

RNA 自我剪接机制

——生物体中无酶催化的奇特反应

文 重 乔 英

(北京大学化学系)

生物体内的一切反应都是在酶作用下进行的。今天对于这一观念,出现了一个反例,1982 年 Thoms R. Cech 和他的同事在 Colorado 大学的一个发现^[1]: 原生动物四膜虫 (*Tetrahymena*) 的核糖体 RNA 前体的编码基因插入序列的剪接无需其它酶类的存在就可以完成。这一发现的意义在于它改变了生物催化的一贯观

目的显然是要获得更适合工业生产条件的、稳定的酶制剂。

随着蛋白质工程的进一步开展,将有更多的酶和蛋白质,从各个不同的方面运用这一技术进行研究,我们对酶的结构与功能的关系就会有更多、更深入的了解,从而有可能从中总结出一些带规律性的东西,进而制定出一套用蛋白质工程改造酶的结构与性质的详细的规则。到那时,人们就可以按照不同的要求来定做各种具有特定性质的新的酶了。

念——具有蛋白质本质的酶促反应。这一发现同时还对生命进化的某些假说以支持,给人们以新的启示。California 大学的 Harry Noller 把 Cech 的工作称作是“十年来最令人激动的发现之一”。

四膜虫嗜热菌大核 rRNA 的自动剪接过程是这样的^[2]: 一个有 413 个 bp (碱基对, base

75, 4268, 1978.

参 考 文 献

- [1] Ulmer, K. M.: *Biotechnology and Biological Frontiers* (ed. by P. H. Abelson) p. 100, The American Association for the Advancement of Science, 1984.
- [2] Smith, M.: *Trends in Biochemical Science*, 440, 1982.
- [3] Zoller, M. J. et al.: *Methods in Enzymol.*, 100, 468, 1983.
- [4] Lederberg, J.: *Science*, 131, 269, 1960.
- [5] Schott, H. et al.: *J. Amer. Chem. Soc.*, 95, 3378, 1973.
- [6] Sanger, F. et al.: *J. Mol. Biol.*, 125, 225, 1978.
- [7] Hutchinson, C. A. et al.: *J. Biol. Chem.*, 253, 6551, 1978.
- [8] Razin, A. et al.: *Proc. Natl. Acad. Sci. U. S. A.*, 75, 4268, 1978.
- [9] Winter, G. et al.: *Nature*, 299, 756, 1982.
- [10] Sigal, I. S. et al.: *Proc. Natl. Acad. Sci. USA*, 79, 7175, 1982.
- [11] Dalbadie-McFarland, G. et al.: *Proc. Natl. Acad. Sci. USA*, 79, 6409, 1982.
- [12] Osinga, K. A. et al.: *Nucleic Acids Res.*, 11, 8595, 1983.
- [13] Carter, P. et al.: *Oligonucleotide Site-directed Mutagenesis in M13*, p. 8, Anglian Biotechnology Ltd. UK, 1984.
- [14] Zoller, M. J. et al.: *DNA*, 3, 479, 1984.
- [15] Carter, P. et al.: *Nucleic Acids, Res.*, 13, 4431, 1985.
- [16] Kramer, W. et al.: *Nucleic Acids Res.*, 12, 441, 1984.
- [17] Schold, M. et al.: *DNA*, 3, 469, 1984.
- [18] Fersht, A. R. et al.: *Angew. Chem. Inter. Ed. Engl.*, 23, 467, 1984.
- [19] Fersht, A. R. et al.: *Nature*, 314, 253, 1985.
- [20] Wells, T. N. C. et al.: *Nature*, 316, 656, 1985.
- [21] Wilkinson, A. J. et al.: *Nature*, 307, 187, 1984.
- [22] Carter, P. et al.: *Cell*, 38, 835, 1984.
- [23] Winter, G. et al.: *Biochemical Soc. Transactions*, 12, 224, 1984.
- [24] Craik, C. S. et al.: *Science* (in press).
- [25] Sigal, J. S. et al.: *J. Biol. Chem.*, 259, 5327, 1984.
- [26] Robey, E. A. et al.: *J. Biol. Chem.*, 259, 11180, 1984.

[本文于 1985 年 11 月 23 日收到]

pair) 的插入序列 IVS(intervening sequence) 或称内含子 (intron) 的前体 rRNA, 经纯化, 去蛋白后与二价正离子 ($MgCl_2$) 及一个鸟苷因子一起保温, 核糖 RNA 前体就会发生剪接, 剪接反应的特点是: 鸟苷向 IVS 的 5' 端的亲核进攻, IVS 被准确地剪切下, 两段外显子 (exon 编码序列) 接在一起, 而 IVS 则共价环合 (图 1)。

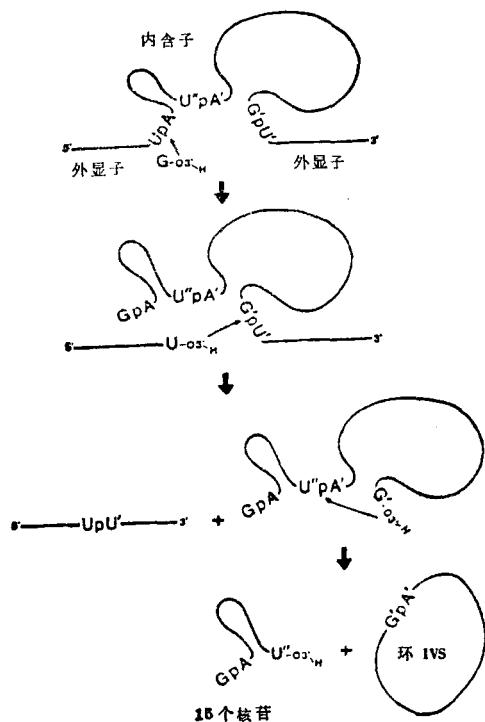


图 1 大核 rRNA 的剪接反应^[4]

这一反应是在无酶, 无其它外来能量 (如 ATP 或 GTP) 下发生的。反应的奇特之处在于四膜虫 rRNA 本身是引起反应的内在根源, 正因 rRNA 具有这样一个特点, Cech 和其同事为它起了一个 “ribozyme” 的名字。本文所介绍的是四膜虫 rRNA 剪接的机制及自动催化对结构的要求。

一、rRNA 在反应部位的结构

如图 1 所示, 这是一个转磷酸酯反应, 每一个 P—O 键的断裂, 都伴随着一个新的 P—O 键的形成, 所以从热力学角度考虑, 自由能的变化

不大, 但键断裂反应总需要活化能, 这一能量是怎样获得的? rRNA 分子的反应活性部位在哪里? 对于这个无需蛋白质, 无需外来能量的反应, 回答只能是: 这一切取决于 rRNA 的特殊的二级和三级结构。事实也正是如此, 四膜虫 rRNA 在高温下或高浓度的甲酰胺, 尿素等条件下, 就会失去活性。因此可以设想: 1. 这种特殊的结构降低了特殊键断裂和形成的活化能。2. 结构上有一个专一的鸟苷结合部位。3. rRNA 有两个或两个以上的供转磷酸酯反应的部位, 这个活性部位或是一个缝隙, 或是一个洞, 但它应是疏水的, 这样才能在每一步转移磷酸基的反应中不发生水解。

Cech 等人得出的反应部位结构主要是通过以下实验证实的:

1^[2]. 取一单股 rDNA, 其序列与除去 IVS 后的四膜虫 rRNA 为互补关系。实验表明, 这个单股 rDNA 与前体 rRNA 或剪接后的 rRNA 均可杂交, 并形成双股螺旋结构, 这意味着 “ribozyme” 中, U 和 U' 是很靠近的, 由此也可以看出 rRNA 在剪接前和剪接后构象的改变不大, 剪接反应有可能是在构象最小改变下进行的, 因此在力学上是有利的。

2^[3]. 使用各种专一的只切断单股或双股的核酸酶, 降解 IVS, 然后通过顺序凝胶电泳确定断键位置, 结果表明: IVS 的 5' 端有一发卡结构 (hairpin structure) 见图 2, 而后通过计算机, 得出一与核酸酶实验结果相一致的最低自由能的二级结构。

以上结果表明, IVS 的两个末端是靠近的, IVS 有利于使两个编码部分接近, 进而连在一起, IVS 的环化反应也易发生。图 2 是推测的反应部位和内含子的结构。

Cech 推测这一发卡结构可能是一个动力学结构, 这种结构是基于碱基序列自我互补配对而形成。这种动力学结构可能在剪接过程中存在, 而在环合反应后即告消失。

尽管目前还不清楚剪接, 及环化反应到底要求一个什么样的结构, 但这 14 个核苷酸肯定起着很大作用, 从模型上看, 鸟苷的进攻和

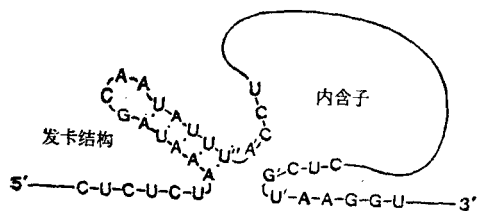


图2 剪接部位的碱基序列^[4]

rRNA 前体的 P—O 键 (U—A 间) 断开皆发生在这个双螺旋臂的一边, IVS 环化发生在臂的另一边。正是这种特殊的结构, 使得这一反应能在无蛋白质存在下进行。

二、自动催化的电子机制

由于剪接及环化反应不需要外来能量, 因此 Cech 等提出磷酸酯键在其断裂过程中, 为新的磷酸酯键的形成提供能量, 如图 3。

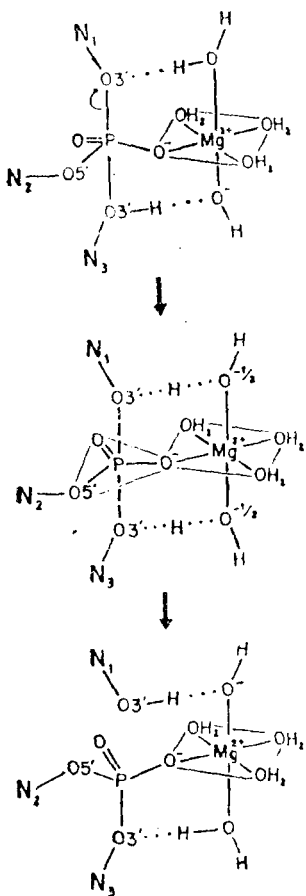


图3 转移磷酸酯反应机制^[4]

Leland C. Allen 等^[4]认为这一机制类似于

S_N2 反应, 与通常的核酸酶 A 催化酯键水解的机制——酸碱催化机制^[5] (图 4) 相类似。在酯键水解反应中, 组氨酸可作为质子的给体——受体起着酸碱催化的作用。

Allen 等提出转移磷酸酯反应也是一种酸碱催化, 实验证明 Mg^{2+} 是剪接反应中所必需的, Allen 认为 Mg^{2+} 很可能与水形成了复合物, 而复合物中的水起了质子的给体-受体作用, 即 Mg^{2+} 先与磷酸中的氧负离子配位, 而后进行催化反应。这种含有磷酸基团作为配位体的金属配位类型, 已为其化合物的晶体结构 X-射线衍射所证实^[6-8]。如果有一个以上磷酸基团(这些磷酸是以它的一个氧与金属配位)在一个 Mg^{2+} 上就会产生较大的键角张力和静电斥力, 因此还没有发现有两个磷酸基团与一个 Mg^{2+} 配位。现所介绍的机制认为 Mg^{2+} 的其它配体

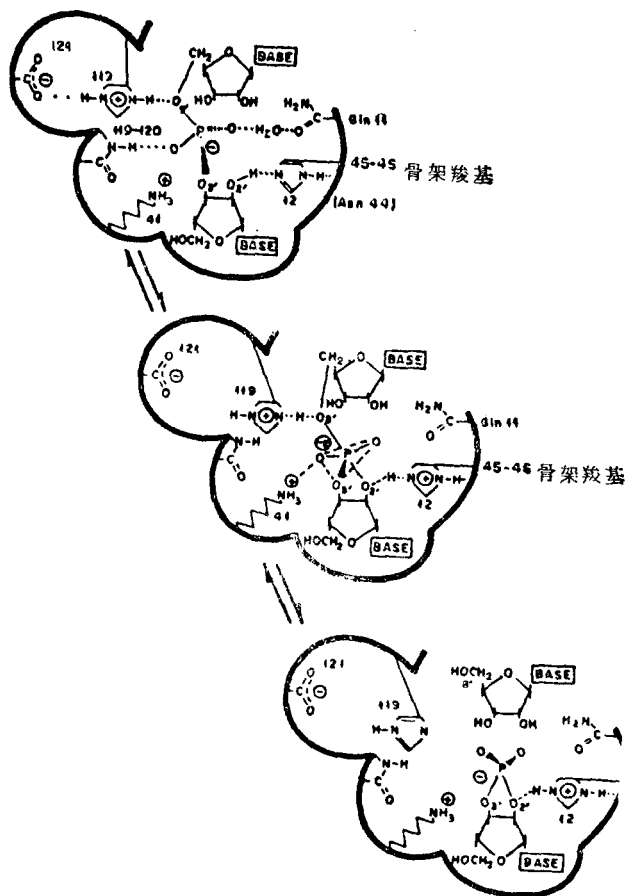


图4 RNase A酶转移磷酸酯反应^[5]

为4个水分子, 和一个氢氧根离子, 形成一个6配位的Mg复合物, 复合物带电为0。在反应时, 图3中(N₃)O3'的亲核方向与将要断裂的P—O3'键“共线”, Mg²⁺复合物的OH⁻接受一个来自于核苷酸3(N₃)O3'的质子, 而与OH⁻

离子在同一轴向的水分子则把其质子供给核苷酸(N₁)O3'位的氧, 产生了一个三角-双锥体的过渡态。(N₁)O3'—P键一经断裂, (N₃)O3'—P键遂形成。短的Mg—OH⁻键接受质子后变长了, 而长的Mg—H₂O键当H₂O失去质子时变

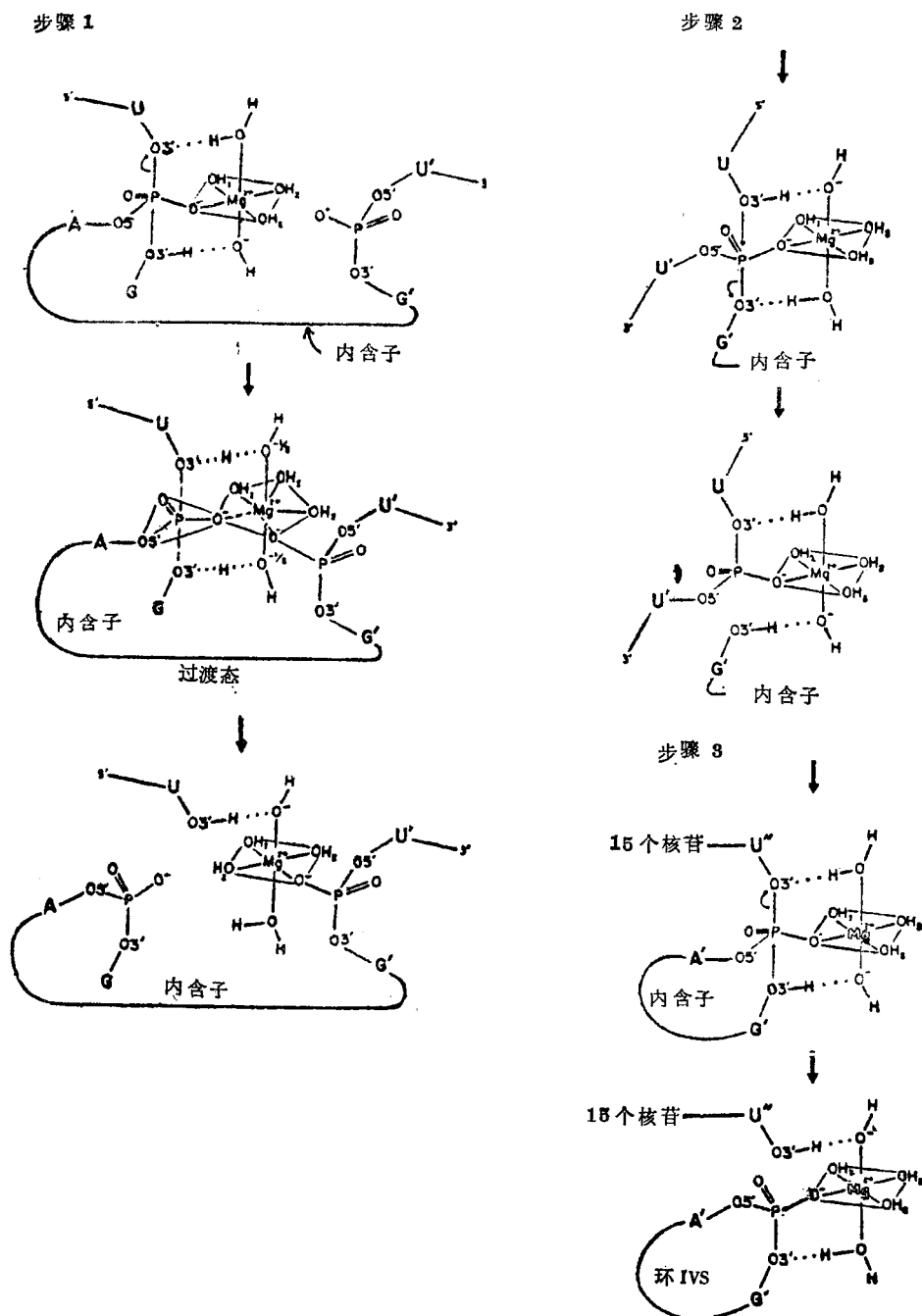


图5 ribozyme 反应的详细机制^[43]

短了,因此 $\text{Mg-H}_2\text{O}$ 复合物的催化机制实质上是酸碱催化。此外, Mg^{2+} 通过从磷酸配位的氧原子上吸引电子,增加了磷原子的正电性,因此也就增加了 $(\text{N}_3)\text{O}3'$ 位氧对磷酸亲核反应的反应性。这一基本机制在 ribozyme 的三步反应中,步步重复。如图 5。

在第一步反应中,外显子 5'-端的切断位于 $(\text{U})\text{O}3'-\text{P}(\text{A})$ 键, G 则连接在插入序列的 5' 端。在这步反应中, Mg^{2+} 原有的磷酸配体与 $\text{G}'\text{pU}'$ 的磷酸基对换,并释放 GpA ,此时该 Mg^{2+} 进入重新使用状态,继续催化第二步反应。在第一步反应中, Mg^{2+} 复合物中的 OH^- 自 G 接受一个 $\text{O}3'$ 上的 H^+ ,变成了水分子,在第二步反应中又把质子供给 G' 的 $\text{O}3'$ 。当 U 的 $\text{O}3'$ 进攻 $\text{G}'\text{pU}'$ 的磷酸基时,两个外显子开始相接。

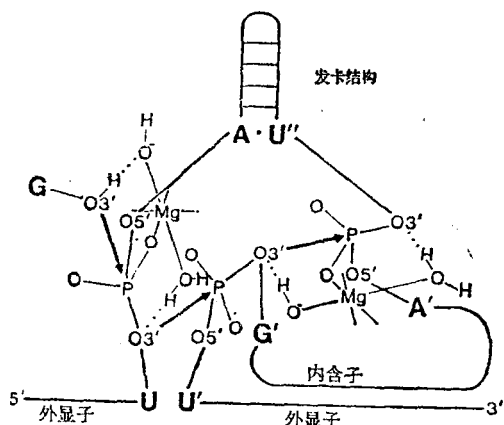


图 6 ribozyme 反应的 Kendrew 骨架模型^[4]

第三步反应是环化反应,这步反应所用的 $\text{Mg-H}_2\text{O}$ 复合物不是前面两步反应所用的复合物,可能因为环化反应同前面两步反应是分开进行的,同时环化反应也需要 Mg^{2+} 以稳定第三个磷酸基的电荷;也可能因为在催化环化反应中, Mg 所处的位置在几何学上难于使用第一个 Mg ,因此第一个 $\text{Mg-H}_2\text{O}$ 复合物在完成第二步反应后,难于过渡到第三步。

从上述的反应机制可搭出 Kendrew 骨架模型(图 6),三个 $\text{O}3'$ 亲核基团皆处于进攻 $\text{P}-\text{O}3'$ 键,发生 $\text{S}_\text{N}2$ 反应的状态,每一步反应的发生都取决于前一步反应的完成,因此,最终归结

于辅助因子鸟嘌呤核苷的连接,引发了剪接反应。

三、催化反应中的分子识别

2',3' 羟基的鸟嘌呤核苷化合物 (GTP, GDP, GMP, 或鸟苷) 在 ribozyme 的反应中起“辅因子 (cofactor)”的作用; ATP, UTP, CTP 及其一、二磷酸化合物,以及不含有 2',3' 羟基的 GTP, GDP, GMP, 鸟苷,则都无活性^[2],因此 ribozyme 能够从不同的核苷中识别出鸟嘌呤核苷的最直接,最简单的方法,就是通过它与鸟嘌呤核苷上的氢原子的结合,即与鸟苷上的氢形成氢键。同所有其它碱基比较,明显的差别是:鸟嘌呤 α 位有一氨基,可以形成氢键,至于鸟嘌呤的 N_3 在腺嘌呤中也有, $\text{H}-\text{N}$, 和 O_6 在空间上类似于尿嘧啶的 $\text{H}-\text{N}_3$ 和 O_4 , 鸟嘌呤的 N_7 也一定不会是氢键形成的位置,因为 N_7 位如被取代,如 7-甲基鸟嘌呤它仍有辅助因子的作用,所以如果形成了两个氢键,鸟苷可被识别的唯一方式是,两个氢键之一是由一个 $\alpha-\text{NH}_2$ 形成的。

那么 ribozyme 的哪个部位与鸟嘌呤相接? 最简单的答案是: 鸟嘌呤 G 与 U 或 A 形成氢键,同样 G' 可与 U'' 或 A' 相接。这个答案可以

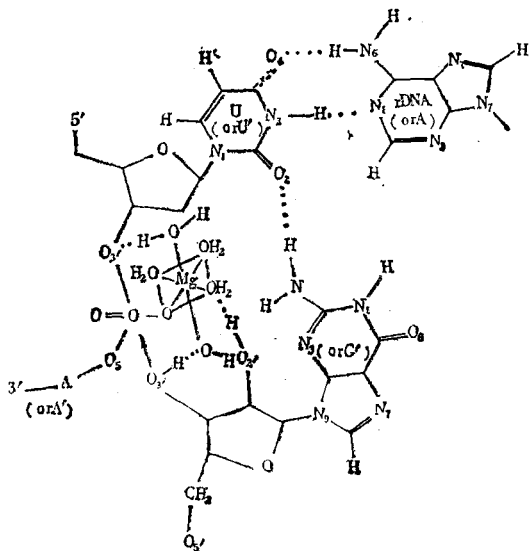


图 7 G 或 G' 在 ribozyme 反应机制中的识别作用^[4]

解释在剪接反应中对鸟嘌呤的要求, 和对尿嘧

啉(或腺嘌呤)的专一性。Leland C. Allen 等, 通过分子模型提出: 唯一合适的氢键形成方式是 G 的 α -NH₂ 的 H 与 U 的 O₃ 成氢键。如图 7。这一方式满足了在 A 与 U'' 碱基配对的假设下, 保证鸟苷的 O3' 处于进攻 UpA 的状态。

鸟嘌呤核苷的 2' 位的 —OH 基团在 ribozyme 反应中也是有所要求的, 模型表明这一个 —OH 基团可以与 Mg²⁺ 复合物中的一个 H₂O 分子形成氢键, 这样有利于提高 ribozyme 对鸟嘌呤核苷的专一识别作用。

四、自动剪接、自动环化反应 在生物体系中的意义

自动剪接反应的重要性在于它为解释早期的生命进化开辟了一个新的领域。可以这样设想: 四膜虫核糖体 RNA 的剪接展现了一个古老的反应机制, 一种曾是生物化学反应中的催化形式。因此它在从事生命起源研究的人中引起震动, 尤其是对于那些主张 RNA 作为最初的遗传物质比 DNA 更为合适的人们。这一发现对他们是一个重大支持。Noller 说: “如果 RNA 既能贮存信息又能进行催化反应, 那么很明显它适于作为最早的基因物质。”

IVS 的环化作用^[10]: 在一些四膜虫品系的 rRNA 中没有插入序列, 因为在它们的 rDNA 中没有 IVS, 这一事实排除环形 IVS RNA 有某种生物功能的可能; IVS 的消除和插入序列的环合可分开进行, 也否定了环化反应在剪接过程中起作用的假定。以上是 Cech 等人进行体外实验的结果, 体内实验大概也是如此。因此认为环化反应的作用是: 它破坏了 rRNA 剪接反应的产物而使反应正向进行, 降低逆反应速度。如果逆反应发生了, 主要会出现两种情况, 一种是 IVS 重新插入已经剪接过的 rRNA 中, 这将降低 rRNA 成熟的速度。另一种可能

的逆反应结果, IVS 插入到 rRNA 的另一位置 [发生转位], 这种结果害处更大。而环化反应则通过变线型 IVS 为两部分, 降低了 IVS 逆反应的活性, 阻止了逆反应的进行。

1983 年 10 月, Cech 等根据以往的工作, 按照剪接部位的差异, 将断裂基团的剪接方式分为以下三类:

1. 剪接反应部位由内含子和外显子相接的一短的碱基顺序决定。如真核 mRNA。

2. 剪接反应的发生, 取决于外显子的结构。如真核的 tRNA。

3. 剪接反应主要取决于内含子的结构。如核中的 rRNA 和线粒体 RNA。

由上可见, 对核内基因, 三类 RNA 的转录和剪辑体系各不相同。所以这种核基因的分类有利于追溯基因的演化。可以预期, 在今后几年的研究中, 不仅内含子的消除机制将会有一个圆满的阐述, 对内含子的作用还将有更深入的认识。

参 考 文 献

- [1] Lewin, R.: *Science*, **218**, 872, 1982.
- [2] Kruger, K., et al.: *Cell*, **31**, 147, 1982.
- [3] Cech, T. R., et al.: *Proc. Natl. Acad. Sci. USA*, **80**, 3903, 1983.
- [4] Karen Hydock and Leland C. Allen: “Self-Splicing RNA”, Princeton University's lecture note, 1982, (unpublished).
- [5] Deakyne, C. A. and Allen, L. C.: *J. Amer. Chem. Soc.*, **101**, 3951, 1979.
- [6] Hingerty, B., et al.: *J. Mol. Biol.*, **124**, 523, 1978.
- [7] Aoki, K., et al.: *Biochim. Biophys. Acta*, **425**, 369, 1976.
- [8] Aoki, K.: *Biochim. Biophys. Acta*, **447**, 379, 1976.
- [9] Cech, T. R., et al.: *Cell*, **27**, 487, 1981.
- [10] Zaug, A. J., et al.: *Nature*, **301**, 578, 1983.
- [11] Cech, T. R.: *Cell*, **34**, 713, 1983.

[本文于 1985 年 9 月 12 日收到]