〔19〕中华人民共和国专利局

[11]公开号 CN 1065681A

〔12〕发明专利申请公开说明书

[21] 申请号 92100297.1

[51] Int.Cl[5]

C12N 15/52

〔43〕公开日 1992年10月28日

[22]申请日 92.1.17

[30]优先权

[32]91.1.17 [33]US [31]642,330

[71]申请人 综合医院有限公司

地址 美国马萨诸塞州

[72]发明人 J·哈塞尔福 A 布兰德

N·波瑞蒙 H·M·古德曼

[74]专利代理机构 中国国际贸易促进委员会专利代理部

代理人 黄革生

C12N 9/00 C12N 15/79

C12N 15/10 C07H 21/00

C12P 19/34 A01N 57/16

A01N 63/00

说明书页数: 51 附图页数: 23

[54]发明名称 利用转拼核酶摘除细胞

[57]摘要

本发明介绍了根据I组内含子的催化活性而设计的能够进行自我催化的转拼的新核酶．利用这一设计，有可能构建出能够以非常精确的方式将新的3′外显子序列有效地拼接至任何选定的靶RNA序列中去的核酶．还介绍了利用本发明的新转拼核酶摘除细胞的方法，该方法能够在体内以定向、调节方式为宿主细胞提供毒性产物．还介绍了无活性的核酶原．

权　利　要　求　书

1. 一种多核苷酸分子，该分子编码转拼核酶，该核酶的序列为融合 *RNA*，该融合 *RNA* 的序列包括：

(1) 第一 *RNA* 序列，该第一 *RNA* 序列足以使该核酶与靶 *RNA* 杂交，

(2) 第二 *RNA* 序列，由于该核酶转拼活性的结果，该第二 *RNA* 序列能够被共线性地转染至靶 *RNA* 中；其中该多核苷酸分子的表达与转录激活子蛋白的表达进行了可操纵地连接，且其中所说的第一 *RNA* 序列是与编码所说转录激活子蛋白的 *RNA* 杂交的序列。

2. 根据权利要求 1 所述的多核苷酸分子，其中所说的转录激活子是 *GAL4*。

3. 根据权利要求 1 所述的多核苷酸分子，其中所说的第二 *RNA* 序列包括编码对宿主细胞有毒的肽的序列。

4. 根据权利要求 3 所述的多核苷酸分子，其中所说的肽是 *DTA* 肽。

5. 根据权利要求 4 所述的多核苷酸分子，其中所说的 *DTA* 肽是突变肽序列。

6. 根据权利要求 5 所述的多核苷酸分子，其中所说的突变肽

序列包括由 *SEQ ID No.* 40 编码的氨基酸。

7. 根据权利要求 5 所述的多核苷酸分子，其中所说的突变肽序列包括由 *SEQ ID No.* 41 编码的氨基酸。

8. 根据权利要求 1 所述的多核苷酸分子，其中所说的第一 *RNA* 序列是与 *GAL4* 杂交的序列，且其中所说的第二 *RNA* 序列是编码 *DTA* 肽的序列。

9. 根据权利要求 1—8 中任一项所述的多核苷酸分子，其中所说的分子是 *RNA*。

10. 根据权利要求 1—8 中任一项所述的多核苷酸分子，其中所说的分子是 *DNA*。

11. 一种包含核酶表达盒的多核苷酸分子，该盒能稳定地插入至宿主基因组中，且该盒包括能够在宿主中发挥功能且与权利要求 1—9 任一项的多核苷酸的编码序列可操纵地连接的启动子序列。

12. 一种包含权利要求 11 多核苷酸分子的宿主细胞。

13. 根据权利要求 12 所述的宿主细胞，其中所说的宿主细胞是病毒细胞。

14. 根据权利要求 12 所述的宿主细胞，其中所说的宿主细胞是原核植物细胞。

15. 根据权利要求 12 所述的宿主细胞，其中所说的宿主细胞是真核生物细胞。

16. 根据权利要求 15 所述的宿主细胞，其中所说的真核生物细胞是植物细胞。

17. 根据权利要求 15 所述的宿主细胞，其中所说的真核生物细胞是动物细胞。

18. 根据权利要求 17 所述的宿主细胞，其中所说的动物是果蝇。

19. 根据权利要求 17 所述的宿主细胞，其中所说的动物是哺乳动物。

20. 根据权利要求 19 所述的宿主细胞，其中所说的动物是人。

21. 一种用于体外转拼的方法，该方法包括以下步骤：

(1)向转拼反应混合物中提供权利要求 9 的多核苷酸分子，该多核苷酸分子包括能够与第二多核苷酸杂交的序列；

(2)向该反应混合物中提供所说的第二多核苷酸；

(3)在所说的条件下催化所说的第二多核苷酸的转拼。

22. 一种用于体内转拼的方法，该方法包括以下步骤：

(1)为宿主细胞提供权利要求 9 的多核苷酸；

(2)在所说的宿主细胞中表达由所说分子编码的所说的核酶；

(3)在所说的宿主细胞中表达所说核酶的底物；

(4)在所说的宿主细胞中用该底物催化该核酶的转拼。

23. 一种灭活靶 *RNA* 活性的方法，该方法包括：

(1)向转拼反应混合物中提供权利要求 9 的多核苷酸，所说的核酶对靶 *RNA* 具有催化活性，该催化活性导致该靶 *RNA* 功能的失活；

(2)向该混合物中提供所说的靶 *RNA*；

(3)提供条件，使得该多核苷酸表达该催化活性。

24. 一种在体内为宿主细胞提供所需遗传序列的方法，该方法包括：

(1)为该宿主细胞提供权利要求 9 的多核苷酸，该多核苷酸序列在该宿主细胞中对靶 *RNA* 具有催化活性，该核酶能够转拼该所需的遗传序列；

(2)为该宿主细胞提供所说的靶 *RNA*；

(3)提供条件，使得该核酶转拼所需的遗传序列至该靶 *RNA* 的序列中。

25. 一种在多细胞动物或植物中摘除细胞的方法，该方法包括给受精胚宿主细胞提供权利要求 9 的多核苷酸，该核酶编码对该宿主细胞有毒的肽且该核酶能够将该序列转拼至该宿主细胞的靶中。

26. 一种对植物雄性或雌性不育进行操纵的方法，该方法包括为该种植物的胚细胞提供权利要求 9 的多核苷酸，该核酶作用于一 *RNA*，该 *RNA* 所表达的蛋白质是该植物受精所必需的，并导致表达该蛋白的细胞被摘除。

27. 一种对植物进行抗植物病原菌的免疫的方法，该方法包括用权利要求 9 的多核苷酸转化植物细胞，其中所说的多核苷酸所编码的转拼序列能够给该植物提供抗该病原菌的免疫力，且其中被该病原菌感染的植物细胞导致该细胞的摘除。

28. 根据权利要求 1—7 中任一项所说的多核苷酸分子，其中所说的核酶是核酶原。

29. 根据权利要求 8 的多核苷酸分子，其中所说的核酶是核酶原。

30. 根据权利要求 9 的多核苷酸分子，其中所说的核酶是核酶原。

31. 根据权利要求 10 的多核苷酸分子，其中所说的核酶是核酶原。

32. 根据权利要求 11 的多核苷酸分子，其中所说的核酶是核酶原。

33. 根据权利要求 12 的宿主细胞，其中所说的核酶是核酶原。

34. 根据权利要求 21 的方法，其中所说的核酶是核酶原。

35. 根据权利要求 22 的方法，其中所说的核酶是核酶原。

36. 根据权利要求 23 的方法，其中所说的核酶是核酶原。

37. 根据权利要求 24 的方法，其中所说的核酶是核酶原。

38. 根据权利要求 25 的方法，其中所说的核酶是核酶原。

39. 根据权利要求 26 的方法，其中所说的核酶是核酶原。

40. 根据权利要求 27 的方法，其中所说的核酶是核酶原。

说　　明　　书

利用转拼核酶摘除细胞

本发明涉及新的转拼核酶以及利用这些核酶摘除细胞的方法。

Ⅰ．Ⅰ组内含子

具有催化活性的 *RNA* 分子称为核酶或 *RNA* 酶(*Cech*, *T. R.*, *Ann. Rev. Biochem.* 59:543－568(1990))。嗜热四膜虫(*Tetrahymena thermophila*)前体 *rRNA* 含有能够催化其自身切割的内含子(核酶)。该核酶是一组结构上相关的Ⅰ组内含子中的一种。

修饰过的嗜热四膜虫内含子的拼接活性需要有鸟苷辅因子和二价阳离子(Mg^{++}或Mn^{++})存在，并且是通过两个连续的转酯反应(图 1)发生的。首先，游离的鸟苷与核酶结合且其 3′羟基被置于攻击 5′拼接位点的磷原子的位置。鸟苷共价连接于内含子序列并释放出 5′外显子。接着，新释放出的 5′外显子的 3′羟基与位于 3′拼接位点的磷酸二酯键反应，从而产生连接的外显子序列。然后，切除的内含子通过一系列的转酯反应(涉及其 3′羟基和内部序列)形成缩短了的环形分子。

从化学上讲，这些连续反应是相似的并且都发生在单一活性位点。自拼反应的特征在于交替 *RNA* 结构的形成，因为不同的 *RNA* 链均在高度保守的内含子周围形成了类似的构象。拼接需要

对准位于称作"内部指导序列"(或 *IGS*)的互补序列内含于一外显子接点。

在 5′拼接位点进行的首次切割需要在 *IGS* 和与拼接位点相邻的序列之间形成碱基配对的螺旋(*P*1)。该螺旋中 *U*：*G*"变位"碱基对的存在可限定将在核酶的催化反应中破坏磷酸二酯键。该键断裂后，部分 *P*1 螺旋被取代，且由于 *IGS* 和与 3′拼接位点相邻的序列之间的互补性而形成了新螺旋(*P*10)。不变的鸟苷残基位于 3′拼接位点处的磷酸二酯之前，如同 *P*1 序列中被取代的部分一样。因此，外显子的连接仅在新的外显子序列取代了内含子序列之处以与第一次切割反应相反的方向发生。应当指出，紧接外显子连接反应的内含子环化反应同样也涉及 5′序列与 *IGS* 的碱基配对，而且反应是由内含子末端鸟苷残基的 3′羟基介导的(*Been, M. D. et al.*, "*Selection of Circularization Sites In A Group I IVS RNA Requires Multiple Alignments of An Internal Template—Like Sequence*," *Cell* 50:951(1987))。

Ⅰ. 催化活性

为了更好地确定Ⅰ组内含子的结构和催化性质，已从嗜热四膜虫内含子"核心"中剔去外显子序列。*Cech, T. R*, 等(*WO* 88/04300)曾指出，嗜热四膜虫内含子核酶至少具有三种催化活性：(1)脱磷酸活性，它能够以序列特异性方式除去 *RNA*3′末端的磷酸，(2)*RNA* 聚合酶活性(核苷酸基转移酶)，它能够催化寡核糖核苷酸

转化为多核糖核苷酸，(3)序列特异性核糖核酸内切酶活性。

分离的核酶活性能够与底物 *RAN* 进行相互转移作用，且已对这类反应进行过鉴定。例如，当将平末端形式的内含子和对应于 5′ 拼接点的序列一起保温时，该位点受到类似于拼接第一步的鸟苷依赖性切割。底物与核糖核酸内切的内含子 *RNA* 进行碱基配对，形成螺旋 *P*1，且切割发生在第 4—6 位的 *U*：*G* 碱基对之后。通过系统发育比较和突变分析发现，只要维持了 *P*1 螺旋的碱基配对，紧接 5′ 拼接位点处保守尿嘧啶残基的序列的性质对催化作用并不重要（*Doudna, J. A. et al.*, *Proc. Natl. Acad. Sci. USA* 86:7402—7406(1989)）。

选择 3′ 拼接位点所需的序列似乎主要位于内含子本身的结构内，它包括螺旋 *P*9 0 以及随后的确定 3′ 内含子边界的鸟苷残基。但是，如同突变分析所显示的一样，3′ 外显子内的侧翼序列是形成螺旋 *P*10 以及进行有效拼接所必需的（*Suh, E. R. et al.*, *Mol. Cell. Biol.* 10 :2960—2965 (1990)）。另外，利用平末端形式的内含子和“外部”指导序列以及已被 5′ 鸟苷残基延长的寡核苷酸，已将寡核苷酸转移连接。在连接之前，通过在外部模板上形成类似于 *P*10 的螺旋而单独对准对应于 3′ 外显子序列的底物寡核苷酸（*Doudna, J. A. et al.*, *Nature* 339:519—522(1989)）。

通过将分散的“杂交”区操纵至核酶而将核酶的切割活性对准于特定的 *RNA*，该杂交区能够与所需的 *RNA* 特异杂交。例如，*Ger-*

lach， *W*。 *L* 等(*EP*321 201)构建出一种含有与靶*RNA* 互补的序列的核酶。通过增加该互被序列的长度可提高该序列对靶 *RNA* 的亲和性。但是，该核酶的杂交区和切割区彼此整合。通过互补区与靶 *RNA* 杂交后，核糖酶的催化区切割该靶。它提示核酶将可在体内用于灭活或切割靶 *RNA*，从而可用于例如治疗由于外源宿主 *RNA* 的产生所引起的人体疾病。但是，由核酶指导的转拼(与转切相反)则未曾提到或揭示过。

已对各种天然存在的核酶的核糖核酸内切酶活性(切割活性)进行过广泛研究。对这些核酶的结构和序列进行分析后发现，切割位点周围的某些核苷酸非常保守，而其侧翼序列则并非如此。这一情况导致人们设计出自然界尚未发现的新的核糖核酸内切酶活性。例如，*Cech* 等人已构建出改变了底物序列特异性的新核酶(*Cech*, *T. R. et al.*, *WO* 88/04300; *Koizumi*, *M. et al.*, *FEBS Lett.* 239:295—288 (1988); *Haseloff*, *J. et al.*, *Nature* 334:585—591 (1987);及 *Heus*, *H. A. et al.*, *Nucl Acids Res.* 18:1103—1108 (1990))。根据自身切割植物类病毒和卫星 *RNA* 方面的早期研究(*Buzayan*, *J. M. et al.*, *Proc. Natl. Acad. Sci. USA* 83:8859—8862(1986))，已为能够以高度序列特异性方式转移切割其它 *RNA* 分子的核酶的设计建立了指南(*Haseloff*, *J. et al.*, *Nature* 334: 585—591 (1988))。但是，这些构建物并不能够有效地催化针对转拼的反应。

分离 *RNA* 上含有的外显子的连接(即转拼)在自然界发现于 *snRNP* 介导的以及自我催化的 I 组和 II 组内含子中。在锥虫及 *Caenorhabditis elegans mRNA* 中,共同的 5′前导序列是由各自的基因上转录并被拼接到 *mRNA* 的 3′部分(*Agabian, N., Cell* 61:1157—11600 (1990); *Hirsh, D. et al., Mol. Biol. Rep.* 14: 115 (1990))。这些小的"拼接前导"*RNA*(*slRNA*)由已融合至在哺乳动物 *snRNA* 拼接提取物中能够从功能上取代 *U1 snRNA* 的序列的 5′外显子构成。

另外,I 组和 II 组自拼内含子均能够在人工系统中转移连接外显子(*Been, M. D. et al., Cell* 47:207—216 (1986); *Galloway—salvo, J. L. et al., J. Mol. Biol.* 211: 537 — 549 (1990); *Jacquier, A. et al., Science* 234:1099—1194 (1986);和 *Jarrell, K. A. et al., Mol. Cell Biol.* 8:2361—2366 (1988))。I 组内含子的体内转拼发生在叶绿体的拼接基因中(*Kohchi, T. et al., Nucl. Acids Res.* 16:10025—10036 (1988)),且 I 组内含子的体内转拼在大肠杆菌的人工拼接基因中得到显示(*Galloway—Salvo, J. L. et al., J. Mol. Biol.* 211 :537—549 (1990))。在后一情况下,含有 I 组内含子的噬菌体 T_4 胸苷酸合成酶基因(*td*)在连接内含子螺旋 *p6a* 的环处被断开。已证实 *td* 基因片段的转录物经过体外转拼拼营救发生机能障碍的大肠杆菌宿主细胞。已知的碱基配对(*p3*、*p6* 和 *p6a*)和内含子片段之间可能的三级反应使得基因片段能够

正确装配和加工。

在体外，四膜虫核酶能够催化单链寡核糖核苷酸底物的转拼。其中有四个成分是必需的：核酶、3′单链 *RNA*、5′外显子和 *GTP*。在该反应中使用了缩短形式的四膜虫核酶(*L*－21 *Sca* I *IVS RNA*（起始于内部指导序列并终止于 U_{409}）(*Flanegan, J. B. et al., J. Cell. Biochem.* (*Supp.*)12 *part D*: 28 (1988))。由 *GTP* 在 5′拼接位点所进行的攻击释放 5′外显子，接着，该 5′外显子在 3′拼接位点进行的转酯反应中被核酶连接至 3′外显子。

已有人提出在体内用核酶代替反义 *RNA* 对准和破坏特定的 *RNA*(*Gerlach, W. L. et al., EP* 2,21,201; *Cotten, M., Trends Biotechnol.* 8:174--178 (1990); *Cotten, M. et al., EMBO J.* 8: 3861－3866 (1989); *Sarver, N. et al., Science* 247:1222－1225 (1990))。例如，已有人提出，对于 *HIV*－1 感染期间所表达出的 *RNA* 具有核酸内切催化活性的核酶的表达可作为一种抵抗 1 型人免疫缺陷病毒(*HIV*－1)感染的可能治疗方法(*Sarver, N. et al., Science* 247:1222--1225 (1990); *Cooper, M., CDC AIDS Weekly, April* 3, 1989, *page*2; *Rossi, J. J., Abstract of Grant No. IROIAI*29329 *in Dialog's Federal Research in Progress File* 265)。但是，这一想法并未获得成功。

在为调查核酶作为治疗 1 型人体免疫缺陷病毒(*HIV*－1)感染的治疗剂的替在应用而进行的研究中，*hammerhead motif* 核酶

(*Hutchins, C. J. et al., Nucl. Acids Res.* 14:3627 (1986); *Keese, P. et al., in Viroids and Viroid—like Pathogens, J. S. Semancik, ed., CRC Press, Boca Raton, FL*, 1987, *PP*1—47) 所作用的是 *HIV*—1 *gag* 转录本。作用于 *gag* 的核酶在人细胞培养物中的表达导致 *HIV*—1 *gag RNA* 及抗原 *p24* 水平的下降(但不是完全丧失)(*Sarver, N. et al., Science* 247:1222—1225 (1990))。因此,*Sarver* 氏核酶的医疗效果受其低效所限,因为任何逃脱的致病 *RNA* 都会给宿主带来问题。

在体内应用核酶的另一个问题是,为了在核提取物中发挥核酶的抑制功能需要有高比值的核酶/底物,而这一比值难以达到。*Cotton* 等通过显微注射法获得了高比值的核酶/底物,它是将含有与强 *tRNA* 启动子(聚合酶Ⅲ启动子)可操纵连接的核酶产生基因的表达盒与含有核酶切割序列的底物 *RNA* 一起注射进蛙卵母细胞中(*Cotton, M. et al., EMBO J.* 8:3861—3866 (1089))。但是,对于多细胞生物来说显微注射不是一种合适的给药方法。

抗编码大肠杆菌 β—半乳糖苷酶的 *mRNA* 的核酶体内活性已有报道(*Chuat, J. —C. et al., Biochem. Biophys. Res. Commun.* 162 :1025—1029 (1989))。但是,这一活性只有当核酶和靶被转染至细菌细胞内的同一分子上时才能观察到。当作用于自具有 β—半乳糖苷酶基因的靶部分的细菌 *F* 游离基因转录的 *mRNA* 时,核酶的活性无效。

因此，目前对核酶活性的技术应用仅限于利用核酶的切割活性以破坏靶 *RNA* 的活性。不幸的是，为了保证其效力这类应用常常需要完全破坏所有的靶 *RNA* 分子和/或需要相对高比值的核酶/底物，而这一高比值难得达到。更重要的是，现有技术中修饰过的核酶不能够有效、定向地转拼。

因此需要开发高度有效的核酶和核酶表达系统。尤其是，现有技术中没有提到通过将新 *RNA* 序列转拼至宿主 *RNA* 来破坏已有的 *RNA* 序列或改变已有 *RNA* 的编码序列的有效方法。

鉴于设计新核酶的可能性以及对在体内改变高等真核生物遗传特征的有效方法的需要的认识，本发明人对利用核酶在体内改变天然 *RNA* 遗传信息进行过研究。这些努力的结果是开发出高度有效的转拼核酶并为其操作建立了指南。

按照本发明，首先提供了一种 *RNA* 或 *DNA* 分子，该分子编码转拼核酶，而该核酶能够在体外或体内以高度精确的方式将新的 3′外显子序列有效地拼接至任何选定的靶 *RNA* 序列中，并且这种 *RNA* 或 *DNA* 分子在调节任何选定的靶 *RNA* 或 3′外显子序列的能力以及在加进了增强该核酶特异性的至短序列方面是新的。

按照本发明，也提供了一种编码核酶的 *RNA* 或 *DNA* 分子，而为该核酶编码的序列是一融合 *RNA*，该融合 *RNA* 产生出可以使核酶与靶 *RNA* 杂交的第一 *RNA* 序列以及能够转位至靶 *RNA* 中去的第二 *RNA* 序列．且该第二 *RNA* 序列编码出一种异源于靶 *RNA*

序列的 *RNA* 序列。

按照本发明，也提供了一种编码核酶的 *RNA* 或 *DNA* 分子，而编码该核酶的序列为上面所述的融合 *RNA*，由融合 *RNA* 产生的第一 *RNA* 序列是一种使得该 *RNA* 分子与 *GAL4* *RNA* 杂交的序列，且融合 *RNA* 的第二 *RNA* 序列提供白喉毒素 *A* 链(*DTA*)的编码序列。

按照本发明，也提供了一种 *RNA* 或 *DNA* 分子，它编码的是构象被破坏的本发明核酶(核酶原)，该核酶原由底物激活，即，该核酶原在与靶 *RNA* 发生特异性反应而被重新激活之前只有微不足道的或没有自切或转拼活性。

按照本发明，也提供了含有核酶或核酶原表达盒的 *RNA* 或 *DNA* 分子。该表达盒能够稳定地插入至宿主基因组中，该表达盒产生能够在宿主中起作用的启动子序列，而该序列与本发明的核酶或核酶原序列可操纵地连接。

按照本发明，也提供了含有核酶或核酶原表达盒的 *RNA* 或 *DNA* 分子。该表达盒能够稳定地插入至宿主基因组中，该核酶表达盒产生应答 *GAL4* 的启动子序列，而该序列与本发明的核酶或核酶原序列可操纵地连接。

按照本发明，也提供了一种体外转拼的方法，该方法包括以下步骤：(1)在体外提供本发明的核酶或核酶原以及适合于该核酶的底物，(2)提供促进所需的该核酶或核酶原催化活性的体外反应条

件，(3)使该核酶或核酶原与该底物在该条件下反应。

按照本发明，也提供了一种体内转拼的方法，该方法包括以下步骤：(1)将本发明的*RNA*或*DNA*分子提供给宿主细胞，(2)在该宿主细胞中表达由该分子所编码的核酶或核酶原，(3)在该宿主细胞中表达该核酶或核酶原底物，(4)将该核酶或核酶原与该底物在该宿主细胞中反应。

按照本发明，也提供了一种灭活靶*RNA*活性的方法，该方法包括(1)提供本发明的核酶或核酶原，该核酶或核酶原对靶*RNA*具有催化活性，(2)提供该靶*RNA*，和(3)提供使该核酶或核酶原表达其作用于靶*RNA*的催化活性的条件。

按照本发明，也提供了一种在体内为宿主细胞提供所需遗传序列的方法，该方法包括(1)为所需宿主细胞提供本发明的核酶或酶原，该核酶或核酶原在该宿主细胞中对靶*RNA*具有催化活性，(2)提供编码该所需遗传序列的该核酶或核酶原，(3)提供条件使得该核酶或核酶原将该所需的遗传序列转拼至靶*RNA*序列中。

按照本发明，也提供了一种摘除多细胞动物或植物细胞的方法，该方法包括将本发明的核酶或核酶原提供给任何宿主细胞(尤其是受精的胚胎宿主细胞)，该核酶或核酶原编码对该宿主细胞有毒的基因序列，且该核酶或核酶原能够在该宿主细胞中与所需的靶转拼。

按照本发明，也提供了一种对农业上重要的植物种进行雄性或

雌性不育处理的方法，该方法包括利用本发明的核酶或核酶原摘除受精所需的任何细胞。

按照本发明，也提供了一种对植物进行抗病原菌免疫的方法，该方法包括构建能够表达本发明的植物病原菌特异性融合核酶或核酶原的转基因植株，该核酶或核酶原能够摘除被该病原体感染的任何宿主细胞。

按照本发明，也提供了一种转化的、抗病原菌微生物，该微生物被本发明的核酶或核酶原所转化，而该核酶或核酶原对该病原体所表达的核酸分子具有催化活性。

按照本发明，还提供了一种能够将所需的核酶或核酶原活性传递给所需宿主的病毒病原体，该核酶或核酶原活性由本发明的核酶或核酶原传递。

图 1 是 I 组内含子的核酶拼接机制示意图。

图 2 是（*A*）嗜热四膜虫 *rRNA* 内含子；(*B*)本发明靶 *mRNA* 和转拼核酶或核酶原的结构示意图。

图 3(*A*)是 *CAT—LacZ* α—肽转拼核酶的设计图；(*B*)是 *CAT—LacZ* 核酶的完整 *DNA* 编码序列。

图 4 表示黄瓜花叶病毒(*CMV*)*RNA*4 转拼核酶的序列。*A*：病毒 *RNA* 靶序列；*B*：寡核苷酸靶序列；*C*：*CMV RNA*4—白喉毒素 *A* 链转拼核酶。

图 5 是黄瓜花叶病毒 3/4 序列的比较。

图 6(*A*)是 *Gal*4—白喉毒素 *A*(*DTA*)转拼核酶的设计图;(*B*)是具有异亮氨酸取代基的 *Gal*4—*DTA* 核酶的完整编码序列。

图 7 是用于表达 *Gal*4 蛋白的 *p*—成分介导的“增强子—诱捕”法示意图。

图 8 表示了野生型 *DTA* 和 *DTA*3′外显子突变体的部分序列。

图 9 是 *pGaTB* 和 *pGaTN* 的酶切图谱。

图 10 是 *pUAST* 的酶切图谱。

图 11 是表达(*al*4—*DTA* 转拼核酶的果蝇胚表皮制品。

图 12 表示了“核酶原”设计的基本原理。箭头显示核酶切割位点,“反义”区用黑色表示 辐射状阴影表示催化区,浅阴影表示 3′“外显子”序列。在没有靶 *mRNA* 的情况下,转拼核酶可进行瞬时碱基配对,并与异源序列(包括其自身)反应。另外,在“3′外显子”连接区将发生剪切。所构建的无活性“核酶原”含有额外的自身互补序列,该序列引起核酶的催化中心误叠。只有当与所 需 的靶 *mRNA* 进行碱基配对并且干扰二级结构接着被取代后,才会形成具有活性的核酶。

图 13 显示了 *CAT*—*LacZ* 转拼核酶的序列及其估计的二级结构。核酶“核心”序列被涂成阴影(按照 *Cech*,*Gene* 73 ;259—271 (1988))。其中显示了未修饰核酶和核酶原 1 和 2 所作用的螺旋 *p*8,它们分别具有 13 和 18 个核苷酸的互补于“反义”区(已标明)的序列。

图 14(1)显示了活化的 *CAT—LacZ* 核酶，反义，具有螺旋 *p*8 和 3′“外显子”序列的核酶区；(2)(*a*)显示了在修饰过的螺旋 *p*8 的序列与“反义”区之间进行了碱基配对的无活性 *CAT—LacZ* 核酶原 2；(*b*)与 *CAT mRNA* 碱基配对以取代螺旋 *p*1—“反义”配对并重新形成螺旋 *P*1 后具有活性的核酶原。

图 15 显示了 *CAT—LacZ* 核酶原转录本的稳定性。用 *EcoRI* 切割含有 *CAT—LacZ* 核酶和核酶原序列的质粒，并利用 *T7* 或 *SP6 RNA* 聚合酶及〔^{32}p〕*UTP* 进行转录。在 *7M* 尿素和 25%的甲酰胺中通过 5%聚丙烯酰胺凝胶电泳对放射性标记的转录本进行分级分离。核酶转录本受到深度水解(主要在“3′外显子”连接处)。而核酶原则明显地很少反应。

图 16 显示了 *CAT—LacZ* 核酶原的核糖核酸内切酶活性。用 *ScaI* 切割含有 *CAT—LacZ* 核酶和核酶原序列的质粒，并用 *T7* 或 *SP6 RNA* 聚合酶进行转录。在 *40mM Tris—HCl*(*PH*7.5)、*6mM* $MgCl_2$、*2mM* 亚精胺、*10mM NaCl*、*2mM GTP* 中将转录本与放射性标记的 *CAT RNA*(由 *T7 RNA* 聚合酶自 *Puv* I 切割的质粒转录)一起分别于 37℃、45℃和 50℃保温 30 分钟。然后在 *7M* 尿素和 25%甲酰胺中通过 5%聚丙烯酰胺凝胶电泳对产物进行分级分离，并放射自显影。由 *RNA* 介导的 173 核苷酸 *CAT RNA* 的切割产生分别为 76 和 97 个核苷酸的 5′和 3′片段。

图 17 显示了用于核酶原设计的“野生型”和修饰过的螺旋

*p*8，它具有图示的可能的碱基配对。与相应核酶原的"反义"部分互补的碱基以黑体形式表示。在每一螺旋边列出了互补碱基的数目。螺旋由相应的核酶转录本的稳定性控制，如同通过体外转录过程中"3′外显子"水解程度所测定的一样。

图18显示了*GAL4—DTA*核酶原的稳定性。用*XhoI*使含有核酶和核酶原序列的质粒开环，并利用*T7 RNA*聚合酶进行转录。将转录本于50℃在40*mM Tris—HCl*(*PH*7.5)、6*mM* $MgCl_2$、2*mM*亚精胺、10*mM NaCl*、1*mM GTP*中保温60分钟，然后在7*M*尿素及25%甲酰胺中通过5%聚丙烯酰胺凝胶电泳进行分级分离，并进行放射自显影。在上述条件下核酶转录本大量水解，而核酶原1水解较少，核酶原2则很稳定。

1. 定义

在以下的叙述中，使用了重组*DNA*(*rDNA*)技术领域所广泛利用的各种术语。为了对本发明的说明书和权利要求书(包括这些术语的给定范围)有一个清楚和一致的理解，下面对其进行定义。

核酶：遗传上具有催化活性的*RNA*分子。

转拼：一种遗传操作形式，由此，第一多核苷酸的核酸序列被共同线性连接至或插入至第二核苷酸的序列中，并保留了这些多聚核苷酸之间的3′→5′磷酸二酯键。"定向"转拼或"底物特异性"转拼指需要特定*RNA*(即专用于拼接转换序列的*RNA*)作为转拼反应底物的转拼反应。如果设计出针对于相应套*RNA*上出现的靶序列

的 核 酶或核酶原 ，则定向转拼可作用于一个以上的 *RNA*。

靶 *RNA*：作为本发明核酶或核酶原催化活性底物的 *RNA* 分子。

表达盒：为本发明核酶或核酶原的表达提供所需序列的遗传序列。

稳定：将序列"稳定地"插入基因组是指以导致该序列在该基因组的拷贝中遗传的方式进行插入。

可操纵的连接：指将一序列与另一序列相连以改变该序列功能的发挥的连接。例如，通过将核酶或核酶原编码序列可操纵地连接于启动子，就可将核酶或核酶原编码序列的表达置于该启动子的控制之下。如果启动子功能的诱导导致核酶或核酶原编码序列的转录且如果这两个序列之间键合的性质没有（1）导致移码突变的引入(2)影响指导核酶表达的表达调节序列的能力，则可将这两个核酸序列(如核酶或核酶原编码序列及在编码序列 5′端的启动子区序列)称作被可操纵的连接。因此，如果启动子能够引起该核酸序列的合成，则启动子区是与核酸序列可操纵地连接的。

Ⅰ．本发明核酶的操纵

本发明的转拼核酶、核酶原和方法首次提供了能够定向转拼至任何 *RNA* 序列 ，尤其是成熟（不含内含子）的 *mRNA* 中的核酶。这里所述的与靶的互补性得到延长的转拼核酶与现有技术中所述的得自嗜热四膜虫的核糖核酸内切酶活性极不相同。本发明核酶互补

性的增加提高了它对靶的亲和性和特异性，且该互补性并不是催化活性不可缺少的部分。另外，在无变性剂和含有高浓度 Mg^{++} 的情况下切割有效而精确地发生。

这里所述的设计转拼核酶的准则是保守的，它基于 I 组自拼内含子的突出性质，其目的是为任何定向转拼核酶的设计提供一个总方案。因此，这里所提出的准则并不限于嗜热四膜虫前 *mRNA* 的 I 组内含子，本领域技术人员可根据这些准则和已有知识设计出具有其它的 I 组内含子的本发明核酶。

在下面的嗜热四膜虫染色体外 *rDNA* 序列中，天然嗜热四膜虫核酶(内含子序列)位于 53—465 碱基之间。

TGACGCAATT CAACCAAGCG CGGGTAAACG GCGGGAGTAA CTATGACTCT

CTAAATAGCA ATATTTACCT TTGGAGGGAA AAGTTATCAG GCATGCACCT

CCTAGCTAGT CTTTAAACCA ATAGATTGCA TCGGTTTAAA AGGCAAGACC

GTCAAATTGC GGGAAAGGGG TCAACAGCCG TTCAGTACCA AGTCTCAGGG

GAAACTTTGA CATGGCCTTG CAAAGGGTAT GGTAATAAGC TGACGGACAT

GGTCCTAACC ACGCAGCCAA GTCCTAAGTC AACAGATCTT CTGTTGATAT

GGATGCAGTT CACAGACTAA ATGTCGGTCG GGGAAGATGT ATTCTTCTCA

TAAGATATAG TCGGACCTCT CCTTAATGGG AGGTAGCGGA TGAATGGATG

CAACACTGGA GCCGCTGGGA ACTAATTTGT ATGCGAAAGT ATATTGATTA

GTTTTGGAGT ACTCGTAAGG TAGCCAAATG CCTCGTCATC TAATTAGTGA

CGCGCATGAA TGGATTA [SEQ ID NO.1]

(Kan, N.C. *et al.*, *Nucl. Acids Res.* 10:2809-2822 (1982)).

如这里所述，利用该内含子的催化核心对本发明的定向转拼核酶进行操纵。利用已知方法可分离出内含子及其催化核心。内含子的催化核心（即平头内含子）与全长内含子的不同仅在于它在*ScaI*位点被剪平，从而除去了内含子的最后5个核苷酸。该平头内含子*RNA*可按已知技术制备或通过商业途径以药盒的形式得到，例如购自*US Biochemical*，*Clevelend*，*OH*的产品#72000（*RNAzyme*™ *Tet* 1.0*kit*）。该*US Biochemical kit*提供了核酶以及使用方法。可利用转录的*Tet*.1 *cDNA*作为如下所述的聚合酶链反应（*PCR*）突变的底物以产生合成的转拼酶。

本发明核酶的底物特异性（即核酶"对准"作为底物的特定*RNA*的能力）是通过将特异于靶（底物）*RNA*的互补序列融合于核酶的5′末端而产生的。

本发明核酶的定向转拼特异性（即利用靶*RNA*的序列转拼所需外源序列的特异性）是通过在核酶的3′端产生新的3′外显子而获得的。下面提供了设计的详情。

为了改变I组内含子的结构和催化性质，用外显子序列取代该内含子的侧翼序列以使得仅保留了内含子的催化核心（核酶）。所产生的修饰过的核酶能够与底物*RNA*进行转拼反应。当将平末端形式的内含子（即催化"核心"，在*ScaI*位点处平切以除去内含子的最后5个核苷酸）与天然核酶5′拼接连接处相应的序列一起保温时，该位点按与拼接第一步相同的方式进行鸟苷依赖性切割。

对本发明的核酶进行操纵需要考虑到以下四个准则。

首先，拼接位点必须选择在靶 *RNA* 内。在最后的转拼复合物中，只有 *P*1 双螺旋的 5′部分受靶 *RNA* 的作用。在紧接所需拼接位点的 5′端仅需要唯一的一个保守残基(尿嘧啶)，这是对靶 *RNA* 序列的唯一要求。对靶 *RNA* 没有天然结构要求。所作用的可能是成熟的 *mRNA*，并在细胞质中完成转拼反应，而不是在细胞核中作用于前 *mRNA*。从而避免了对细胞核中高浓度核酶的需要。

其次，在选定了特定的靶序列后，必须将补偿序列更换加至核酶的 5′部分，以在靶和核酶 *RNA* 之间形成合适的螺旋 *P*1。所非常需要的是，螺旋 *P*1 应在预期的 5′拼接位点处含有 *U*：*G* 碱基对并应位于从螺旋的底部开始的第 4、第 5(优选)或第 6 位(*Doudna, J. A., et al.,* "*RNA Structure, Not Sequence Determines The* 5′ *Splice—Site Specificity of a Group I Intron,*" *Proc. Natl. Acad. Sci. USA* 86:7402－7406 (1989)，在此引用以作参考))。对于天然嗜热四膜虫内含子，*P*1 比预期的 5′拼接位点延长 3 个碱基对，并且在一优选实施方案中，它在本发明的转拼核酶中被维持。为使转拼有效，底物和核糖核酸内切内含子 *RNA* 必须进行碱基配对的形成螺旋 *P*1，并带有所产生的变位 *U*：*G* 碱基对。靶 *RNA* 的切割发生在紧接 *U*：*G* 碱基对 3′端的磷酸二酯键处。系统发育比较和突变分析表明，只要螺旋 *P*1 的碱基配对得以维持，紧接 5′拼接位点处保守尿嘧啶残基的序列的性质对催化并不重要。

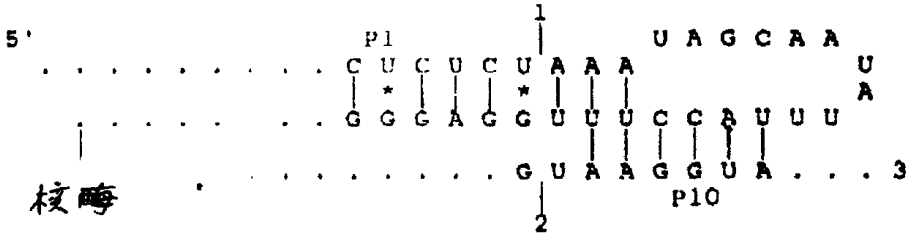
第三，必须选择3′拼接位点侧翼的外显子序列，并根据需要在核酶的5′部分进行调整以形成稳定的*P*10螺旋。尽管*P*10螺旋在某些情况下可以省却，但它的存在可增强拼接，且在本发明核酶的优选实施方案中保留了*P*10螺旋(*Suh, E. R. et al.*, "*Base pairing Betueen The 3′ Exon And An Internal Guide Sequence Increases 3′ Splice Site Specificity in theTetrahymena Self—Splicing rRNA Intron*," *Mol. Cell. Biol.* 10:2960—2965 (1990))。螺旋*P*1和*P*10沿嗜热四膜虫内含子内含子*IGs*重叠且紧接5′和3′拼接位点的第二和第三个残基与*IGS*中的相同残基互补（图2）。虽然这一点可能有某些优越性，但许多天然I组内含子并不受这一约束。因此，对3′外显子序列的选择主要是出于对实验的考虑。这种考虑反映了在选择拼接位点上的广泛灵活性。例如，如果需要将两种序列在某一特定位点进行连接，则该位点处的序列不能因转拼而突变或发生其它变化。如果重叠序列不另外对所需拼接位点进行调节*P*1或*P*10均可被缩短。

选择3′拼接位点所需要的序列似乎主要位于内含子(核酶)本身的结构中，其中包括螺旋*P*9.0和邻接3′端的划定3′内含子边界的鸟苷残基。*P*9.0完全包含在内含子序列中并帮助确定邻接的3′拼接位点。为设计转拼，*P*9.0螺旋及内含子中其余的功能*RNA*成分未被改变。*P*9.0螺旋的结构特征是已知的(*Michael, F. et al.*, "*The Guanosine Binding Site of the Tetrahymona Ribozyme*" *Na-*

ture 342:391－395 (1989))。但是,3′外显子中的侧翼序列是形成螺旋 *P*10 和有效拼接所必需的,如同突变分析所显出的一样。

第四 ,将互补序列区置于转拼核酶的 5′末端以增加其对靶 *RNA* 的亲和性和特异性。如这里所示,使用了约 40 个残基的任意长度。只要不影响所需效果也可使用其它长度。

例如,由如下嗜热四膜虫自拼内含子开始:



(在上图以及整个申请文本的其它核酶图中,"1"和"2"分别表示第 1 和第 2 拼接位点)。

(1)在选定的靶 *RNA* 内尿嘧啶残基的相邻处选择"5′"位点。引起互补性的序列与引起 *P*1 螺旋形成的序列并不是紧密相连,而是分开的,例如被也引起 *P*10 形成的 5 个核苷酸所隔开;

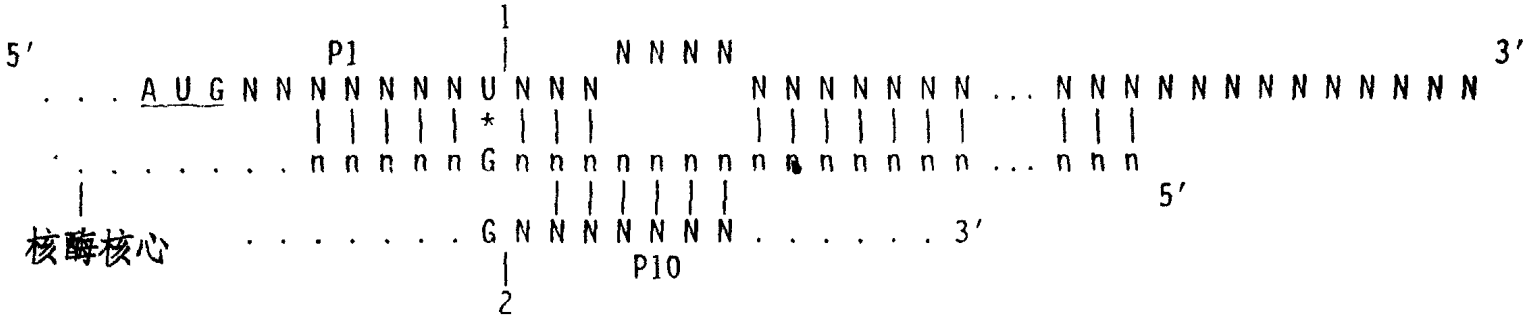
(2)将与选定的 *RNA* 互补的序列融合至自拼 I 组内含子的 5′部分。核酶和靶 *RNA* 之间的碱基配对导致螺旋 *P*1 的形成;

(3)将选定的"3′外显子"融合至核酶的 3′部分,并维持保守的螺旋 *P*10;及

(4)如果需要增加对靶 *RNA* 的亲和性，将延长序列互补性的部分融进核酶的 5′ 部分以形成 30—40 个碱基对。

所产生的转拼核酶与其靶 *RNA* 的排列用下图显示。顶上行表示靶 *RNA* 序列。核酶序列在其下排列，在其下两行 5′ 和 3′ 以及核酶的 *P*1 和 *P*10 的核苷酸杂交的区域内可看到一连续序列。

本发明核酶与靶ＲＮＡ的排列



按照本发明所设计的转拼核酶将转拼几乎任何 *RNA* 序列至任何 *RNA* 靶上。并不需要使该靶含有内含子序列或者使该核酶为靶序列中的内含子。例如，用于这一设计的方案可包括(1)鉴定所需的靶 *RNA*，(2)对该所需的靶 *RNA* 或其一部分进行克隆和/或序列分析，(3)筛选转拼至靶 *RNA* 中去的所需编码序列，(4)利用这里所述的准则构建能够与该靶杂交的本发明核酶，(5)证实本发明的核酶将利用该靶作为所需特定转拼反应的底物，(6)将核酶插入至所需宿主细胞中。

对靶 *RNA* 的选择将反映出所需转拼反应的目的。如果反应的目的是灭活特定的 *RNA*，则该 *RNA* 必须在破坏由该 *RNA* 编码的所有功能肽区的位置以及在不导致不需要的遗传序列连续表达的位置 被转拼 。如果存在一个以上的编码该 *RNA* 的基因的等位基因，则所设计的核酶最好能够在所有表达形式所共有的位点灭活靶 *RNA*。或者，给细胞提供一个以上的核酶，而每一核酶被设计成灭活靶 *RNA* 的特定等位基因的形式。

当仅需要灭活靶 *RNA* 而不需要表达新的所需 *RNA* 序列时，并不 需 要由核酶代表的外源 *RNA* 产生能够被宿主细胞翻译的序列，且可利用含有翻译终止密码子的序列作为平末端内含子，例如在 *ScaI* 位点切平的内含子核酶。

如果转拼反应的目的是为宿主细胞提供遗传特性 ，则对靶

RNA 的选择将反映出所需遗传特性的表达方式。如果需要使遗传特性被宿主连续表达，那么靶 *RNA* 也应被连续表达。如果需要使遗传特性仅在所需的生长、内分泌或环境条件下进行选择性表达，那么靶 *RNA* 也应在同样条件下被选择性地表达。

如果核酶的底物并不另外存在于宿主中，则不需要使核酶本身的表达专门限于在所需的生长、内分泌或环境条件下，因为核酶本身并不被宿主转译。因此，在发生转拼行为之前由本发明核酶所代表的 *RNA* 编码的序列并未在宿主中被翻译，且该转拼行为可被宿主中核酶底物的表达所控制。

如果需要的话，可对核酶的表达进行操纵，使其发生应答于诱导调节的靶表达的相同因素，或对核酶的表达进行操纵以产生另一水平的调节，使得转拼行为的发生仅限于在某些条件下，在该条件下核酶和靶均在细胞中仅被不同因素所选择性地诱导，而这些因素的结合则是不希望有的行为。这种调节将使得宿主细胞在核酶自身不被共同表达的条件下表达核酶靶。

与靶 *RNA* 杂交的核酶区的序列是由靶 *RNA* 的序列决定的。靶 *RNA* 的序列是在克隆编码该 *RNA* 的序列之后或在对由该靶 *RNA* 编码的多肽进行序列分析并由此推断出编码这种肽的 *RNA* 序列之后而确定的。可采用本领域已知的克隆技术对编码靶 *RNA* 的序列进行克隆。

可被转拼到靶 *RNA* 中的预期序列（在此称为“转拼序列”）的

选择可体现转拼的意图。如果期望一种不导致新的基因序列表达的转拼行为，则转拼序列不需要编码可转译的蛋白质序列。如果期望一种确实可导致某一新的基因序列，特别是新的肽或蛋白质序列表达的转拼行为，则转拼序列可进一步提供转译终止密码子，和*RNA*在宿主细胞内正确转译加工所必需的其它信息。如果期望转拼行为的结果是产生一种特定的蛋白质产物，则在所产生的融合体中保持该氨基酸解读密码将是必不可少的。

对本发明核酶的特异性的鉴定和确证可通过检测推断核酶仅在预期的靶序列存在下催化预期的转拼反应的能力来完成。如果存在的唯一*RNA*序列是该核酶不应对其敏感(或不太敏感)的非靶序列，则转拼反应不应发生。可借助于一种标记物进行这样的定性，以便可更容易地监测正确(或不正确)的核酶活性。在多数情况下，体外检测对抗其预期靶物的核酶，然后用它转化宿主细胞以研究其体内作用就足够了。

当需要除去宿主的*RNA*时，这样的删除应该是尽可能完全的。当需要给宿主细胞提供新的基因序列时，本发明的转拼反应不需要是完全的。本发明的优点在于，根据从这样的基因序列转译的肽的生物学活性，转拼活动的效率实际上可以是相当低的，只要发生的转拼可足以为预期的目的提供给宿主足够的*mRNA*，从而编码多肽即可。

本发明的核酶在宿主细胞中的转录发生在核酶基因引入到宿

主细胞之后。如果不需要由宿主细胞稳定保持该核酶，这样的核酶就可被化学合成或酶促合成并用机械方法，例如显微注射、脂质体介导的转染、电击、或磷酸钙沉淀提供给宿主。而当需要稳定保持编码该核酶的基因时，可通过将该核酶的至少一个 *DNA* 拷贝稳定地插入宿主的染色体，或通过提供一种该核酶的 *DNA* 拷贝于一个由宿主细胞稳定地保持的质粒上而实现。

最好是将本发明的核酶插入宿主的染色体作为表达暗盒的一部分，该暗盒可提供转录调节成分，这些成分可控制核酶在宿主细胞中的转录。这样的成分可包括，但不一定局限于，启动子成分、增强子或 *UAS* 成分、和转录终止信号。不转译核酶时，多聚腺苷酸化不是必需的。然而，对于编码被转拼的成分的序列，可提供这样的多聚腺苷酸化信号。

其编码序列已插入宿主染色体的核酶的表达受与核酶编码序列可操纵地连接的启动子序列所控制。可指导核酶表达的启动子可以是任何在宿主细胞中起作用的启动子，原核启动子是在原核细胞中所需要的，真核启动子则是在真核细胞中所需要的。启动子由分立的单元所组成，这些单元可指导启动子在宿主细胞中的转录激活和/或阻遏。在核酶的启动子中可使这样的单元混合和相配以便为核酶在宿主中的适当表达提供条件。真核启动子可以是任何可在真核细胞中发挥作用的启动子，特别是任何具有 *RNA* 聚合酶Ⅰ、Ⅱ、或Ⅲ特异性的启动子。如果需在多种真核宿主细胞中表达核酶，则

应选择可在多种真核宿主细胞中发挥作用的启动子，例如 *rRNA* 或 *tRNA* 启动子，或被广泛表达的 *mRNA* 的启动子，例如肌动蛋白基因，或糖酵解基因的启动子。如果只需在某一种细胞或组织类型中表达该核酶，则应选择只在该细胞或组织类型中发挥作用的细胞特异性（或组织特异性）启动子成分。

转拼反应无论是在体外还是在体内进行在化学上都是相同的。但在体内，由于辅助因子通常已存在于宿主细胞中，所以靶物和核糖酶的存在将足以导致转拼。

本发明的转拼核酶和方法可用于产生适用于靶细胞的基因修饰，和/或细胞死亡的基因活性。例如，本发明的转拼反应可用于将具有毒性的蛋白质引入所需细胞，细胞的感受性将由靶 *RNA* 和支配核酶的表达的调控物的选择来决定。例如，可操纵一种能使编码某一毒性蛋白的 *FNA* 序列转位的核酶，以便该核酶的表达将取决于可操纵连接的启动子的特性。在一个非常优选的实施方案中，白喉毒素肽 *A* 是由转位到宿主内预期靶物中的那部分核酶所编码的。核酶和白喉毒素肽 *A* 链的有条件表达可导致宿主细胞的死亡。其它有用的肽毒素包括蓖麻毒素、外毒素 *A*，和疱疹胸苷激酶（*Evans, G. A., Genes & Dev.* 3:259—263 (1989)）。此外，各种水解酶具有破坏细胞代谢的潜能。例如，真菌的核糖核酸酶可用于引起植物的雄性不育（*Mariani, C. et al., Nature* 347:737—741 (1990)）。由于靶 *RNA* 的有限表达可破坏特定的组织。再者，如果用病毒

RNA 作为靶物，则可操纵新形式的病毒抗性或治疗。

可用本发明的核酶设计一种控制组织特异性基因表达和/或用于异位切除的二元系统。例如，可以使用利用 *P*—成分增强子捕集载体以组织特异性和空间特异性方式表达酵母转录激活子 *GAL4* 果蝇品系。任何转录激活子都可用来代替 *GAL4*，本发明并非只限于 *GAL4*。可将一种编码能够转拼 *DTA* 序列的融合核酶的基因置于 *GAL4—UAS* 启动子的控制之下并以稳定遗传的方式插入果蝇，这样的核酶在无 *GAL4* 存在下不能在果蝇中得以表达。因此，遗传上携带这一核酶构建物的果蝇宿主与以组织特异性方式表达 *GAL4* 的果蝇宿主杂交可导致子代随机编码，当 *GAL4* 表达被诱导时，显示出与 *GAL4* 表达方式相似的细胞死亡方式。

此外，通过将核酶作为与 *GAL4 mRNA* 转拼的对象，该核酶的拼接活性可钝化 *GAL4* 的表达并可自我调节核酶的表达。

核酶原

如上所述，转拼核酶由三种融合序列成分所组成，即与靶 *RNA* 互补的 5′“反义”区，基于自拼 I 组内含子的催化区，和 3′“外显子”序列。5′区可与所选靶 *RNA* 进行碱基配对，使之与 I 组内含子的催化序列接近。I 组内含子的结构可提供一种适于催化靶 *RNA* 与 3′“外显子”序列精确拼接的化学环境。然而，在无合适的靶 *RNA* 存在下，该核酶序列仍能催化在 3′“外显子”连接处的切割（相似的水解可见于 I 组自拼内含子(*Zaug et al.*, *Science*

231:470—475(1986)),并且可能能通过核酶与异源 *RNA* 序列(可包括它们自身)的瞬时碱基配对催化不正常的拼接活动。这样的副反应和不正常的拼接活动是多余的,或许是有害的。例如,如果要将转拼用于毒素在体内的条件释放,则不正常的转拼可能会导致有毒活性的不期望表达。在 3′"外显子"连接处的自发性切割会降低转拼的效率。

为了有助于避免这些问题,已构建了核酶原形式的转拼 *RNA*,例如其中螺旋 *P*8 被破坏。构建的核酶原含有额外的自我互补序列,这些序列可导致核酶催化中心误折。在预期靶 *RNA* 不存在时,核酶原是无活性的;只有在核酶和靶 *RNA* 进行碱基配对一伴随着对核酶内二级结构干扰的取代之后才能形成活性形式。核酶原是指在靶 *RNA* 不存在下无催化活性的种类,从而可消除体外和体内多余的自我切割、自拼和不正常的转拼反应(图 12)。

这里所述的核酶原是构象受到破坏的且因此成为无转拼活性的形式。该核酶原具有很少的自我切割活性。它们只能通过与靶 *RNA* 特异地相互作用重新活化,因而是不大可能催化非预期靶 *RNA* 的转拼的底物活化核酶。期望将转拼核酶用于新基因活性在体内的释放,在不希望的副反应或不正常拼接的程度上的任何减少都是可取的,而且可能是必需的。

尽管对于转拼核酶原这里已举例说明了螺旋 *P*8 的破坏,但也可能已应用了催化活性所需的其它螺旋。

以这样的方式，即只与预期的底物 *RNA* 进行碱基配对可形成活性核酶，破坏催化上重要结构的构象的同样方法可应用于其它核酶的设计。例如，"锤头"型核糖核酸内切酶的环序列(*Haseloff et al.*，*Nature* 334:585—591 (1988))可被扩展并使之与核酶的"反义"臂之一互补一与螺旋 *P8* 的以上修饰相似。只有在与所选的靶 *RNA* 进行碱基配对，对二级结构的破坏被取代，和催化所需的干环结构重新形成之后才会显示出核糖核酸内切酶活性。这样将有效地增强核酶对其靶物的特异性。

此外，核酶原的激活不需依赖于与底物本身进行碱基配对。附加的碱基配对或 *RNA* 与蛋白质间的相互作用可能是活性核酶复合 物的形成所需的 。这些额外组分的可得性将决定核酶的活性，并可被用来改变核酶的选择性。

本发明的核酶或核酶原可被引入任何原核或真核宿主细胞，特别是植物或哺乳动物宿主细胞，尤其是人的细胞，既可是细胞培养物也可是活体细胞，所用的技术是本领域中已知的适合于上述宿主的方法。也可操纵本发明的核酶以破坏病毒。在一个优选的实施方案中，在病毒侵袭之前以一种遗传上稳定的方式将本发明的核酶或核酶原提供给宿主细胞。合适病毒的感染，或潜伏性病毒在这种宿主细胞中的表达（导致宿主细胞中出现核酶或核酶原的靶 *RNA*），将通过可引起病毒感染细胞死亡的转拼激发该核酶的催化活性和病毒 *RNA* 靶的破坏和/或毒素的产生。在另一个实施方案

中，可操纵核酶或核酶原并包装到病毒自身中。这样的实施方案特别适用于为研究目的所用病毒的设计，其中设计出的核酶或核酶原可破坏特异性病毒 *RNA* 的功能，因而可在这种 *RNA* 不存在时研究病毒的功能。携带核酶的病毒也可用作转染具有所需核酶或核酶原活性的宿主细胞的载体。

可利用本发明的核酶或核酶原在农学重要性物种中操纵雄性或雌性不育。例如，可通过将 *TA*29 或 *TA*13 *mRNA*（烟草花 药特异性基因；*Seurinck*，*J. et al.*，*Nucl. Acids Res.* 18：3403 (1990)）作为对象，用可将 *DTA*3′外显子转拼到那些靶物中的本发明的核酶或核酶原操纵烟草的雄性不育。

可通过用本发明的核酶或核酶原对组织进行选择性破坏或修饰控制农作物的形式。例如，可通过将种子贮存蛋白 *mRNA* 作为对象，用可将 *DTA*3′外显子转拼到该靶物中的本发明的核酶或核酶原获得无籽果实。

可通过病毒特异性核酶或核酶原的表达杀死感染细胞而保护转基因植物抵抗感染。这将成为一种人工形式的“超敏反应”。例如，可用本发明的核酶或核酶原（可将 *DTA*3′外显子转拼到靶物中）以黄瓜花叶病毒外壳蛋白 *mRNA* 作为对象。

经引入转拼核酶或核酶原可使微生物群体变得耐受特异性病原体。例如，通过将细菌毒素基因或由本发明的核酶或核酶原所提供的 3′外显子编码的水解酶作用于噬菌体 *RNA* 可使干酪产生细

菌耐受噬菌体感染。例如在噬菌体感染后通过引起过早溶菌现象而干扰噬菌体复制。

可构建通过转拼释放有毒活性的病毒病原体。以这种方式，可以特定的细胞类型作为切除的对象，如进行癌症或病毒治疗。例如，可通过携带 *DTA*3′外显子的本发明的核酶或核酶原以 *HIV mRNA* 作为对象进行病毒或脂质体释放。

下列实施例只是为说明目的，不能看作是对本发明范围的限制。

实施例 1

CAT—*LacZ* 转拼核酶的构建和定性

Ⅰ. 本发明核酶的 *PCR* 扩增和克隆

按照上述指导路线，设计一种转拼融合核酶，该酶可将大肠杆菌 β—半乳糖苷酶(*LacZ*)*mRNA* 的部分氨基末端编码序列拼接到氯霉素乙酰基转移酶 (*CAT*)*mRNA* 中的某一位点上(图 3)。将侧接嗜热四膜虫(*T. thermophila*)核酶核心和 3′外显子的新序列的各区域合成为寡核苷酸。然后通过连续聚合酶链反应，用合成连接寡核苷酸作为引物与核酶和 β—半乳糖苷酶 *DNA* 模板一起装配完整的核酶序列(尽管存在其它可利用的方法，但该方法是最便利的)。

为了构建一种能够将 β—半乳糖苷酶(*LacZ*)α—肽编码序列拼

接到氯霉素乙酰基转移酶(*CAT*)的5′编码序列中的某一位点上的核酶，合成了三种寡核苷酸。

寡核苷酸1

5′—*GGCCA AGCTT CTTTA CGATG CCATT GGGAT ATATC AACGG TGGTA TAAAC CCGTG GTTTT TAAAA GTTAT CAGGC ATGCA CC*—3′ [SEQ ID NO. 2]

寡核苷酸2

5′—*GATTA GTTTT GGAGT ACTCG TACGG ATTCA CGGCC GTCGT TTTAC AA*—3′ TTTAC AA-3 [SEQ ID NO. 3]

寡核苷酸3

5′—*GGCCG AATTC TTACA ATTTC CATTC AGGCT GCGCA ACTGT TGG*—3′ [SEQ ID NO. 4]

将寡核苷酸2和3(各200*pmoles*)与0.1*μg* *Pvu* Ⅰ切割的*pGEM4* *DNA*(含有*LacZ* *α*—肽序列)合并，以100*μl*的体积进行*PCR*扩增：

50*mM* *KCl*.

10*mM* *Tris*-*HCl* *PH*8.3，

1.5*mM* $MgCl_2$，

0.4*mM* *dNTPs*

0.1%明胶，和

5*U* *TaqI* *DNA*聚合酶，

并于94℃下保温1分钟，50℃下保温2分钟，72℃下保温2分钟，这一保温循环共进行30次。

质粒 *pGEM4* 可购自美国 *Promega* 公司（*Madison WI, USA*）。

用低胶凝温度琼脂糖电泳法纯化210碱基对的扩增产物，并用作第二轮 *PCR* 扩增的引物。

在第二轮 *PCR* 扩增之后，将 2.0μg 210碱基对的扩增产物、200*pmoles* 寡核苷酸1和0.1μg 含有嗜热四膜虫 *IVS* 的450碱基对的片段混合并采用上述条件进行 *PCR* 扩增。用限制性核酸内切酶 *EcoRI* 和 *Hind* Ⅲ消化所得到的660碱基对的产物，并克隆到质粒载体 *pGEM4* 中。*CAT—LacZ* α—肽核酶 *DNA* 序列的完整序列用 *SEQ ID No.5* 表示，并示于图 *3B*。

用本领域中已知的技术（*Maniatis, Molecular Cloning, A Laboratory Guide, 2nd edition*, 1989, *Cold Spring Harbor Laboratory, Publishers*），将含有克隆序列的克隆载体转化到细菌宿主 *XL1/Blue*（*Strategene, La Jolla; California*）中并在那里增殖。然而，任何能够稳定维持该载体的细菌宿主均可使用，例如 *JM*109。

为了进一步分析，可用本领域中公知的技术（*Maniatis, Molecular Cloning, A Laboratory Guide, 2nd, edition*, 1989, *Cold Spring Harbor Laboratory, Publishers*）从宿主细胞中

提取质粒。

Ⅰ. 克隆核酶和靶 *RNA* 的体外转录

采用标准方法，从细菌宿主中纯化克隆序列并用不切割核酶序列的限制性核酸内切酶（例如，*EcoRI*）将质粒线性化，用 T_7 *RNA* 聚合酶进行转录，体积为 100*μl*，含有：

5*μg* 线性质粒 *DNA*，

40*mM* *Tris*－*HCl* *PH*7.5，

6*mM* $MgCl_2$，

2*mM* 亚精胺，

10*mM* *NaCl*，

10*mM* *DTT*，

1*mM* *NTPs*（含有 20*μCi*〔$\alpha-^{32}P$〕*UTP*，如果需要被标记的 *RNA* 转录本），

100*U* *RNA* 酶，和

50*U* T_7 *RNA* 聚合酶，

反应物在 37℃下保温 2 小时。

使用前用 5% 聚丙烯酰胺凝胶电泳法（*TBE*，7*M* 尿素凝胶）纯化 *RNA* 转录本。含有活性嗜热四膜虫 *IVA* 序列的 *RNA* 在转录期间经历在 3′ 内含子－外显子连接处某种自发性切割。经电泳纯化移出这些片段用于在后面的转拼测定过程中清楚地进行分析。

Ⅱ. 体外转拼反应条件

将靶物和/或转拼核酶在下列条件下保温：0.1—0.5μg *RNA* 组分（其用量取决于实验的类型，通常核酶的量比靶物量多 5 倍），

30*mM* *Tris*—*HCl* *PH*7.5，

100*mM* *NaCl*，

2*mM* *GTP*，

5*mM* $MgCl_2$，

体积为 5μl，42℃下，60 分钟。

用 95μl 0.1*mM* Na_2EDTA，200*mM* *NaCl* 稀释反应物，并用乙醇沉淀，然后在 5%含有 *TBE* 缓冲液，7*M* 尿素和 25%甲酰胺的聚丙烯酰胺凝胶上分析 *RNA*，并放射自显影。

Ⅳ. 核酸内切酶解活性的测定

核酶和靶物进行碱基配对之后，转拼的第一步是鸟苷介导的在预定 5′ 拼接位点靶 *RNA* 的切割。退火和转拼可在缓冲液如 30*mM* *Tris*—*HCl*（*PH*7.5），100*mM* *NaCl*，5*mM* $MgCl_2$，2*mM* *GTP* 中，在 42℃下进行。由于 3′拼接位点是这一反应所必需的，所以平末端转拼核酶应表现得如同高度特异性核糖核酸内切酶一样，为了测试这一活性，将上述 *CAT*—*LacZ* α—肽转拼核酶的缩短了的体外转录本（*SEQ ID No.* 5 和图 3）与 *CAT mRNA* 序列一起保温，*CAT*—*LacZ* 核酶暗盒在 *Hind* Ⅲ—*EcoRI* 片段上。*ScaI* 切割位点可标明 3′拼接位点上游 5 位碱基。该核酶特异地在所期望的单一位点切割靶 *RNA* 以产生所期望大小的片段。

Ⅴ. 转拼反应

为了确证 *CAT－LacZ* α—肽核酶催化 3′外显子序列在 5′拼接位点的连接的能力，将不同的形式与放射标记的 *CAT RNA* 一起保温。核酶转录本由 *DNA* 模板合成，它在由核酶核心末端至外显子序列之间的几个位置被 3′平切。与标记过的 *CAT* 一起保温导致预期拼接产物的形成，这些产物的长度依 3′外显子序列的长度而异。

此外，一定比例的 *CAT－LacZ* α—肽核酶分子在体外转录期间经历了在 3′拼接位点的自发性切割，与完整嗜热四膜虫内含子相似。这些切割形式（在邻近 3′拼接位点的鸟苷残基处终止）也与 *CAT RNA* 一起保温。在这种情况下，核酶自身可与 *CAT RNA* 的 3′部分连接，以产生大小约为 550 个核苷酸的产物。这一反应与完整内含子的自我成环相似，同样的连接反应见于其它转拼反应。

Ⅵ. 转拼的精度

将得自 *CAT－LacZ* α—肽转拼反应的产物逆转录，并通过聚合酶链反应，用两种与预定拼接位点任一侧上的序列互补的寡核苷酸通过聚合酶链反应进行扩增。将扩增的序列克隆并进行序列分析。各个重组体从拼接产物的预期序列来看都无变异。如对完整内含子的研究中所见，拼接似乎是相当精确的。

因此，以上研究表明按照本发明的指导路线设计的转拼核酶能够在体外精确、有效的转拼。

实施例 2

提供植物病毒抗性的转拼核酶的设计

黄瓜花叶病毒(*CMV*)是一种有很多已知毒株的广泛流行的病毒。在 *RNA*3 和亚基因组 *mRNA*4 所编码的病毒外壳蛋白顺反子的起始区域中显示了病毒株的九条序列(*SEQ ID Nos.* 7—25 图 4(*A*)和图 5)。已选出两个位点,它们保存在序列中,位于外壳蛋白的 *AUG* 起始密码子的下游。用于构建能够将异亮氨酸突变型 *DTA* 转拼到 *CMV* 外壳蛋白 *mRNA* 中的核酶的寡核苷酸示于图 4*B* 且在下文中进行了讨论。

以图 4*C* 和 4*D* 中所示的转拼核酶作用于图 4*B* 中所示的 *CMV* 病毒序列,不但会导致 *CMV RNA* 分子的切割,而且会导致受感染的细胞中白喉毒素 *A*—链的表达。可用本领域中已知的技术将图 4 所示的转拼暗盒转化到任何 *CMV* 感性植物中,并使转基因的子代受到 *CMV* 感染。应这样设计核酶,即病毒感染是引发通过 *RNA* 转拼产生毒素所必需的,因为核酶本身不被转译。由毒素的表达而造成的被感染细胞的局部死亡可限制病毒在显示人为超敏反应的植物体内的复制和蔓延。

实施例 3

*Gal*4—白喉毒素 *A* 链转拼核酶的构建和定性

按照本发明和实施例 1 所述的方法,已设计出一种融合核酶,

它是一种 *Gal4*—白喉毒素 *A* 链转拼核酶(图 6)(该核酶的序列用 *SEQ ID No.* 6 表示)。*GAL4—DTA* 核酶暗盒是 *SalI—XhoI* 片段。*ScaI* 位点可标明 3′拼接位点上游 的 5 位碱基 。这种酶能够将白喉毒素 *A* 链的编码序列拼接到 *GAL4 mRNA*5′区的某一位点上 。这种转拼活性在体外(如上)和体内(见下)均有活性。成功设计 *GAL4—DTA* 核酶,和任何可转拼编码毒性产物的序列的转拼核酶的主要标准在于不但有效和精确催化转拼,而且毒性产物(如 *DTA*)的表达不在无转拼的情况下发生。

按照上述设计构建核酶的催化部分 ,然后分别在 *GAL4* 和 *DTA* 的 5′编码区选择 5′和 3′拼接位点。3′外显子序列相应于已用于在真核细胞中表达的 *DTA* 基因的 3′外显子序列,所不同的是删除了第一个 *AUG* 密码子和几个近端氨基酸。原始白喉棒杆菌形式的 *DTA* 在这个 5′区也不同,它是利用 *CUG* 密码子启动转译。原始 *DTA* 序列也含有一个后来不存在的信号肽前导序列。

这些核酶分子可经历在 3′拼接位点的自发性切割 。假使 *DTA* 的毒性极强,任何释放的 3′外显子序列都不产生毒性转译产物 这一点是重要的 。3′外显子在 13 位含有一个码组内蛋氨酸,设想它可能会产生一种平头但有毒性的多肽。为了消除这种可能性 ,在两个独立的核酶构建中将野生型序列 (*Rz—DTAmet*) 该位置上 的蛋氨酸改为异亮氨酸 (*Rz—DTAile*)或亮氨酸(*Rz—DTAleu*) (图 6)。

实施例 4

本发明核酶的体内活性

Ⅰ. 简介

按本文所提供的指导路线设计的核酶的体内活性和这种核酶传给宿主细胞新的基因活性的能力，用所述的 *Gal*4—白喉毒素 *A* 链转拼核酶(实施例 3 和图 6)将高毒性白喉毒素 *A* 产物传给宿主细胞得到了证实。在该系统中，果蝇是选定的宿主，期望在果蝇宿主中以组织特异性方式控制本发明核糖酶的表达。

白喉毒素是由对 *B* 噬菌体为溶源性的白喉棒杆菌分泌的。该毒素以单个多肽形式产生，该多肽经蛋白水解产生 *A* 和 *B* 链。*A* 链(*DTA*)含有很强的 *ADP* 核糖基酶活性，该活性对真核细胞转译延伸因子 *EF*—2 具有特异性。即使仅有该酶的几个分子存在就足以在多种真核细胞中引起转译停止和最终死亡。*B* 链允许毒素通过结合甘露糖残基附着于细胞表面受体上在胞内释放，被细胞内吞并经泡囊融合进入细胞质。

在 *B*—链不存在时，*A*—链在胞外存在时毒性小得多。这一特性，和它的高毒性已提示了它在异位切除实验中的用途。例如，用视蛋白启动子引发在发育眼球中的表达，已在转基因小鼠中表达了编码 *DTA* 的序列。产生的小鼠双目失明，眼球畸形(*Breitman, M. L.*, *Science* 238:1563—1565 (1987))。在其它研究中，进行了小鼠胰腺的切除(*Palmiter, R. D. et al.*, *Cell* 50:435—443 (1987))，

Wert, S. E. et al., Am. Rev. Respir. Dis. 141(*no.* 4, *part* 2):*A* 695 (1990)描述了通过使用由人表面蛋白 *C* 基因(在 I 型肺泡细胞中被表达)的启动子和 5′侧翼序列和 *DTA* 基因组成的嵌合基因切除肺泡细胞。

然而，用这种类型的方法，不可能维持或繁殖可能具有更严重，或致死性表型的转化生物体。此外，目前未有用完整 *DTA* 序列对某些物种（如果蝇）进行转化的报道。在这样的转化过程中 *DTA* 基因的渗漏表达可导致即刻死亡。

I. 果蝇系统

已经发展了一种靶基因表达的一般方法。首先，该系统可使各个品系迅速产生。其中异位基因表达可涉及不同的组织或细胞类型：利用增强子检测剂技术(*O'Kane, C. J. and Gehring, W. J., Proc. Natl. Acad. Sci. USA*:9123—9127 (1987); *Bellen et al., Genes and Development* 3:1288—1300 (1989); *Bier et al., Genes and Development* 3:1273—1287 (1989))以各种方式在胚胎、幼体和成体内表达转录激活子。第二，该方法可在不同的品系中从它的靶基因中分离出激活子，以确保各个亲代品系都是活的：在一个品系中，激活子蛋白存在，但没有可激活的靶基因，在第二个品系中，靶基因是不活动的。当使这两个品系杂交时，只在杂交的子代中激活靶基因，从而使显性表型(包括致死性)可被方便地研究。

为了只异位表达令人感兴趣的基因，需要一种在果蝇中没有内

源靶的转录激活子。一种来自酵母的激活子，*Gal4* 能激活在果蝇中的转录，但转录只从具有 *Gal4* 结合位点（*Fischer et al.*，*Nature* 332：853—865（1988））的启动子开始。对于靶基因表达，以两种方式将 *Gal4* 限制到特定的细胞：或者用特征化果蝇启动子启动 *Gal4* 转录，或者将无增强子的 *Gal4* 基因随机整合到果蝇基因组中，使之在多种排列形式的基因组增强子的控制之下。为了测定通过 *Gal4* 的转激活，使可表达 *Gal4* 的果蝇与其转录是由 *Gal4* 结合位点（*Fischer et al.*，*Nature* 332：853—865（1988））启动的带有 *LacZ* 基因的果蝇杂交。β—半乳糖苷酶只在 *Gal4* 先被表达的那些细胞中表达。*LacZ* 的组织特异性和细胞特异性转激活已在 *Gal4* 被表达和各种方式被建立的品系中得到了证实。

用该系统，有可能：1）将 *Gal4* 结合位点置于任何编码序列上游；2）只在 *Gal4* 被表达的细胞中激活该基因和 3）观察这种异常表达对发育的影响。在异位表达是致死性的情况下，这种方法可使两个亲代品系（一个表达 *Gal4*，另一个携带不活动基因，在它的启动子中具有 *Gal4* 结合位点）得以稳定地繁殖。然后在杂交的子代中研究这些表型。

Ⅲ. 载体

在这些研究中用作起始材料的载体包括：

1）*pGATB* 和 *pGATN*（图 9）：这些载体可被用于克隆启动子和无启动子 *Gal4* 基因上游的增强子。

构建载体，其中独特的 *NotI* 或 *BamHI* 位点被插入 *Gal4* 编码区的上游。一旦将启动子连接到 *Gal4* 编码序列上，就可从 *pH-SREM* 载体骨架上切下该基因(*Knipple and Marsella—Herrick, Nucl. Acids Res.* 16：7748 (1988))并将其移入 *P*—成分载体。*Ph2* 启动子已被克隆(*Mismer et al.*, *Genetics* 120：173—180(1988))到该载体中并已产生 *Gal4* 只在单眼中被表达的果蝇。

2) *pGawB*：这是一种在增强子检测中使用的 *Gal4* 载体。

将无增强子的 *Gal4* 基因亚克隆到载体 *plwB*(*Wilson et al.*, *Genes and Development* 3：1301—1313 (1989))以产生 *pGawB*。*plwB* 先用 *Hind* Ⅲ 消化以除去 *lacZ* 基因和 *P* 转位酶基因的 *N* 末端。用 *P* 转位酶基因的 *TATA* 盒之后的整个 *Gal4* 编码区置换这些基因。

3) *pUAST*（图 10）：这个质粒被用于克隆 *Gal UAS* 下游的编码序列。

在 *Gal4 USA*(上游激活序列)之后可亚克隆进基因的载体在 *P* 成分载体 *pCaSpeR*$_3$(*C. Thummel, Univ. of Utah Medical Center, Salt Lake City, Utah, Personal communication*)中被构建。插入了 5 个 *Gal4* 结合位点，接着是 *hsp70TATA* 盒和转录起始密码子，多连接子，和 *SV40* 内含子和多聚腺苷酸化位点。基因或 *cDNA* 可被插入的独特位点包括：*EcoRI*, *Bgl* Ⅰ, *NotI*, *XhoI*, *KpnI* 和 *XbaI*。

Ⅳ. 果蝇品系

本文所述的用于定性在这些研究中所用的果蝇品系的遗传学技术是本领域中众所周知的(“*Genetic Variations of Drosophila melanogaster*,”*D. Lindsley and E. H Grell, eds*)。

用携带△2－3的“跳跃起始子(*jumpstarter*)”品系调动 *p* 成分转位子,△2－3是在可表达高水平的组成活性转位酶的第三染色体上的缺陷性 *p* 成分(*Robertson et al.*, *Genetics* 118:451－470(1988))。通常用于产生插入品系和制图的三个原种保藏于果蝇贮存中心(*the Drosophila Stock Center, Indiana University Department of Biology, Jordan Hall A*503, *Bloomington, Indiana* 47405):

1 :*yw*;＋/＋;*Sb P*[*ry*$^{+}$,△2－3*J*/*TM*6,*Ubx*

2:*w*;＋/＋;*TM*3,*Sb*/*CxD*(保藏号 3665)

3:*w*;*Cyo*/*Sco*;＋/＋(保藏号 3666)

其中三个染色体的遗传学特征用分号隔开。因此,例如,在品系 1 中,第一染色体(*X* 染色体)是黄色和白色纯合的(“*yw*”),第二染色体是野生型(“＋/＋”),第三染色体携带残余基因(“*Sb*”),和 *P* 成分转位于玫瑰色基因(“*ry*”)和△2－3,而第二三染色体携带平衡器倒位(“/*TM*6,*Ubx*)。

Ⅴ. 产生 *Gal*4 表达模式的方法

A. 用于分离转化体的方案

将构建物注射到由原种产生的胚胎中；

♀♀ y w/y w ; Δ2-3,Sb/TM6,Ubx X ♂♂ y w/Y ; Δ2-3,Sb/TM6,Ubx

F0；建立单个品系

♀ y w/y w ; Δ2-3,Sb/TM6,Ubx X ♂ y w/Y; +/+

or

♂ y w/Y ; Δ2-3,Sb/TM6,Ubx X ♀ y w/y w; +/+

F1；选择〔*w*⁺〕和〔*Sb*⁺〕子代并建立原种

♀ y w/y w ; +/TM6,Ubx X ♂ y w/Y; +/+

or

♂ y w/Y ; +/TM6,Ubx X ♀ y w/y w; +/+

B. 用于跳过无增强子 *Gal4* 插入体的方案

1. 从 *X*—染色体跳跃

♀♀ FM3/FM7,w; +/+ X ♂♂ y w/Y; Δ2-3,Sb/TM6,Ubx

♀♀ FM7,w/ P[Gal4,w^+] X ♂♂ FM7/Y ; Δ2-3,Sb/+

♀♀ FM7,w/ P[Gal4,w^+]; Δ2-3,Sb/+ X ♂♂ y w/Y; +/+

♀ FM7,w/ y w ; Δ2-3,Sb/+ X ♂ y w/Y; +/+

选择〔w^+〕和〔*B*〕子代并建立原种

2. 从∧2—3—染色体跳跃

♀♀ y w/y w X ♂♂ y w/Y ; P[Gal4,w^+], Δ2-3,Sb/+

选择〔w^+〕和〔Sb^+〕子代并建立原种。

C. 染色体分离

为了分析插入体的分离，使用两个原种：*w*；＋/＋；*TM3*，*Sb*/*CxD* 和 *w*； *Cyo/Sco*；＋/＋。

方法

为了产生大量可以细胞特异性和组织特异性方式表达 *Gal4* 的品系，已构建了携带不同形式的 *Gal4* 基因的增强子检测载体。两个编码全长度蛋白质或平末端蛋白质的基因已被克隆到玫瑰色（ry^+）和白色（w^+）*P* 成分载体中（*P* 成分载体为修饰形式的 *plArB* 和 *plwB*；*Wilson et al. Genes and Development* 3：1301—1313(1989)）。用 ry^+ 或 w^+ 作为筛，已通过引入△2—3 基因（*Robertson et al.*, *Genetics* 118：461—470(1988)）调动了这些载体。为了显现 *Gal4* 的表达方式，使 *Gal4* 插入品系与携带在 *Gal4 UAS*（*Fischer et al.*, *Nature*332：853—865(1988)）的控制下的 *lacZ* 基因的品系杂交。通过用 *X*—半乳糖作为底物的酶测定法，或通过用抗 β—半乳糖苷酶的单克隆抗体染色检查由这些杂交产生的胚胎、幼体和成体中 β—半乳糖苷酶的表达。由 *UAS*—*lacZ* 构建物编码的 β—半乳糖苷酶位于细胞质中。

已筛选出大约 500 个 *Gal4* 插入品系并且已鉴定了可用于激活特定组织中的基因的许多品系，例如鉴定了它们的表皮条纹，中胚层，中枢神经系统和外周神经系统。许多品系可在唾腺以及其它组

织中表达β—半乳糖苷酶。在构建无增强子 *Gal4* 转位子时有可能偶然产生位置依赖性唾腺增强子。

Ⅵ. 样品筛选

品系号	未着色	唾腺	其它组织	%
9	+	—	—	5.8
45	—	+	—	28.8
81	—	+	+	51.9
21	—	—	+	13.5
156				100.0

为了以特异性方式激活基因(*X* 基因),选择 *Gal4* 插入品系,与携带在 *GAL UAS* 之后被克隆的 *X* 基因的品系杂交。

Ⅶ. 对无核酶的 *GAL4/UAS* 系统的概述

Gal4/UAS 系统是一个用于控制基因活化的两部分系统。该方法是多方面适用的,可以是组织特异性的并且不会显示基础水平的表达,也许如本文所述,*UAS—DTA* 构建物例外。可用该方法异位表达特征化的基因,表达否则对生物体将是致死性的被修饰的基因和表达来自其它物种的基因以研究它们对果蝇发育的影响。由于该方法使产生显性,功能获得性突变成为可能,所以上位显性试验和对可见或致死表型的增强子或抑制基因的筛选可得以进行。*Gal4* 系统也允许有毒产物表达,便于研究细胞特异性和组织特异性切除

的后果。

Ⅷ. 含有本发明的 *DTA* 核酶的 *Gal4* 表达果蝇的应用

将携带编码 *DTA* 蛋白质的序列的融合核酶的表达置于 *pUAST*(图 10)中的 *GAL4UAS*(上游激活序列)的控制之下。如上所述,用上述被修饰的 *P* 成分增强子捕集载体,构建了大量稳定的果蝇品系,每个品系都可在发育的果蝇中以特异性时空方式表达酵母转录激活子。任何在 *GAL4* 上游激活子序列(*UAS*)控制下的基因都可单独被转化并维持,然后在特定的果蝇组织中通过与可表达 *GAL4* 的品系(图 7)杂交被诱导。然而,在不进一步修饰的情况下利用 *GAL4* 系统表达 *DTA* 本身是不可能的,因为通过 *DTA* 的渗漏表达产生 *UAS—DTA* 转化体是困难的。

已发现该二元系统用作通过转拼核酶(图 6)条件性表达 *DTA* 的工具克服了这些问题。在可表达 *GAL4* 的那些细胞中,*GAL4* 蛋白质可提供核酶转录所必需的活性,而 *GAL4 mRNA* 可提供 *DTA* 产生所必需的靶。

可得用本领域中已知的技术,给果蝇胚胎注射如上所述的置于 *UAS* 启动子控制之下的核酶序列。注射了 *Rz—DTAmet* 构建物的胚胎不能存活,而正常的转化果蝇得自注射了 *Rz—DTAile* 和 *Rz—DTAleu* 的胚胎。这一结果表明内部 *AUG* 密码子确实成为注射后的有毒产物转译的启动密码子。该密码子与在 *DTA* 序列中所提出的 NAD^+ 结合位点相邻,并与远缘外毒素 *A*(来自铜绿假单胞

菌的另一种 *EF*－2 特异性 *ADP*－核糖基酶)中所保留的序列相邻。

含有在 *GAL4 UAS* 控制下的 *Rz*－*DTAile* 和 *Rz*－*DTAleu* 的转基因果蝇与可以特定的表达方式产生 *GAL4* 的果蝇杂交。例如，在一个特征化品系 1*J*3 品系中，将 *GAL4* 基因插在毛发基因附近，并反映它的表达方式。在胚胎发生过程中在偶数编号腹节的表皮条纹中产生毛发基因产物。当将 *UAS* 启动的 *LacZ* 基因引入 1*J*3(其中 *GAL4* 以与毛发基因产物相同的方式表达)时，发现 *β*－半乳糖苷酶局限在偶数编号条纹中。当含有 *Rz*－*DTAleu* 基因的果蝇与这种 *GAL4* 表达品系杂交时，产生正常子代。然而，当含有 *Rz*－*DTAile* 的果蝇与 *GAL4* 表达品系杂交时，在胚胎发生中子代的发育受到抑制。颜色较深的条带在胚胎的表皮上明显可见，这与潜在细胞的死亡是一致的。当检查表层制备时，偶数编号的小齿状突起条带被破坏或消失，特别是第 4,6 和 8 条纹的小齿状突起条带(图 11)。当含有 *Rz*－*DTAile* 的果蝇与含有不同的 *GAL4* 表达基因的果蝇杂交时，观察到细胞死亡的其它特异性方式。

设计

实施例 5

核酶原的设计

作为一种核酶原的设计试验，修饰了前面描述的 *CAT*－*LacZ* 转拼核酶。系统发育比较和突变分析(参见 *Cech*, *Ann Rev.*

Biochem. 59:543--568 (1990))已表明 I 组自拼内含子的核心区被高度保留，它对活性是重要的(图 8)。为了构建转拼核酶原，紧接在该区旁的螺旋 *P*8 被破坏。在最初的实验中，将新序列的 13 或 18 个核苷酸引入螺旋 *P*8 的 5′股和环，分别产生核酶原 1 和 2。额外的核苷酸与核酶的 5′“反义”部分互补，同时调整侧翼序列以保留(1)在 *P*8 碱基处的真实序列，和(2)在 *P*8 内可能的碱基配对的程度(图 13)。插入到螺旋 *P*8 中的序列与核酶原的 5′“反义”区的自我互补程度可期望在新生转录本中优先于螺旋 *P*8 形成这种新螺旋，也可期望这一更替螺旋的形成破坏核酶的催化核心内侧面二级或许三级相互作用。因此，核酶原的误折会使它失去催化活性（图 14）。然而，核酶原与预期靶 *RNA* 的碱基配对可替代 *P*8“反义”碱基配对，隔绝该“反义”序列并允许 *P*8 螺旋和活性催化区的重新形成。*P*8“反义”螺旋的置换可产生更大数目的碱基对和允许催化区的适当折叠，因此应是很受欢迎的。

CAT—LacZ 核酶原

用如上所讨论的 *PCR* 诱变构建相应于两个 *CAT—LacZ* 核酶原的克隆序列，并通过体外转录产生 *RNA*。据观察 *CAT—LacZ* 转拼核酶在转录期间经历 3′拼接处的切割，这是由内含子序列催化的水解造成的。相似的水解见于未修饰的嗜热四膜虫内含子的体外转录。相比之下，不同的 *CAT—LacZ* 核酶原的转录本较稳定，在同样条件下几乎无明显的切割(图 15)。这表明核酶原是无活性的，若

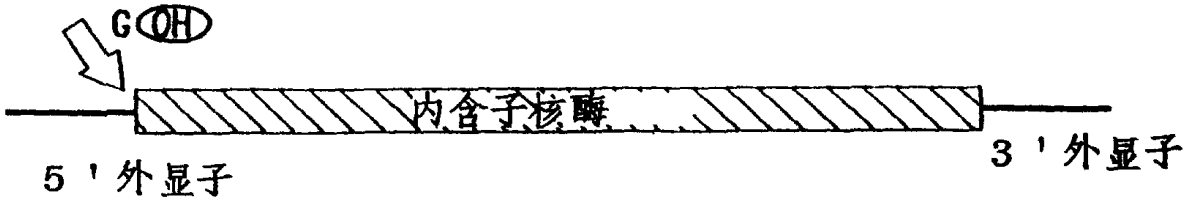
催化序列误折，便可期望核酶原失活。测试了平末端形式的核酶原针对*CAT RNA*的特异核糖核酸内切酶活性。由在*ScaI*位点被平切的模板转录(*AT—LacZ*核酶原，以删除3′拼接序列和*LacZ*序列。删除3′拼接位点后，核酶和核酶原都是稳定的。平末端的核酶原与*CAT RNA*一起保温导致了靶*RNA*的特异性切割，从而产生预期大小的片段(图16)。在37°，45°和50°下可看出特异的切割活性。

也构建了*GAI*4—*DTA*转拼核酶的核酶原形式（图17)。将含20个核苷酸的区域(与"反义"区互补)插入螺旋*P*8的5′股和环。这两种核酶原在修饰的螺旋*P*8中可能的碱基配对程度不同，*GAL*4—*DTA*核酶原1具有较长的干，在环中有较少(3个)可及的碱基。*GAL*4—*DTA*核酶原2的螺旋*P*8更类似于*CAT—LacZ*核酶原2的螺旋*P*8，具有较大的环（14个碱基），环中含有与"反义"区互补的序列。*GAL*4—*DTA*核酶原的转录本比未修饰的核酶的转录本稳定。特别是，核酶原2在可导致核酶形式的基本完全的自我切割的条件下保温（50℃下30分钟，10*mM* $MgCl_2$，2*mM GTP*，见图18)之后大体上是完整的。

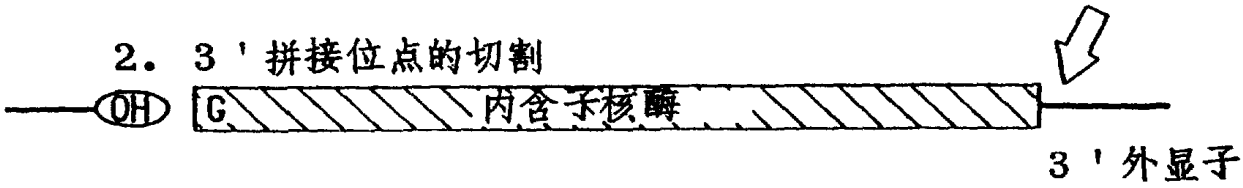
上面已对本发明作了充分的描述，本领域的技术人员在不影响本发明或其任何实施方案的精神和范围的情况下，可在很宽和等同的条件，参数等范围内完成本发明。

说　明　书　附　图

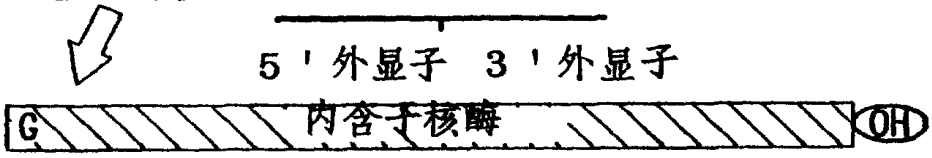
1．鸟苷介导的5＇拼接位点的切割



2．3＇拼接位点的切割



3．外显子片段的连接



4．内含子的环化

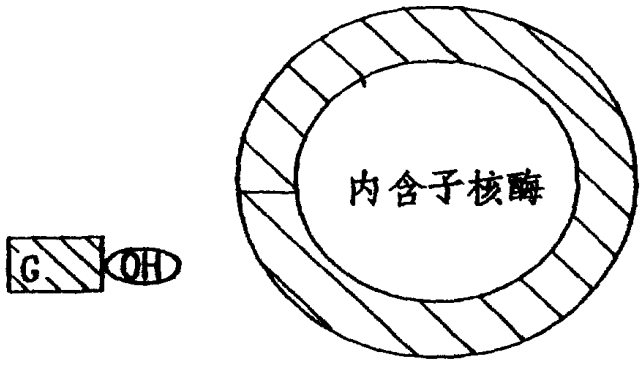
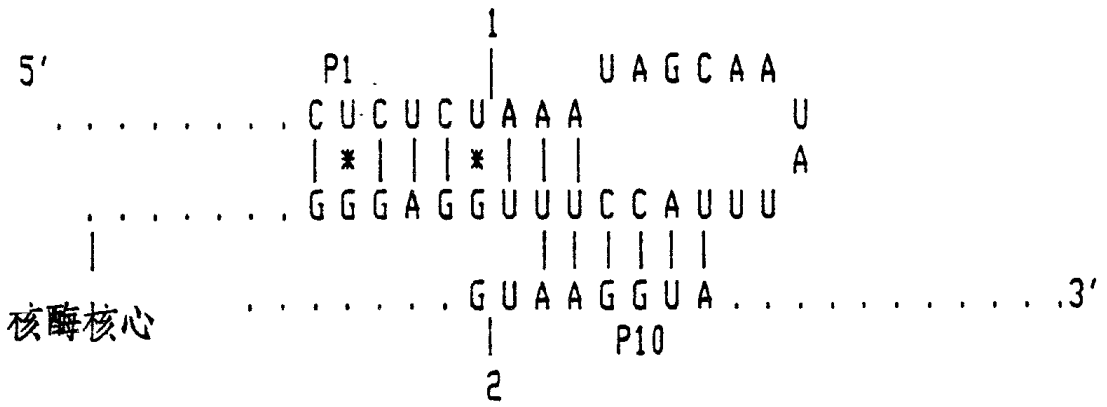
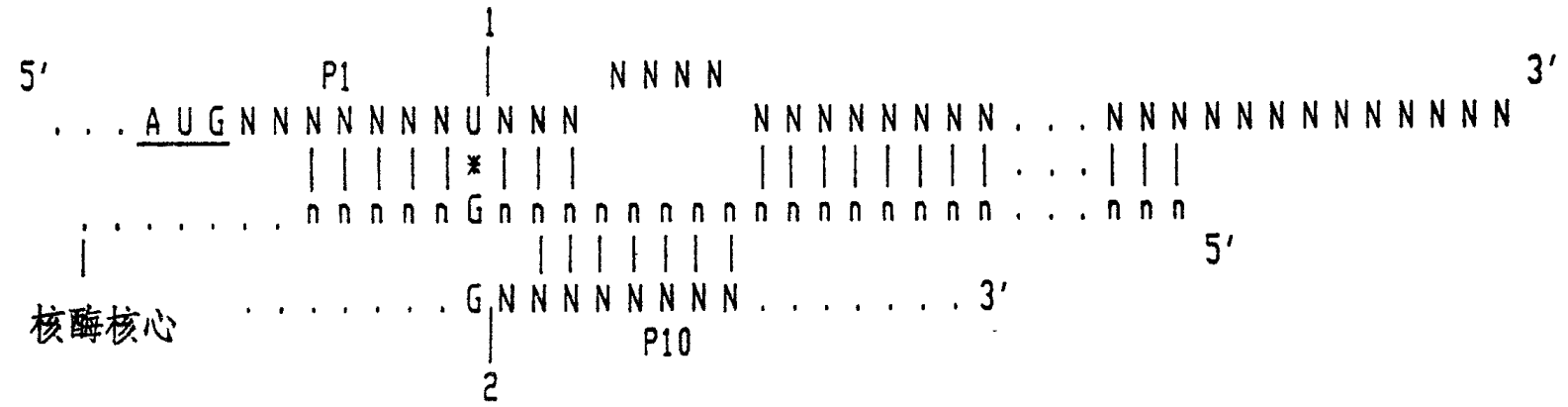


图.1



```
                                1
5'                P1            |     U A G C A A
  . . . . . . . . . C U C U C U A A A           U
                  | * | | | * | | |           A
    . . . . . . . G G G A G G U U U C C A U U U
    |                           | | | | | |
核酶核心      . . . . . . . G U A A G G U A . . . . . . . . . . . . 3'
                                |       P10
                                2
```

图.2A



```
                                1
5'                P1            |     N N N N                                                  3'
  . . . A U G N N N N N N N U N N N         N N N N N N N N . . . N N N N N N N N N N N N N N
                  | | | | | * | | |         | | | | | | | | . . . | | |
    . . . . . . . n n n n n G n n n n n n n n n n n n n n n n . . . n n n
    |                           | | | | | | |                                  5'
核酶核心      . . . . . . . G N N N N N N N N . . . . . . . 3'
                                |       P10
                                2
```

图.2B

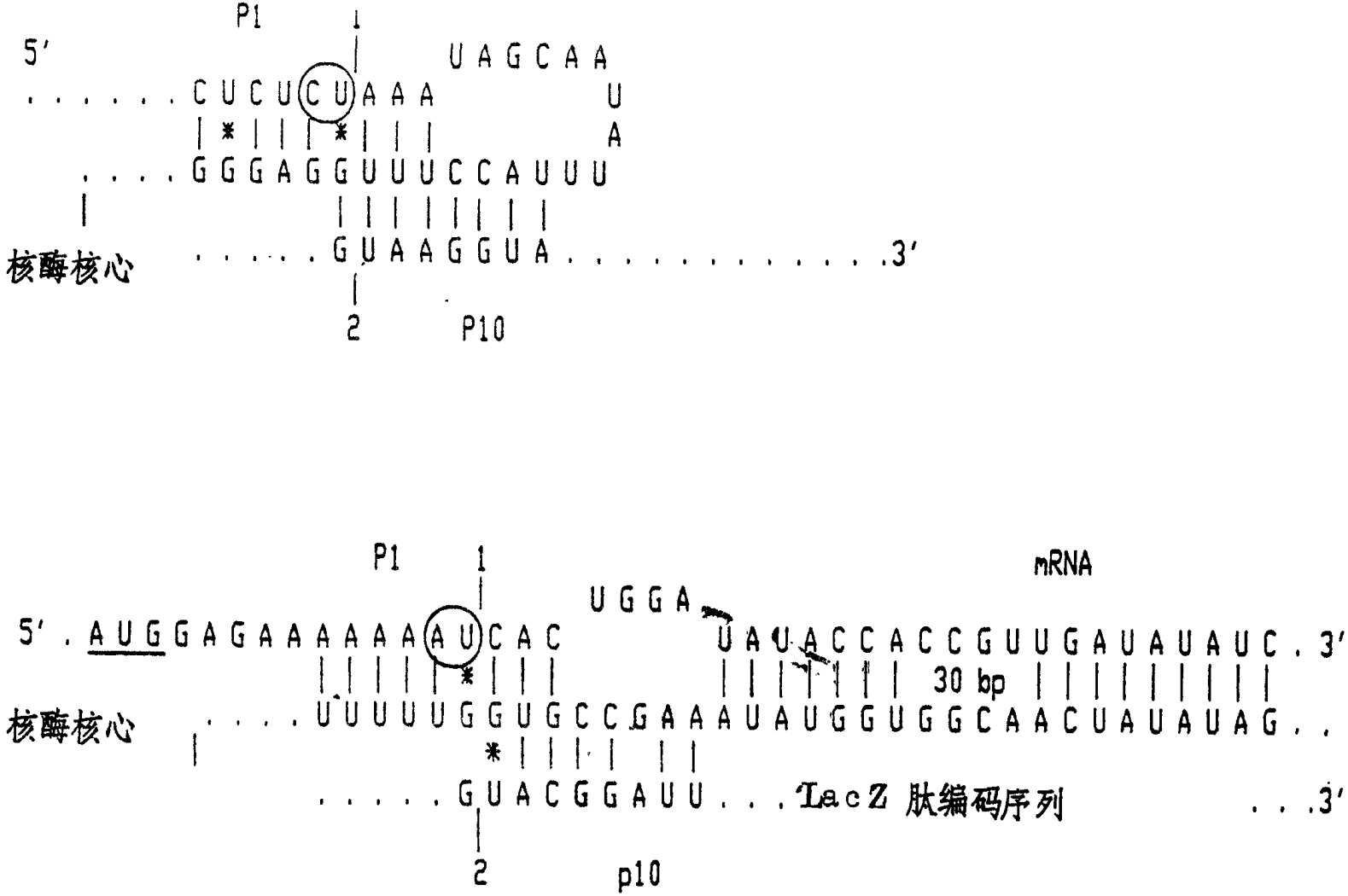

P1
1
5′
U A G C A A
. . . . . C U C U C U A A A U
. . . . G G G A G G U U U C C A U U U
核酶核心 . . . . . G U A A G G U A . . . . . .3′
2
P10
P1
1
mRNA
U G G A
5′ . A U G G A G A A A A A A A U C A C U A U A C C A C C G U U G A U A U A U C . 3′
30 bp
核酶核心 . . . . U U U U U G G U G C C G A A A U A U G G U G G C A A C U A U A U A G . .
. . . . . G U A C G G A U U . . . LacZ 肽编码序列 . . .3′
2
p10
met glu lys lys ile thr asp ser leu ala val val leu gln arg arg asp.....
5′ . AUG GAG AAA AAA AUU ACG GAU UCA CUG GCC GUC GUU UUA CAA CGU CGU GAC.....

图.3A

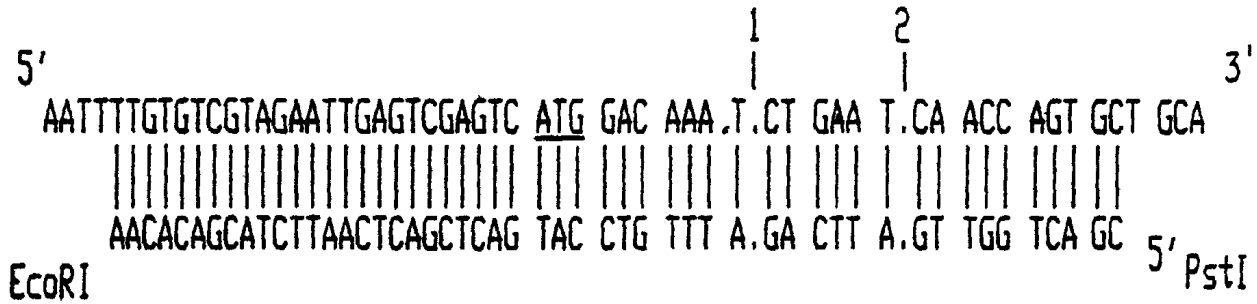
```
        |   10      |   20      |   30      |   40      |   50      |   60
  1 GGGAGACCGG AAGCTTCTTT ACGATGCCAT  TGGGATATAT CAACGGTGGT ATAAAGCCGT 60
 61 GGTTTTTAAA AGTTATCAGG CATGCACCTG  GTAGCTAGTC TTTAAACCAA TAGATTGCAT 120
121 CGGTTTAAAA GGCAAGACCG TCAAATTGCG  GGAAAGGGGT CAACAGCCGT TCAGTACCAA 180
181 GTCTCAGGGG AAACTTTGAG ATGGCCTTGC  AAAGGGTATG GTAATAAGCT GACGGACATG 240
241 GTCCTAACCA CGCAGCCAAG TCCTAAGTCA  ACAGATCTTC TGTTGATATG GATGCAGTTC 300
301 ACAGACTAAA TGTCGGTCGG GGAAGATGTA  TTCTTCTCAT AAGATATAGT CGGACCTCTC 360
361 CTTAATGGGA GCTAGCGGAT GAAGTGATGC  AACACTGGAG CCGCTGGGAA CTAATTTGTA 420
421 TGCGAAAGTA TATTGATTAG TTTTGGAGTA  CTCGTACGGA TTCACTGGCC GTCGTTTTAC 480
481 AACGTCGTGA CTGGGAAAAC CCTGGCGTTA  CCCAACTTAA TCGCCTTGCA GCACATCCCC 540
541 CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG  CCCGCACCGA TCGCCCTTCC CAACAGTTGC 600
601 GCAGCCTGAA TGGAAATTGT AAG                                          623
        |   10      |   20      |   30      |   40      |   50      |   60
```

图.3B

CWV 分离物

```
                                                  1         2
                                                  |         |
WD      UUUGCGUCU.......UAGUGUGCCU AUG GAC AAA U.CU GGA U.CU CCC AAU GCU AGU
Q       UUUGCGUCUCAG.......UGUGCCU AUG GAC AAA U.CU GGA U.CU CCC AAU GCU AGU
Fny     UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
M       UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAU AAA U.CU GAA U.CA ACC AGU GCU GGU
I       UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
O       UUUACGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
Y       UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
D       UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
C       UUUGUGUCGUAGAAUUGAGUCGAGUC AUG GAC AAA U.CU GAA U.CA ACC AGU GCU GGU
     5'
```

图. 4A



```
                                                  1       2
5'                                                |       |                 3'
  AATTTTGTGTCGTAGAATTGAGTCGAGTC ATG GAC AAA.T.CT GAA T.CA ACC AGT GCT GCA
      ||||||||||||||||||||||||| ||| ||| ||| | || ||| | || ||| ||| ||
      AACACAGCATCTTAACTCAGCTCAG TAC CTG TTT A.GA CTT A.GT TGG TCA GC
EcoRI                                                                 5' PstI
```

图. 4B

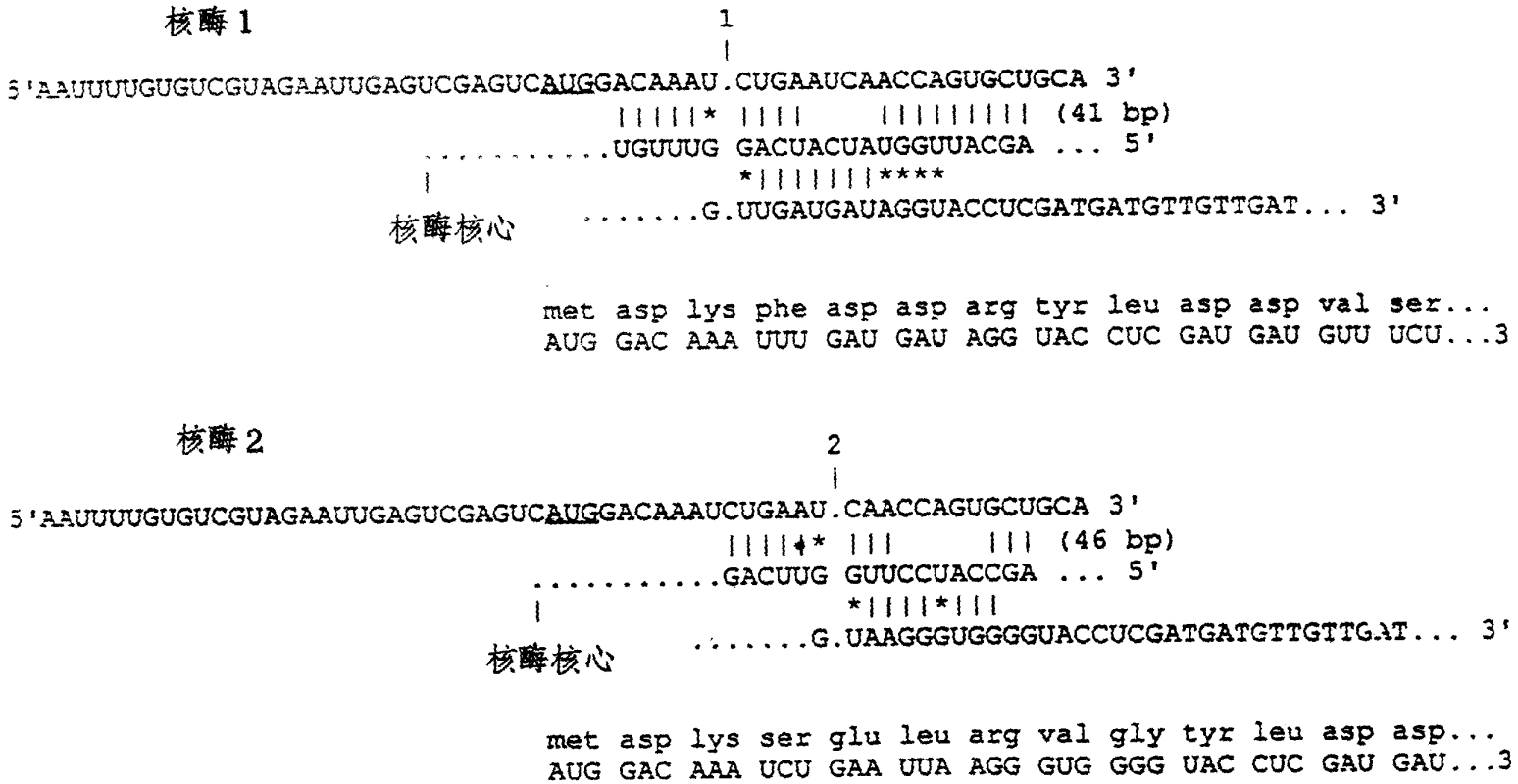

核酶 1
1
5'AAUUUUGUGUCGUAGAAUUGAGUCGAGUCAUGGACAAAU.CUGAAUCAACCAGUGCUGCA 3'
|||||* |||| |||||||| (41 bp)
.....UGUUUG GACUACUAUGGUUACGA ... 5'
*|||||||****
核酶核心 .....G.UUGAUGAUAGGUACCUCGATGATGTTGTTGAT... 3'
met asp lys phe asp asp arg tyr leu asp asp val ser...
AUG GAC AAA UUU GAU GAU AGG UAC CUC GAU GAU GUU UCU...3
核酶 2
2
5'AAUUUUGUGUCGUAGAAUUGAGUCGAGUCAUGGACAAAUCUGAAU.CAACCAGUGCUGCA 3'
|||||* ||| ||| (46 bp)
.....GACUUG GUUCCUACCGA ... 5'
|||||||
核酶核心 .....G.UAAGGGUGGGGUACCUCGATGATGTTGTTGAT... 3'
met asp lys ser glu leu arg val gly tyr leu asp asp...
AUG GAC AAA UCU GAA UUA AGG GUG GGG UAC CUC GAU GAU...3

图 4C

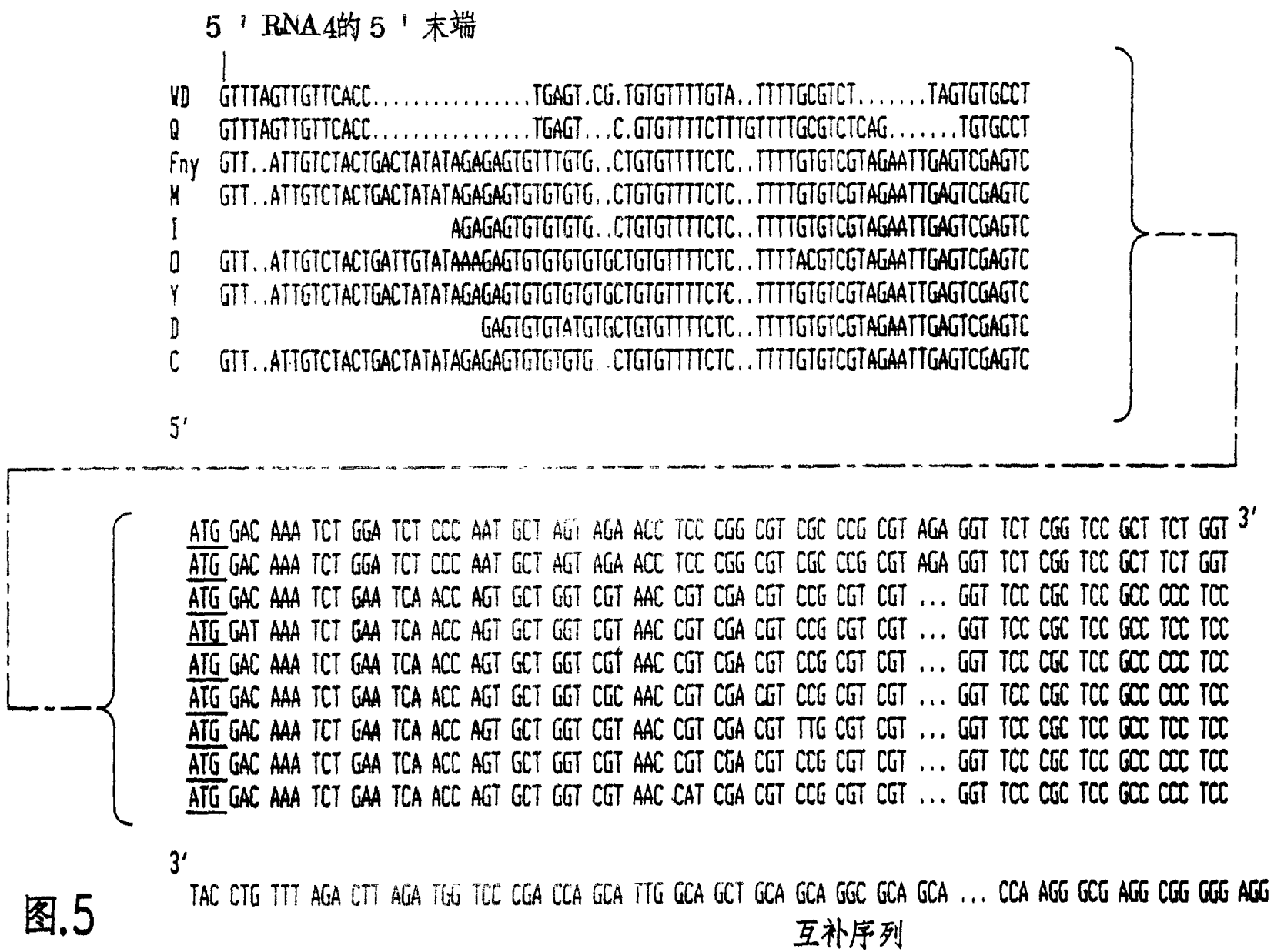
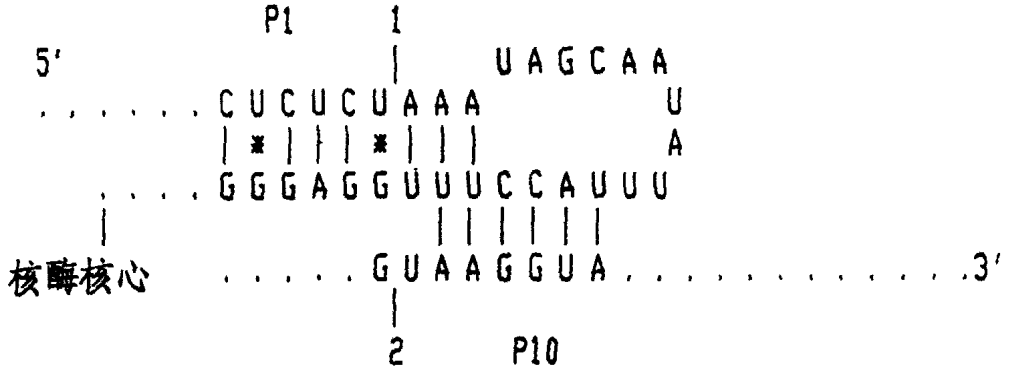

5′ RNA4的5′末端
VD GTTTAGTTGTTCACC.....TGAGT.CG.TGTGTTTTGTA..TTTTGCGTCT.....TAGTGTGCCT
Q GTTTAGTTGTTCACC.....TGAGT...C.GTGTTTTCTTTGTTTTGCGTCTCAG.....TGTGCCT
Fny GTT..ATTGTCTACTGACTATATAGAGAGTGTTTGTG..CTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
M GTT..ATTGTCTACTGACTATATAGAGAGTGTGTGTG..CTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
I AGAGAGTGTGTGTG..CTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
O GTT..ATTGTCTACTGATTGTATAAAGAGTGTGTGTGTGCTGTGTTTTCTC..TTTTACGTCGTAGAATTGAGTCGAGTC
Y GTT..ATTGTCTACTGACTATATAGAGAGTGTGTGTGTGCTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
D GAGTGTGTATGTGCTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
C GTT..ATTGTCTACTGACTATATAGAGAGTGTGTGTG..CTGTGTTTTCTC..TTTTGTGTCGTAGAATTGAGTCGAGTC
5′
ATG GAC AAA TCT GGA TCT CCC AAT GCT AGT AGA ACC TCC CGG CGT CGC CCG CGT AGA GGT TCT CGG TCC GCT TCT GGT 3′
ATG GAC AAA TCT GGA TCT CCC AAT GCT AGT AGA ACC TCC CGG CGT CGC CCG CGT AGA GGT TCT CGG TCC GCT TCT GGT
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CGT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC CCC TCC
ATG GAT AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CGT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC TCC TCC
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CGT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC CCC TCC
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGC AAC CGT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC CCC TCC
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CGT CGA CGT TTG CGT CGT ... GGT TCC CGC TCC GCC TCC TCC
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CGT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC CCC TCC
ATG GAC AAA TCT GAA TCA ACC AGT GCT GGT CGT AAC CAT CGA CGT CCG CGT CGT ... GGT TCC CGC TCC GCC CCC TCC
3′
TAC CTG TTT AGA CTT AGA TGG TCC CGA CCA GCA TTG GCA GCT GCA GCA GGC GCA GCA ... CCA AGG GCG AGG CGG GGG AGG
互补序列

图.5



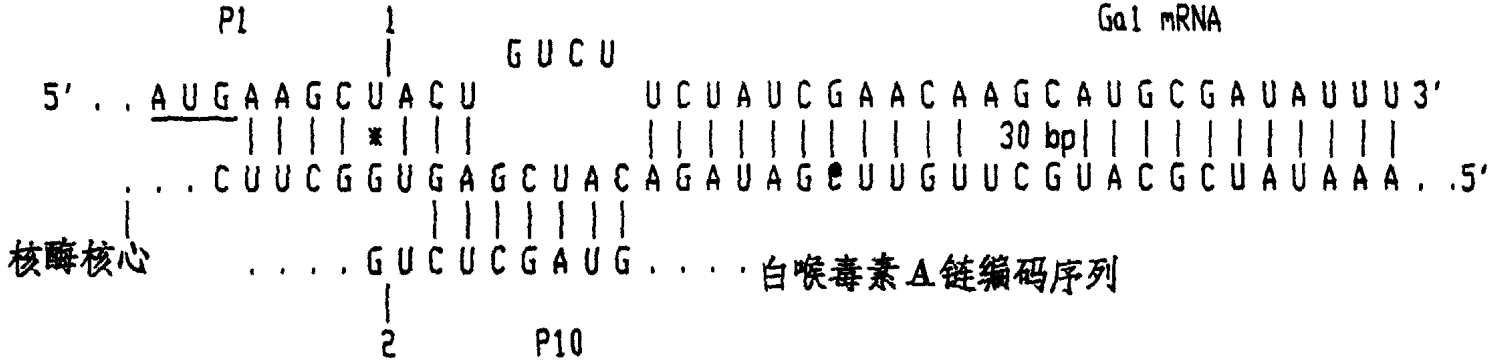


图.6A

met lys leu leu asp asp val val asp ser ser lys ser phe val met glu asn phe ser .....

5'. . AUG AAG CUU CUC GAU GAU GUU GUU GAU UCU UCU AAA UCU UUU GTG AUG GAA AAC UUU UCU .....3'

```
      |   10      |   20      |   30      |   40      |   50      |   60
   1 GTCGACCTTT TTAAGTCGGC AAATATCGCA TGTTTGTTCG ATAGACATCG AGTGGCTTCA 60
  61 AAAGTTATCA GGCATGCACC TGGTAGCTAG TCTTTAAACC AATAGATTGC ATCGGTTTAA 120
 121 AAGGCAAGAC CGTCAAATTG CGGGAAAGGG GTCAACAGCC GTTCAGTACC AAGTCTCAGG 180
 181 GGAAACTTTG AGATGGCCTT GCAAAGGGTA TGGTAATAAG CTGACGGACA TGGTCCTAAC 240
 241 CACGCAGCCA AGTCCTAAGT CAACAGATCT TCTGTTGATA TGGATGCAGT TCACAGACTA 300
 301 AATGTCGGTC GGGGAAGATG TATTCTTCTC ATAAGATATA GTCGGACCTC TCCTTAATGG 360
 361 GAGCTAGCGG ATGAAGTGAT GCAACACTGG AGCCGCTGGG AACTAATTTG TATGCGAAAG 420
 421 TATATTGATT AGTTTTGGAG TACTCGTCTC GATGATGTTG TTGATTCTTC TAAATCTTTT 480
 481 GTGATTGAAA ACTTTTCTTC GTACCACGGG ACTAAACCTG GTTATGTAGA TTCCATTCAA 540
 541 AAAGGTATAC AAAAGCCAAA ATCTGGTACA CAAGGAAATT ATGACGATGA TTGGAAAGGG 600
 601 TTTTATAGTA CCGACAATAA ATACGACGCT GCGGGATACT CTGTAGATAA TGAAAACCCG 660
 661 CTCTCTGGAA AAGCTGGAGG CGTGGTCAAA GTGACGTATC CAGGACTGAC GAAGGTTCTC 720
 721 GCACTAAAAG TGGATAATGC CGAAACTATT AAGAAAGAGT TAGGTTTAAG TCTCACTGAA 780
 781 CCGTTGATGG AGCAAGTCGG AACGGAAGAG TTTATCAAAA GGTTCGGTGA TGGTGCTTCG 840
 841 CGTGTAGTGC TCAGCCTTCC CTTCGCTGAG GGGAGTTCTA GCGTTGAATA TATTAATAAC 900
 901 TGGGAACAGG CGAAAGCGTT AAGCGTAGAA CTTGAGATTA ATTTTGAAAC CCGTGGAAAA 960
 961 CGTGGCCAAG ATGCGATGTA TGAGTATATG GCTCAAGCCT GTGCAGGAAA TCGTGTCAGG 1020
1021 CGATCTTTGT GACTCGAG                                              1038
      |   10      |   20      |   30      |   40      |   50      |   60
```

图. 6B

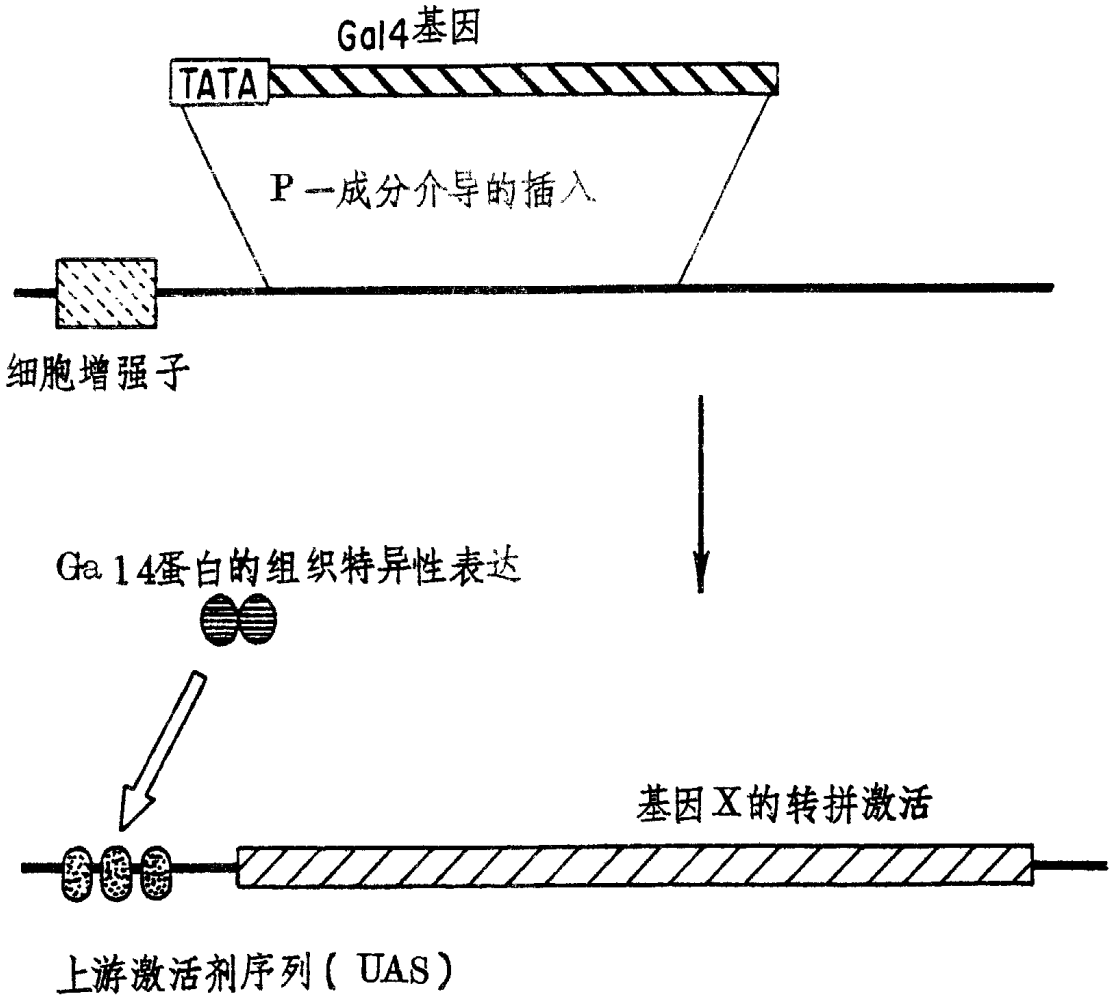

Gal4基因
TATA
P一成分介导的插入
细胞增强子
Ga l4蛋白的组织特异性表达
基因X的转拼激活
上游激活剂序列（ UAS）

图.7

转拼的Gal 4—DTA mRNA的转译

5′ ↓ 3

ATG AAG CTT CTC GAT GAT GTT GTT GAT TCT TCT AAA TCT TTT GTG ATG GAA AAC TTT TCT TCG TAC CAC GGG ACT AAA ... ...

met lys leu leu asp asp val val asp ser ser lys ser phe val met glu asn phe ser ser tyr his gly thr lys

3′“外显子”的转译

Rz-DTA$_{met}$

5′ 3′

TCTCGATGATGTTGTTGATTCTTCTAAATCTTTTGTG ATG GAA AAC TTT TCT TCG TAC CAC GGG ACT AAA ... ...

met glu asn phe ser ser tyr his gly thr lys

Rz-DTA$_{ile}$

5′ 3′

TCTCGATGATGTTGTTGATTCTTCTAAATCTTTTGTG ATT GAA AAC TTT TCT TCG TAC CAC GGG ACT AAA ... ...

ile

Rz-DTA$_{leu}$

5′ 3′

TCTCGATGATGTTGTTGATTCTTCTAAATCTTTTGTG TTG GAA AAC TTT TCT TCG TAC CAC GGG ACT AAA ... ...

leu

图.8

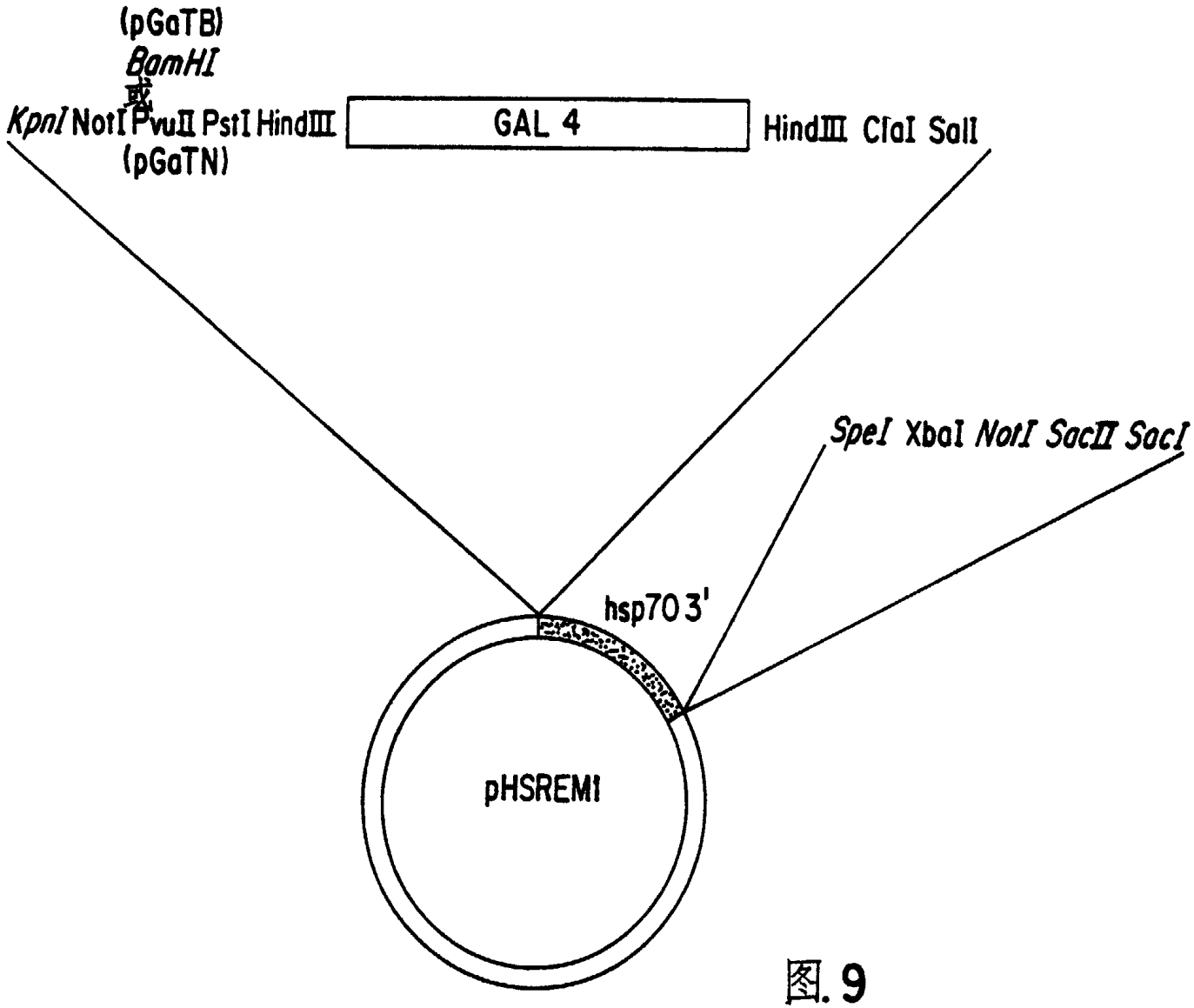

(pGaTB)
BamHI
或
KpnI NotI PvuII PstI HindIII
(pGaTN)
GAL 4
HindIII ClaI SalI
SpeI XbaI NotI SacII SacI
hsp70 3'
pHSREM1

图. 9

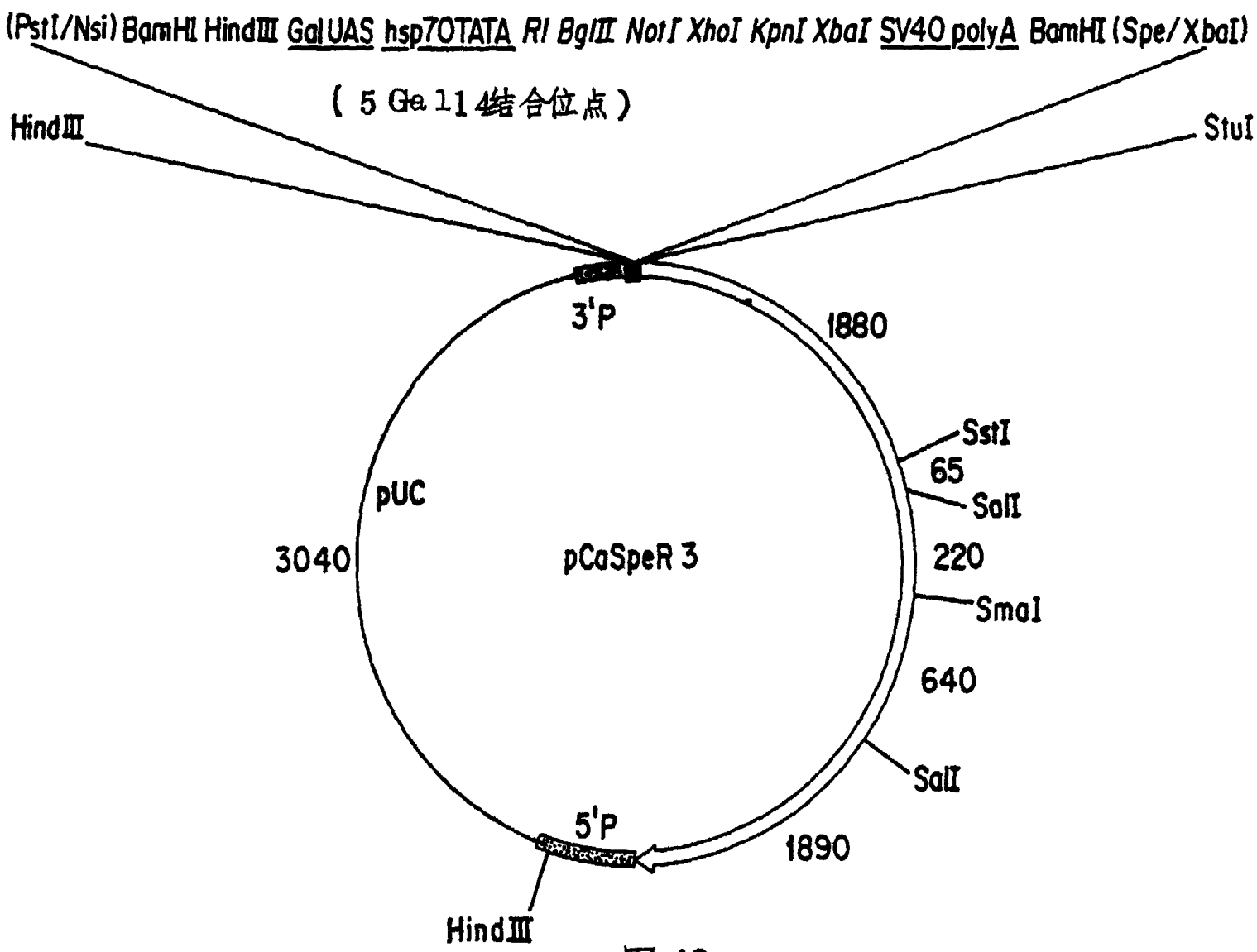

(PstI/NsiI) BamHI HindIII GalUAS hsp70TATA RI BglII NotI XhoI KpnI XbaI SV40 polyA BamHI (SpeI/XbaI)
(5 Gal 14结合位点)
HindIII
StuI
3'P
1880
SstI
65
SalI
220
SmaI
640
SalI
1890
5'P
HindIII
pUC
3040
pCaSpeR 3

图. 10

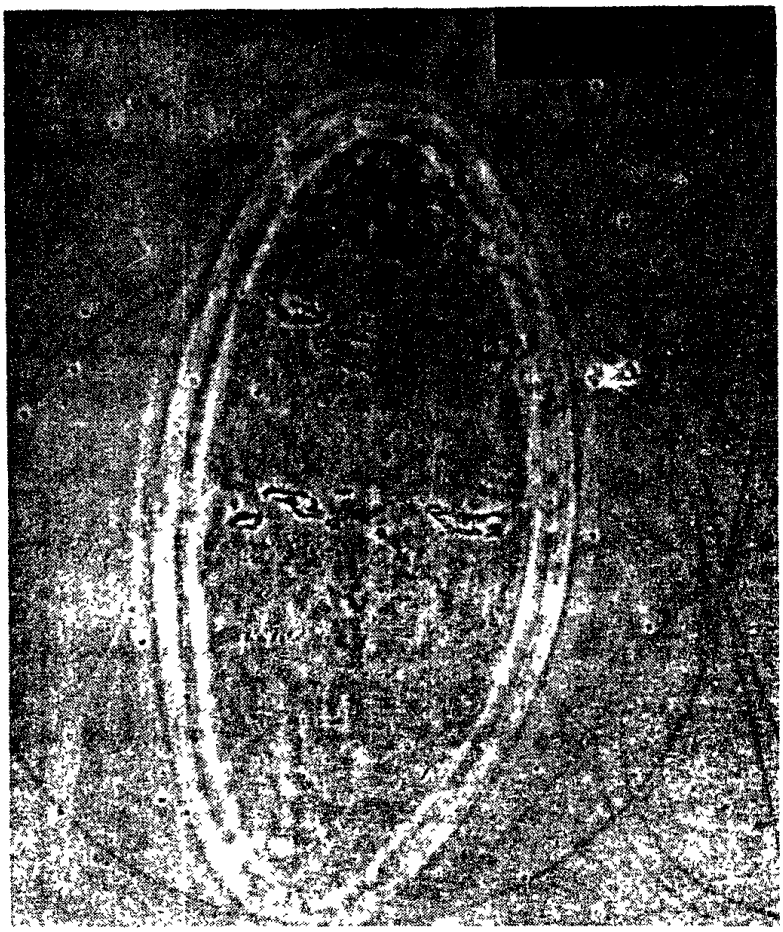


图. 11

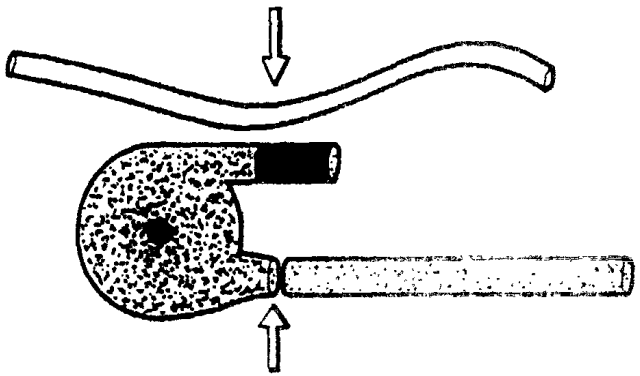


图. 12A

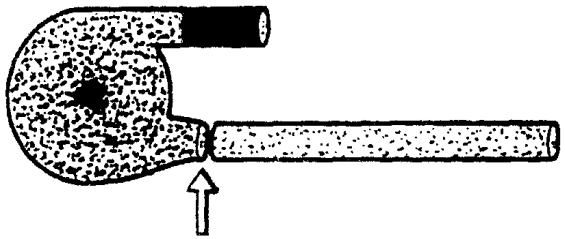


图. 12B

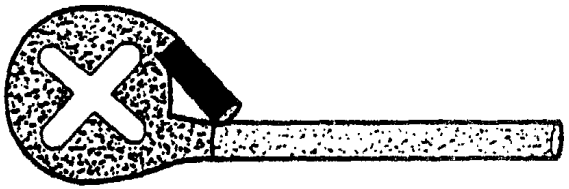


图. 12C

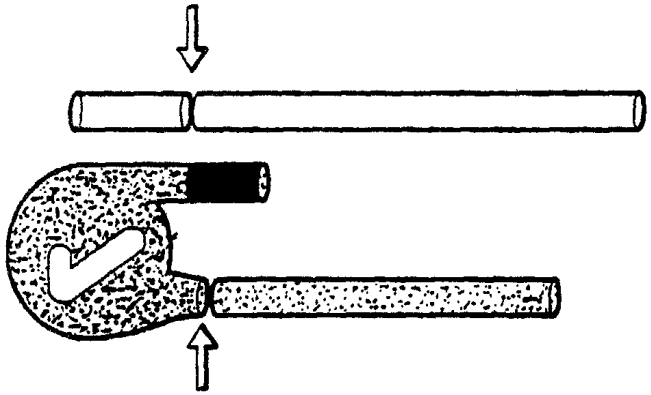


图. 12D

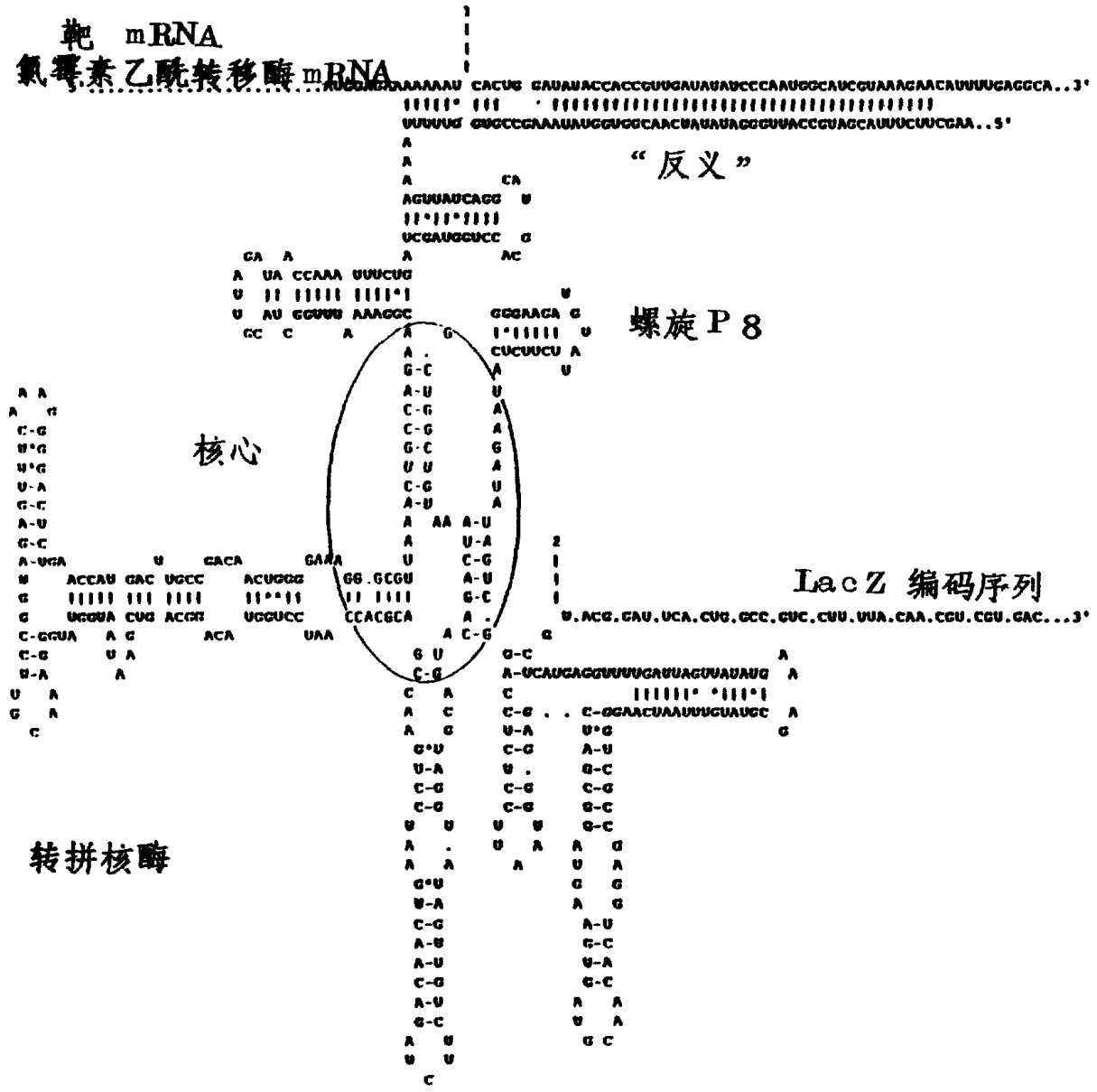
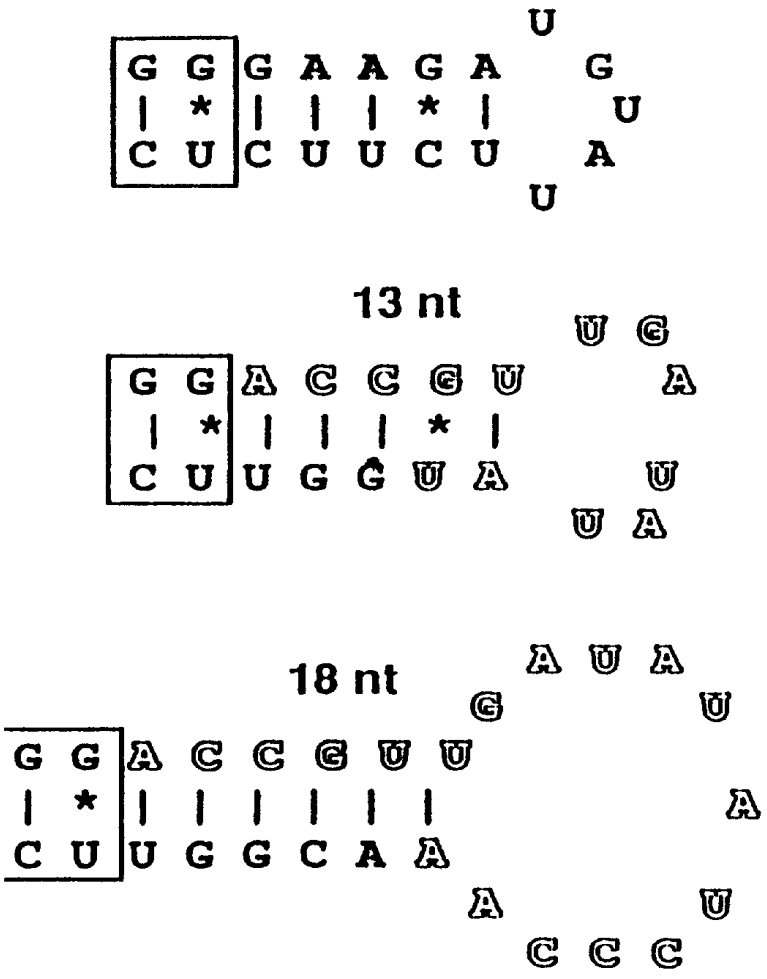


图. 13A


U
G G G A A G A G
| * | | | * | U
C U C U U C U A
U
13 nt
U G
G G A C C G U A
| * | | | * |
C U U G G U A U
U A
A U A
18 nt
G U
G G A C C G U U
| * | | | | | | A
C U U G G C A A
A U
C C C

N＝与“反义”互补

图. 13B

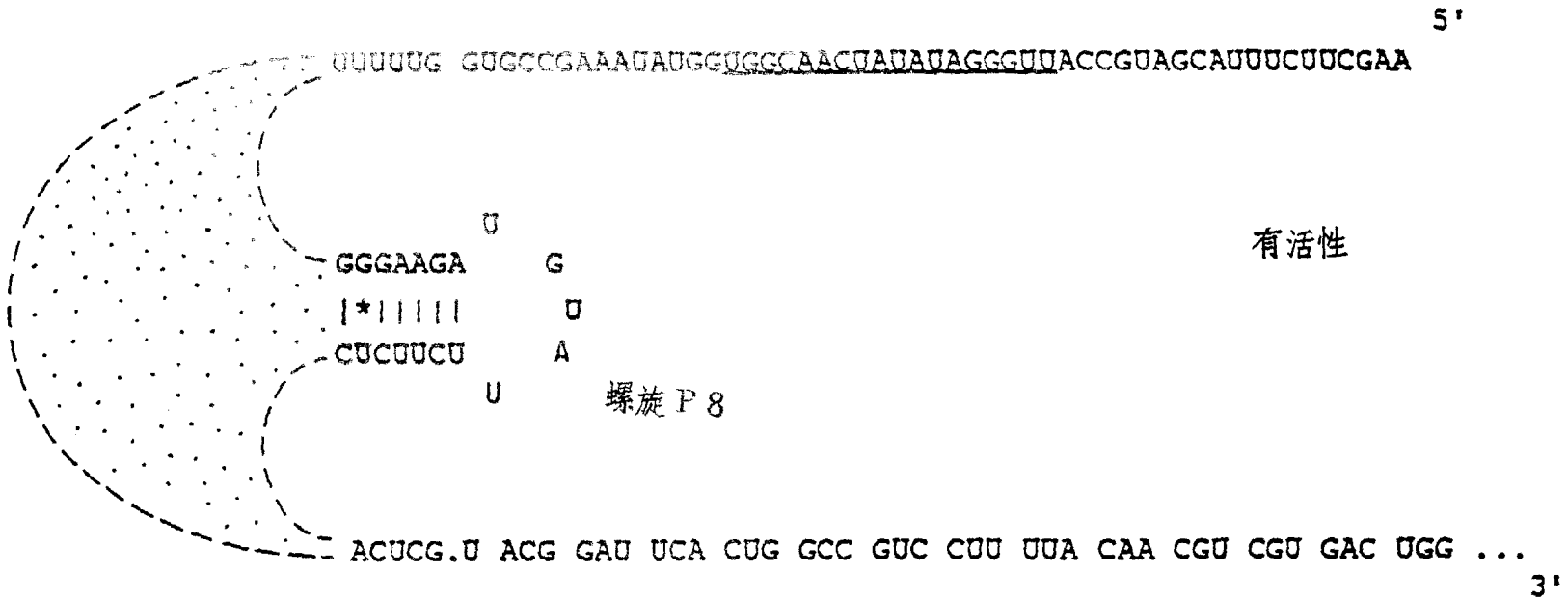

5'
UUUUUG GUGCCGAAAUAUGGUGGCAACUAUAUAGGGUUACCGUAGCAUUUCUUCGAA
U
GGGAAGA G
|*|||||| U
CUCUUCU A
U
螺旋P8
有活性
ACUCG.U ACG GAU UCA CUG GCC GUC CUU UUA CAA CGU CGU GAC UGG ...
3'

图. 14A

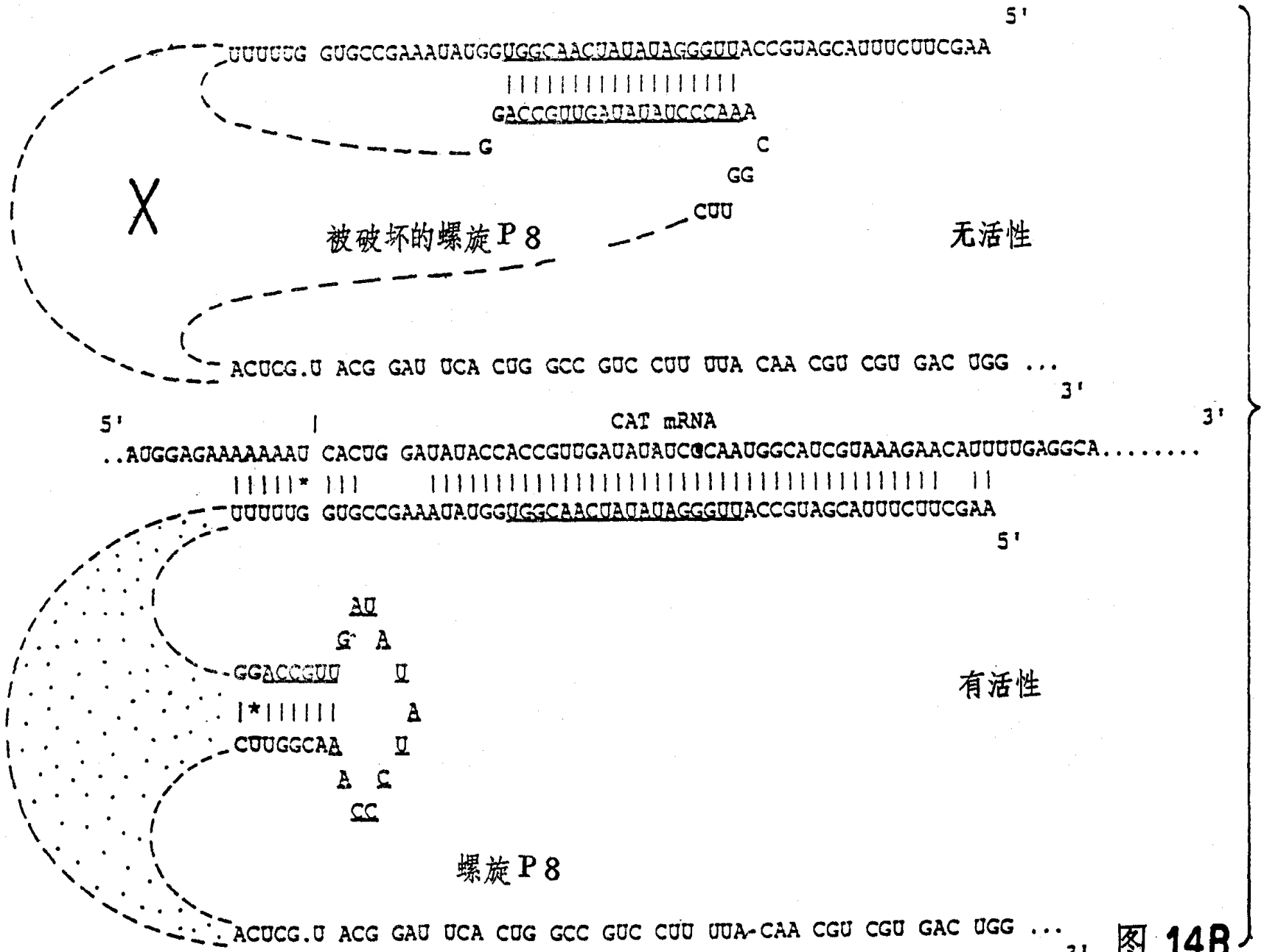

5'
UUUUUG GUGCCGAAAUAUGGUGGCAACUAUAUAGGGUUACCGUAGCAUUUCUUCGAA
GACCGUUGAUAUAUCCCAAA
G
C
GG
CUU
X
被破坏的螺旋P8
无活性
ACUCG.U ACG GAU UCA CUG GCC GUC CUU UUA CAA CGU CGU GAC UGG ...
3'
5'
CAT mRNA
3'
..AUGGAGAAAAAAAU CACUG GAUAUACCACCGUUGAUAUAUCGCAAUGGCAUCGUAAAGAACAUUUUGAGGCA.....
UUUUUG GUGCCGAAAUAUGGUGGCAACUAUAUAGGGUUACCGUAGCAUUUCUUCGAA
5'
AU
G A
GGACCGUU U
A
CUUGGCAA U
A C
CC
螺旋P8
ACUCG.U ACG GAU UCA CUG GCC GUC CUU UUA-CAA CGU CGU GAC UGG ...
3'

图.14B

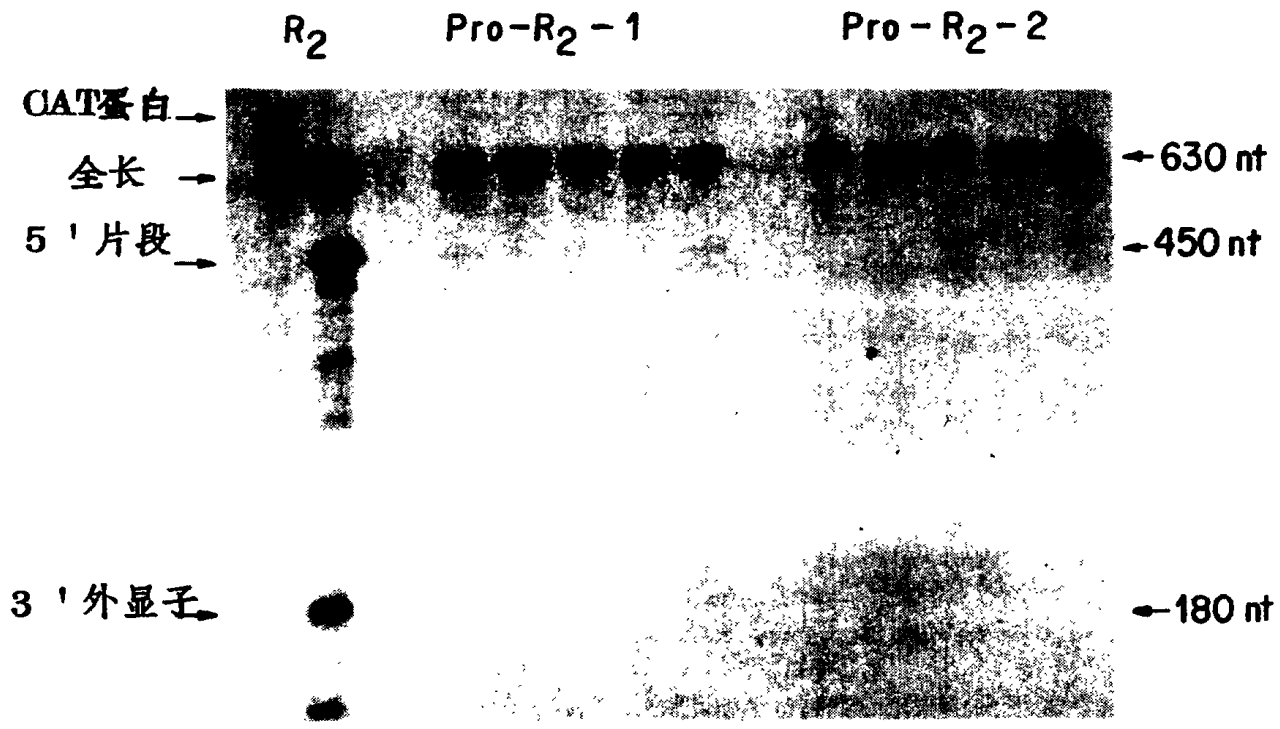


图.15

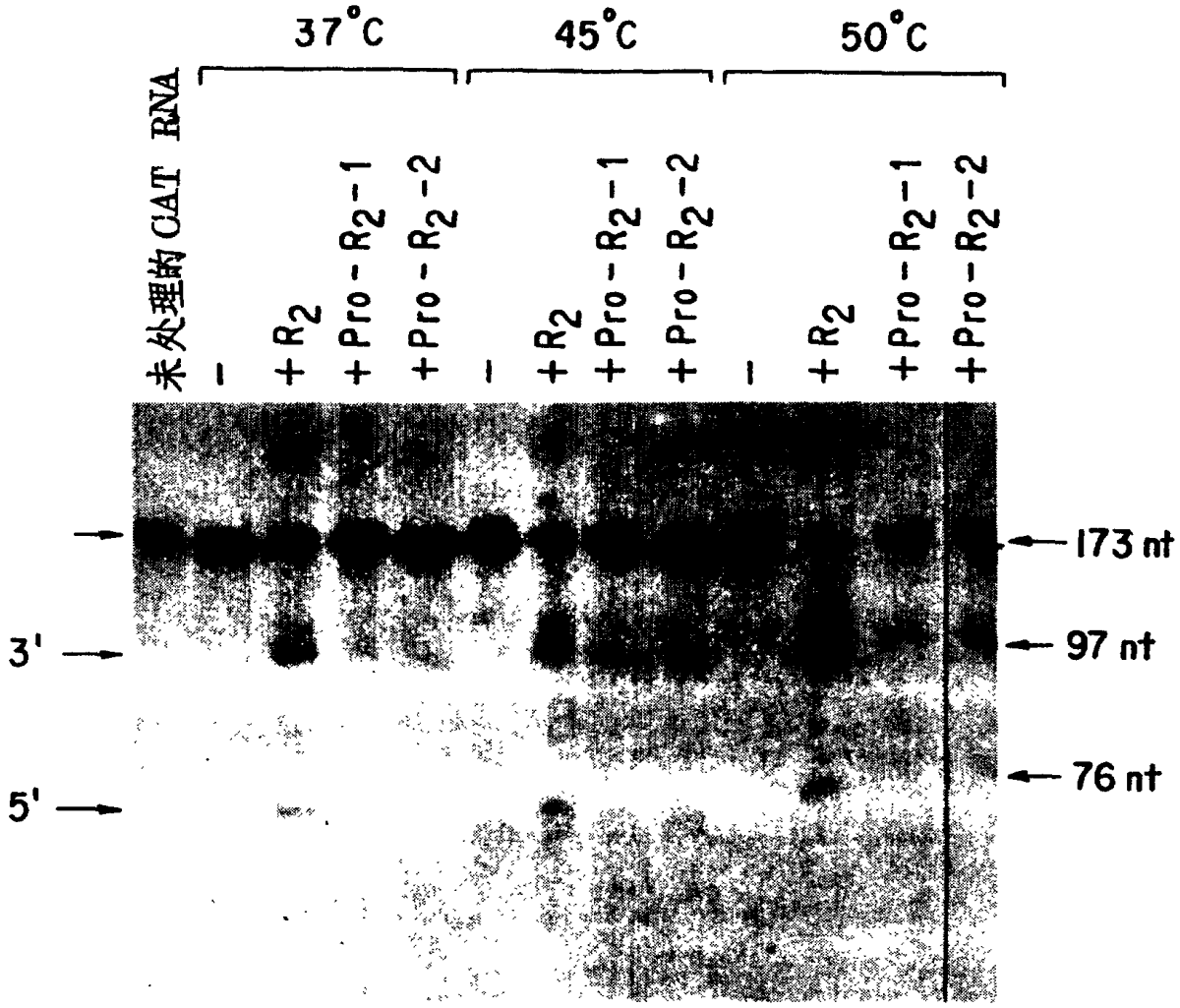


图. 16

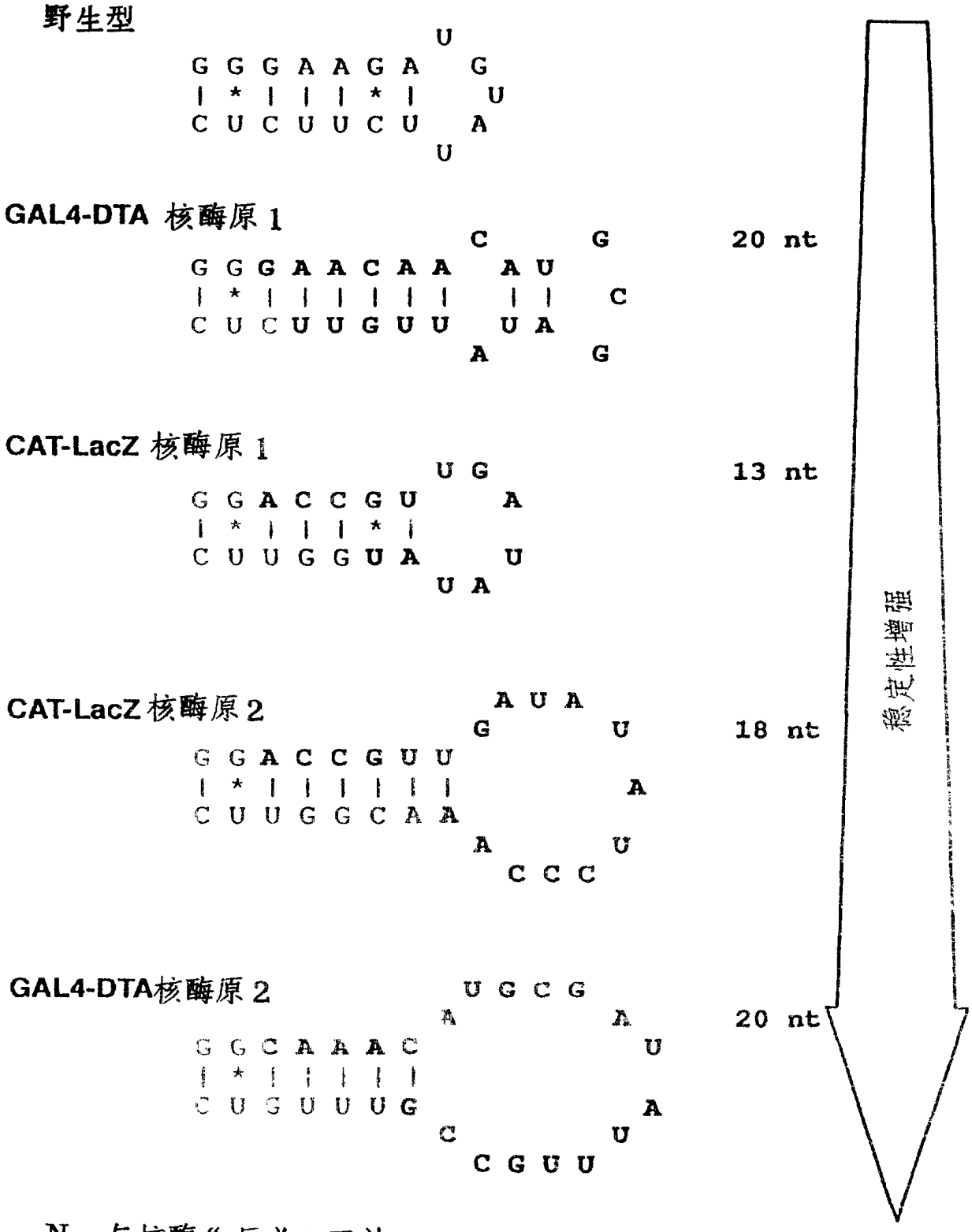

野生型
GAL4-DTA 核酶原 1
20 nt
CAT-LacZ 核酶原 1
13 nt
CAT-LacZ 核酶原 2
18 nt
GAL4-DTA核酶原 2
20 nt
靶定基质
N＝与核酶“反义”互补

图. 17

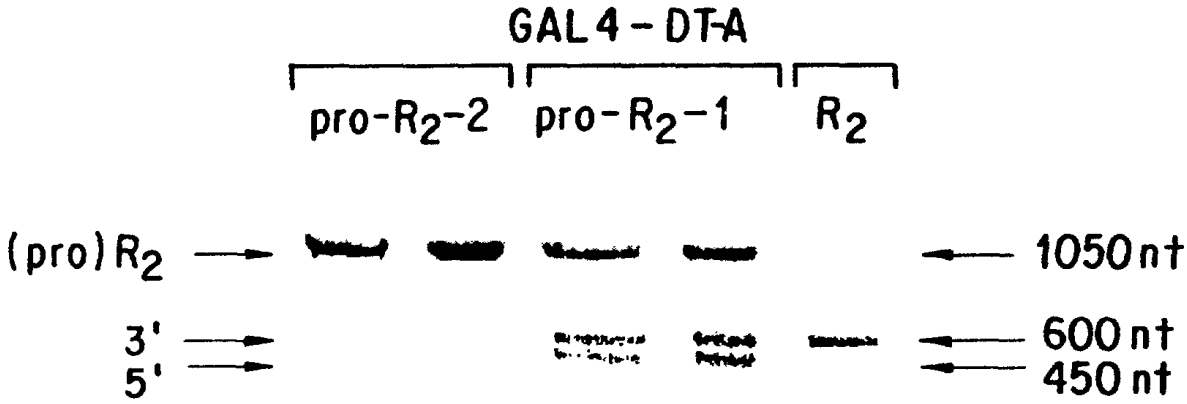


图. 18