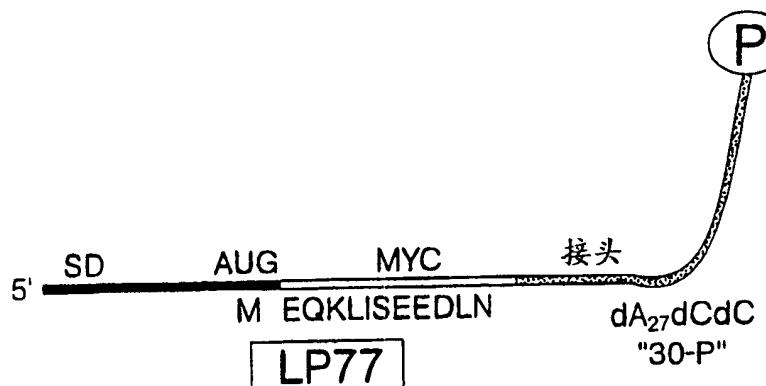


图 7A

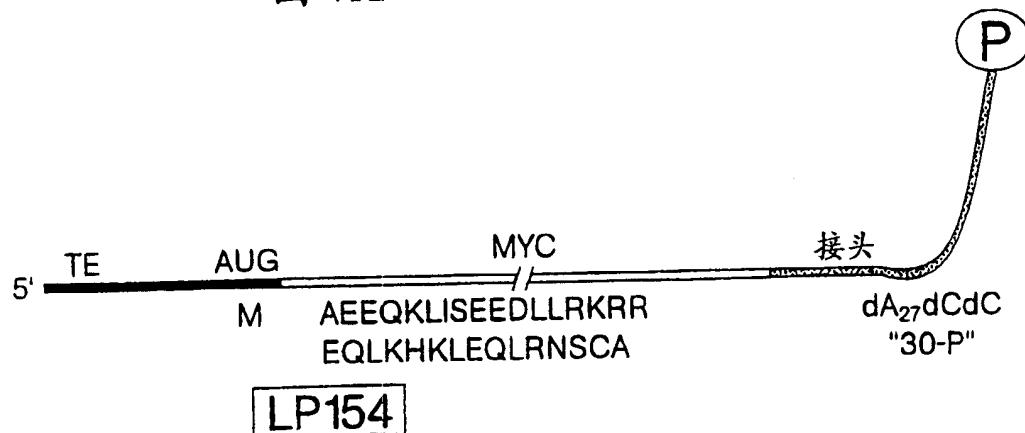


图 7B

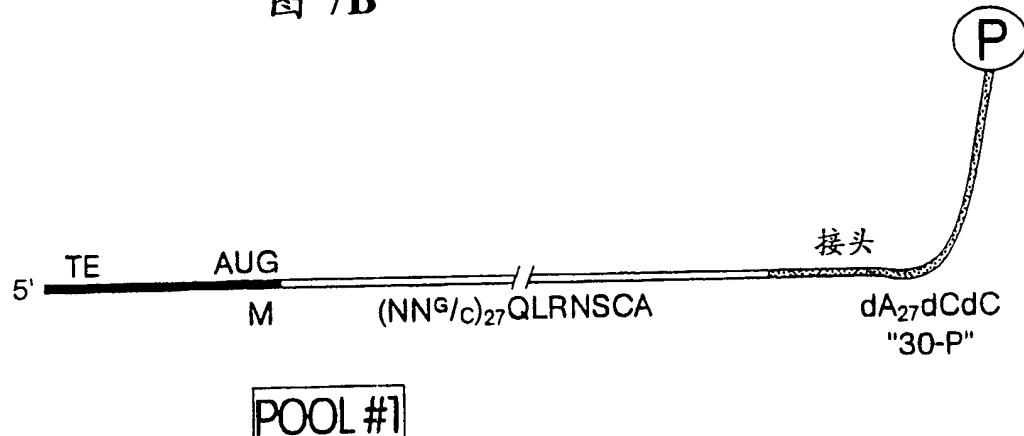


图 7C

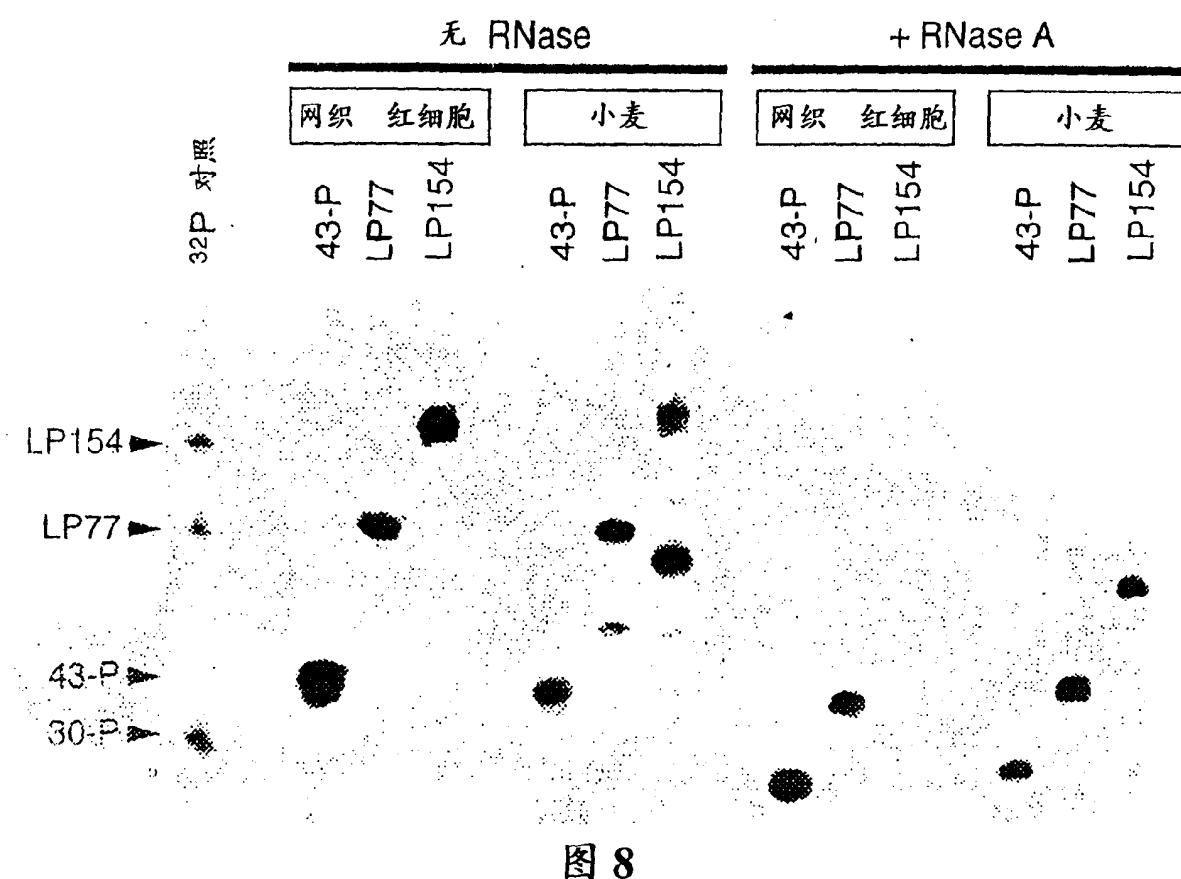


图 8

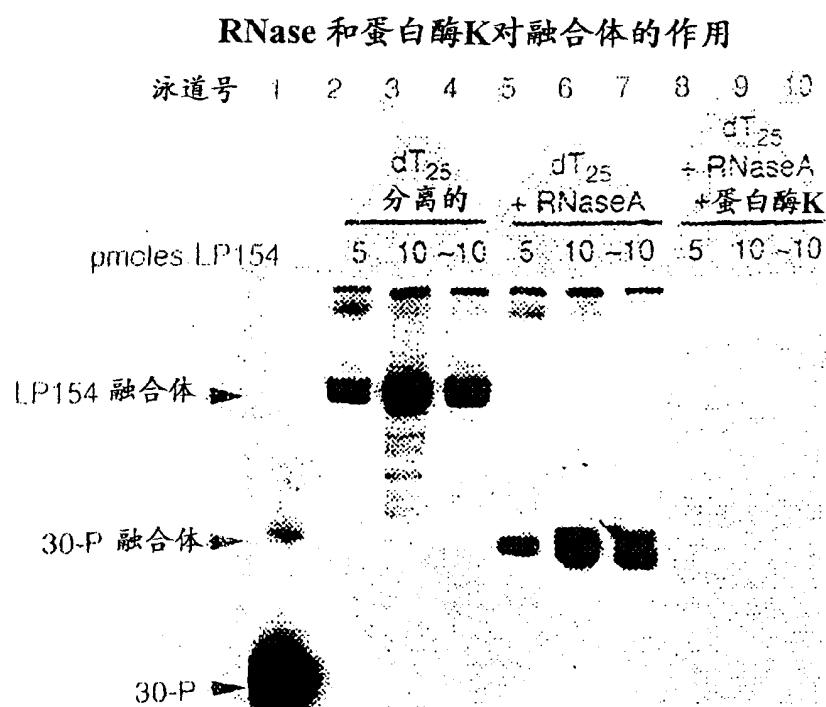


图 9

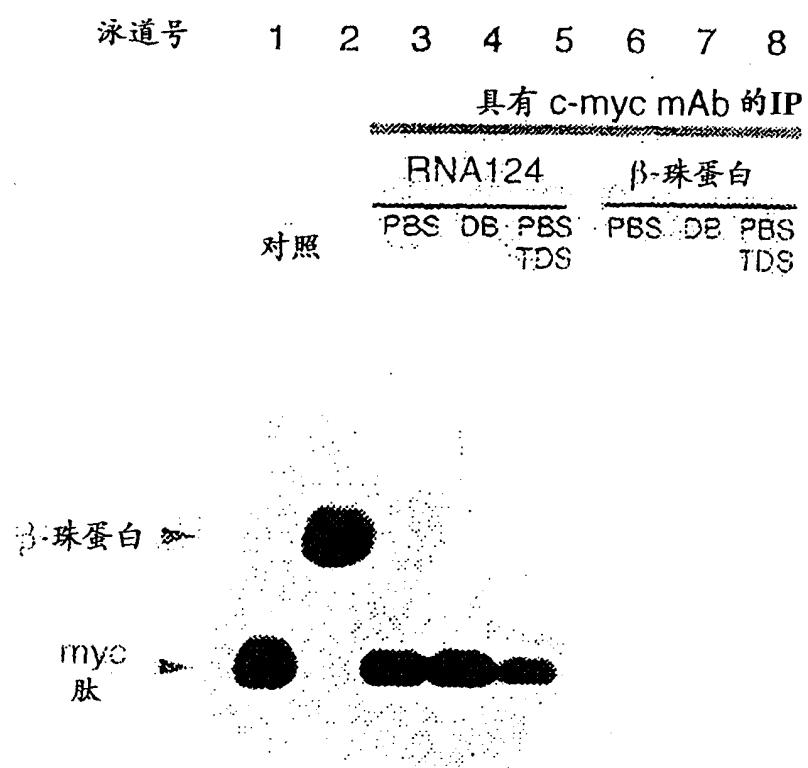


图 10

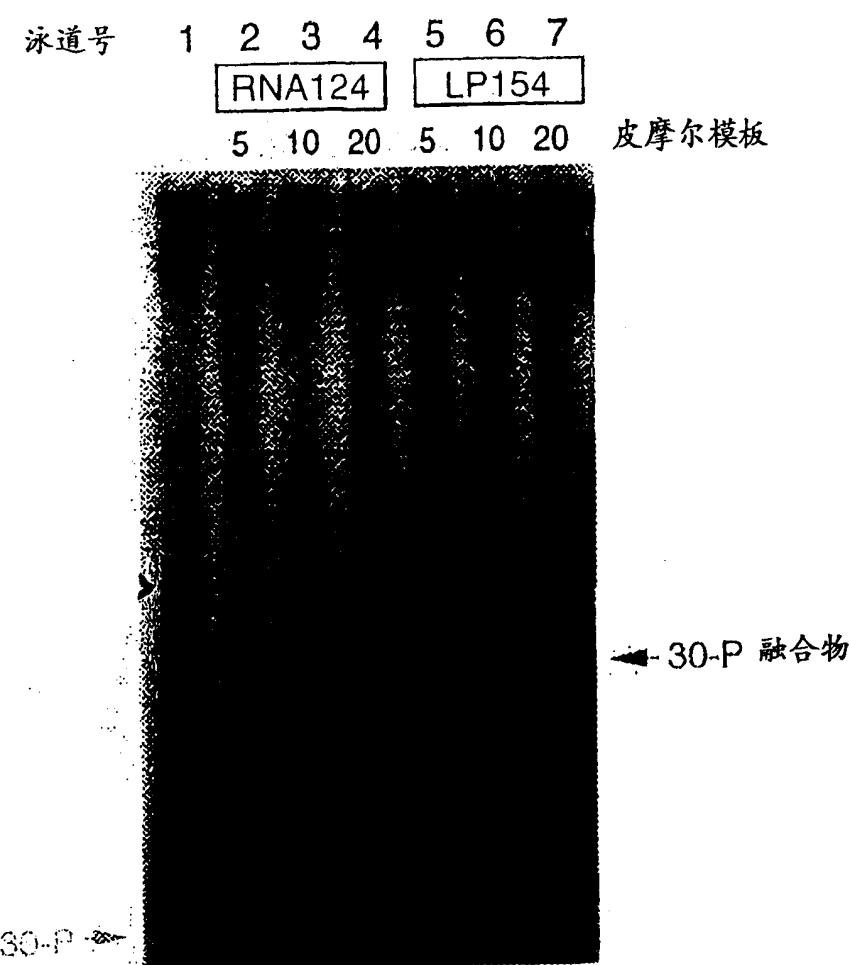


图 11

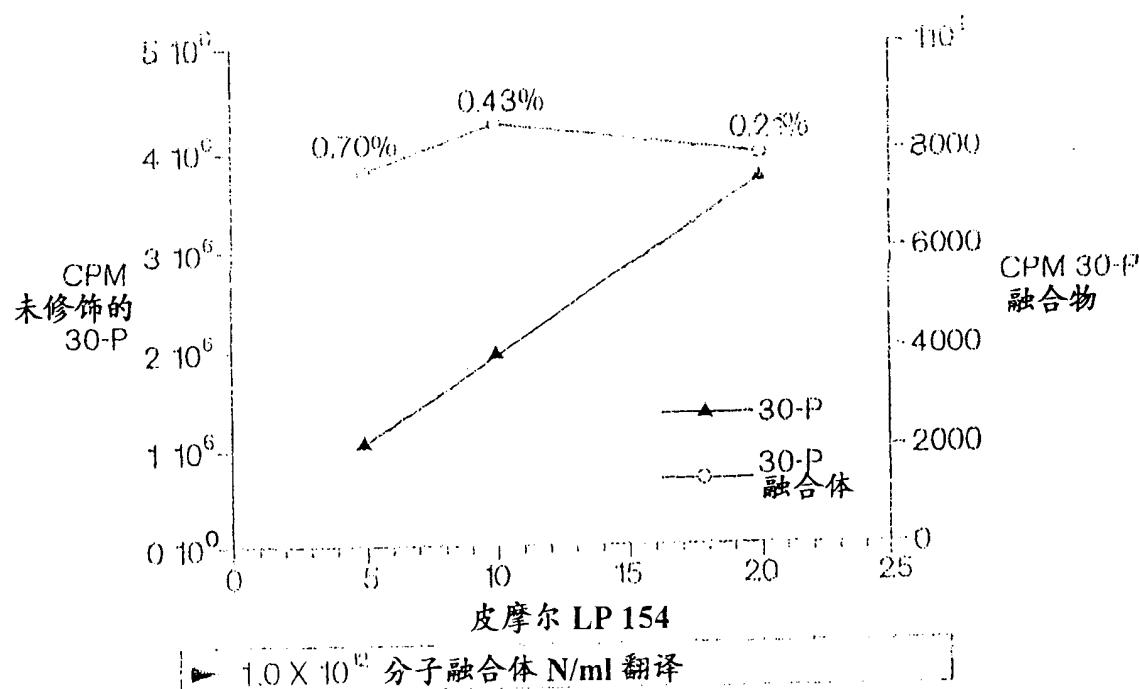


图 12

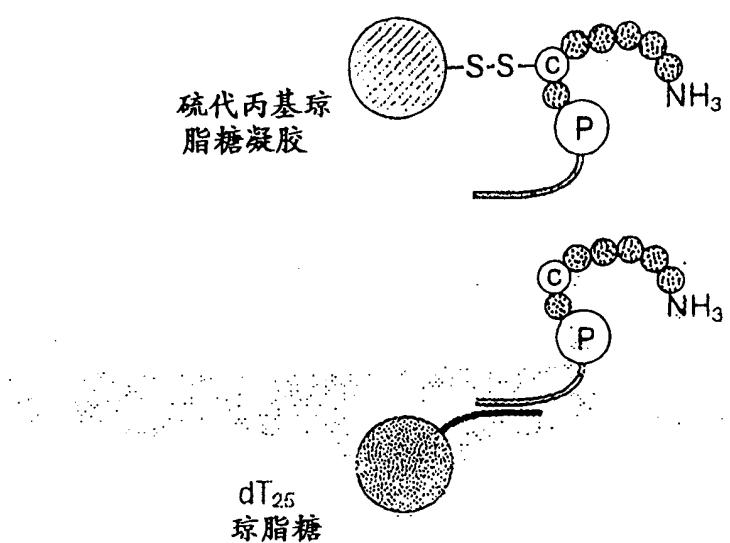


图 13

泳道号 1 2 3

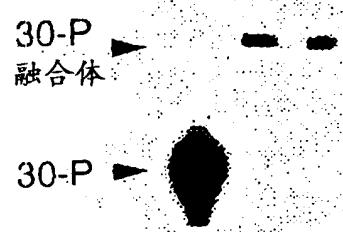


图 14

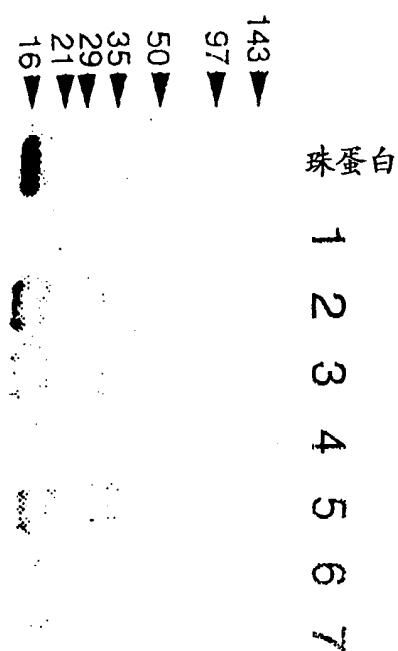


图 15A

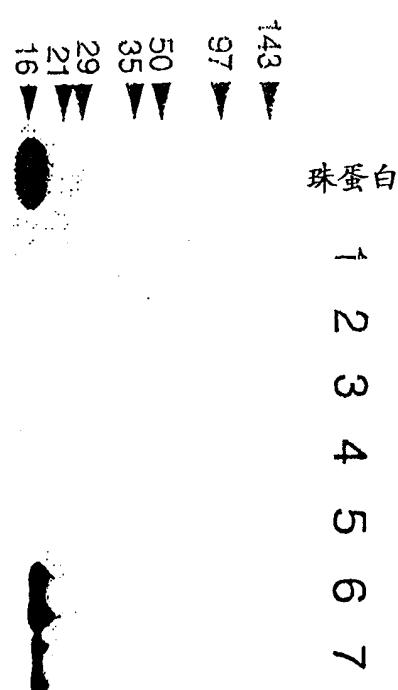


图 15B

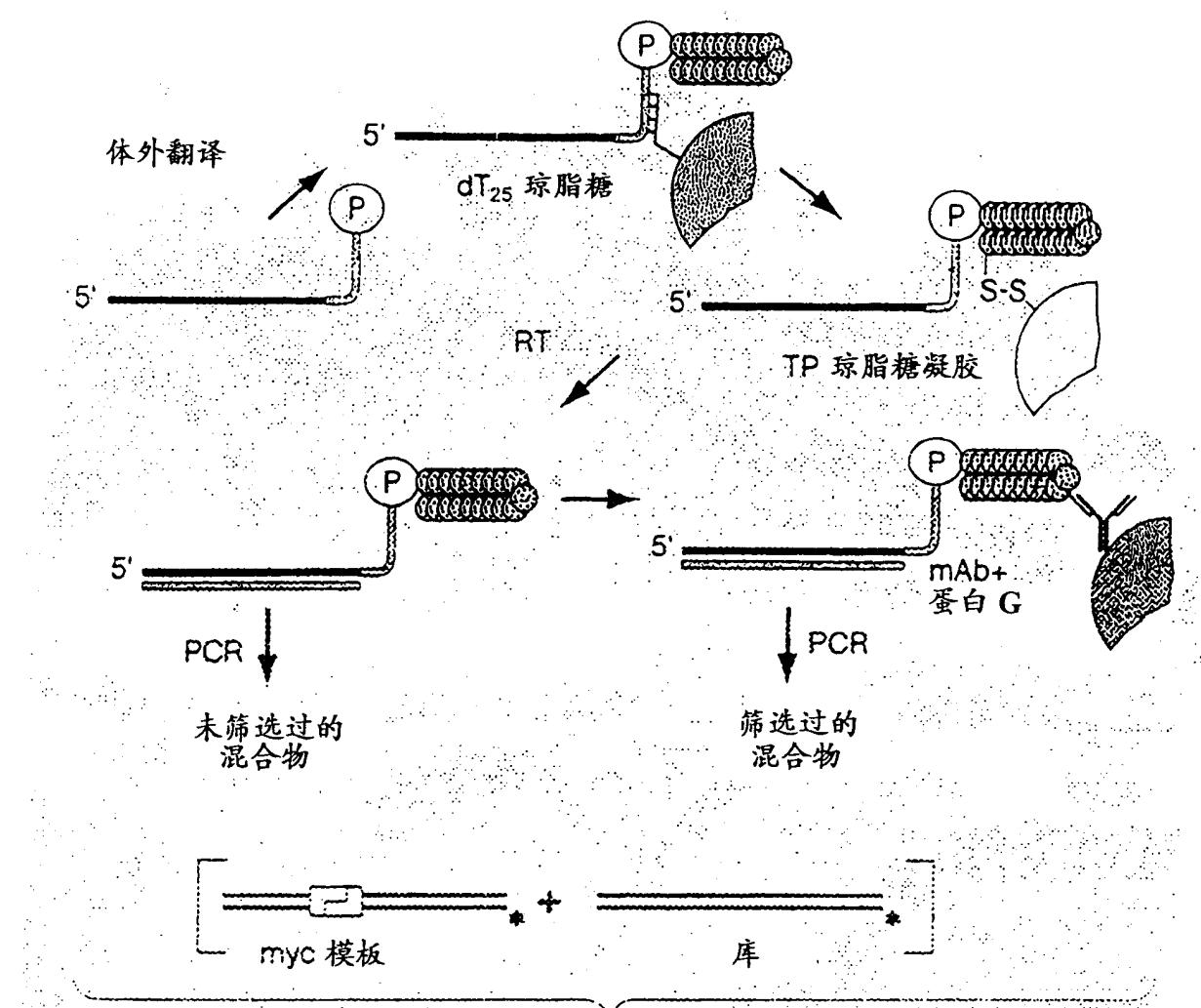


图 16A

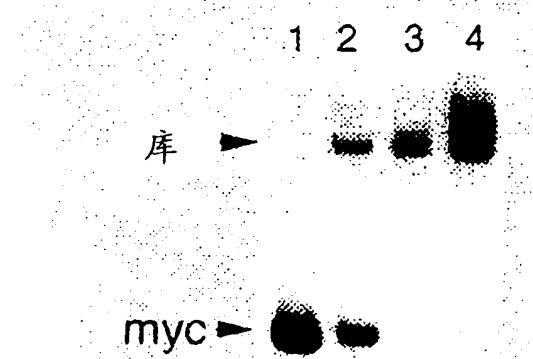


图 16B

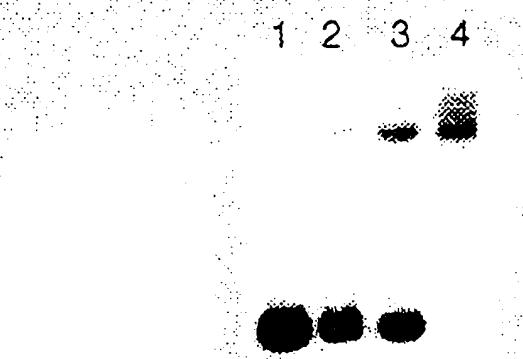


图 16C

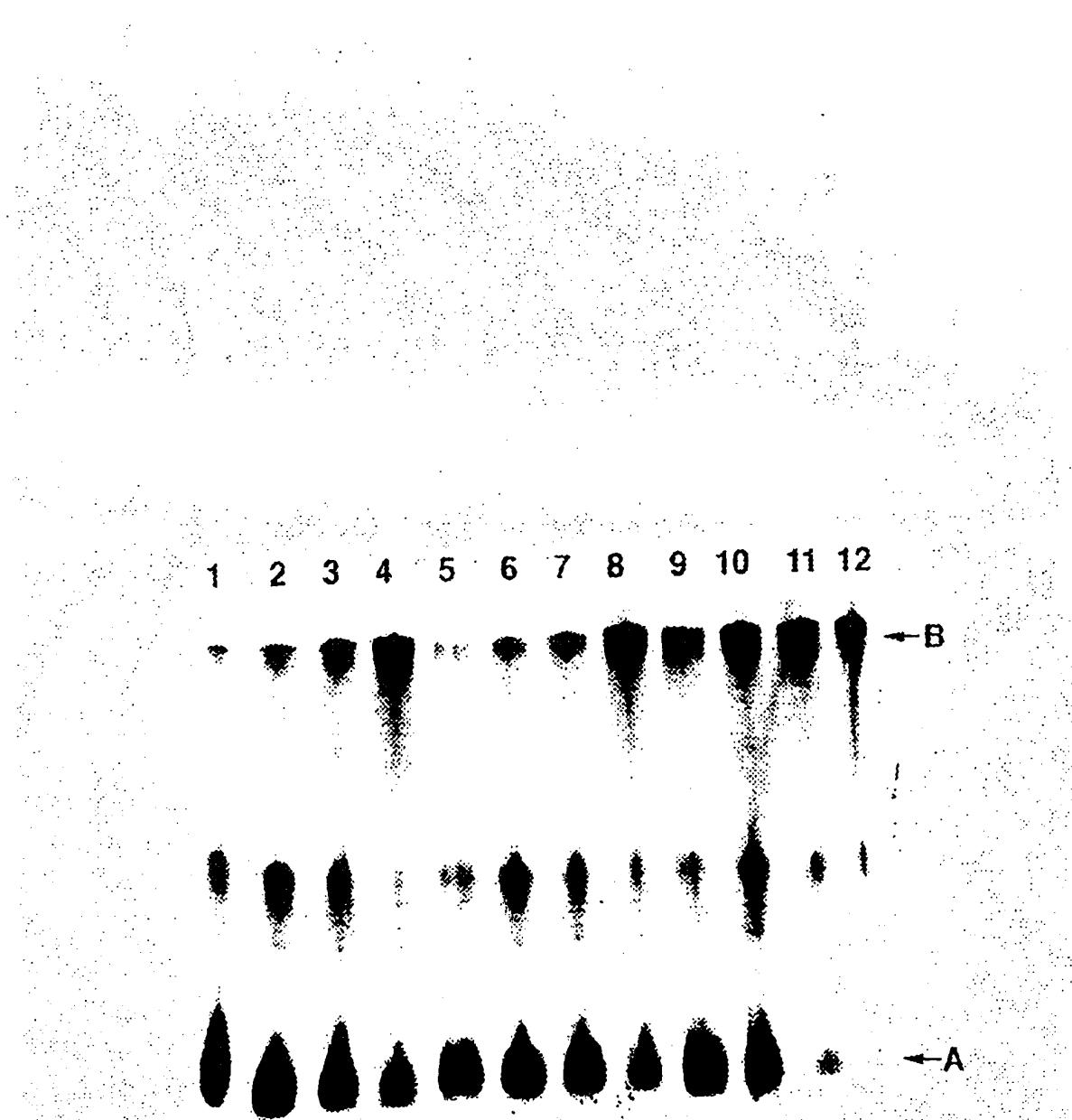


图 17

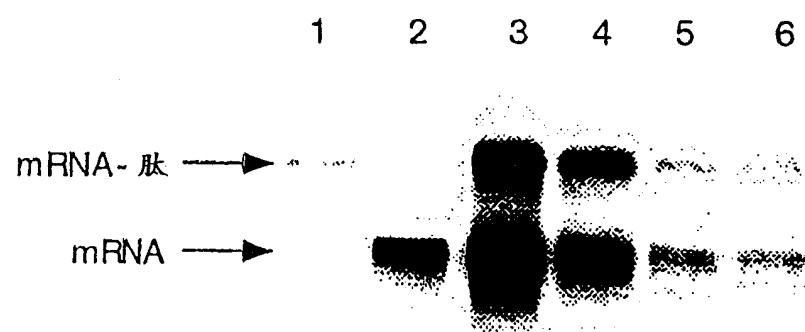


图 18

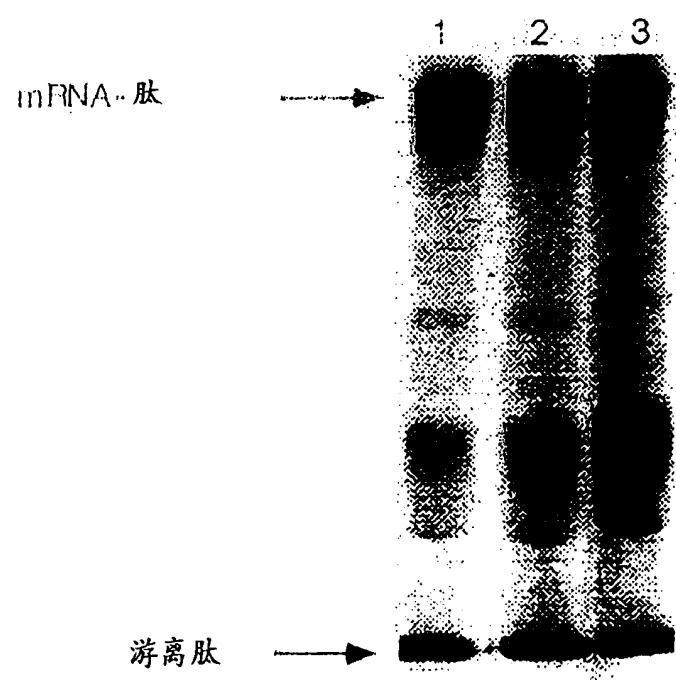


图 19