

文章编号: 1004-0374(2009)04-0524-07

自我剪接内含子结构与剪接的关系

张相萍¹, 林 瑛², 屈 艾¹, 孟 清^{2*}

(1 徐州师范大学生命科学学院, 徐州 221116; 2 东华大学生物科学与技术研究所, 上海 201620)

摘 要: 自我剪接内含子是指具有催化活性的RNA分子, 可介导自身的剪切和两侧外显子的连接。自我剪接内含子包括I型内含子和II型内含子两类, 主要存在于原生生物、真菌、藻类、植物细胞器以及细菌和古细菌(II型内含子)基因组中。尽管核苷酸序列保守性很低, 但两类自我剪接内含子均能分别形成不同的保守二级结构, 且它们都可通过两步连续的转酯反应完成剪接。但由于其结构存在较大差异, 导致其剪接机制也各不相同。对其结构与剪接关系的深入研究将有助于进一步探索自我剪接内含子在生物体内的功能、起源和进化。

关键词: 自我剪接内含子; ORF; 结构; 剪接; 成熟酶

中图分类号: Q522 **文献标识码:** A

Relationship between structure and splicing of self-splicing introns

ZHANG Xiang-ping¹, LIN Ying², QU Ai¹, MENG Qing^{2*}

(1 School of Life Science, Xuzhou Normal University, Xuzhou 221116, China; 2 Institute of Biological Sciences and Biotechnology, Donghua University, Shanghai 201620, China)

Abstract: Self-splicing introns are catalytic RNAs and divided into two groups, group I intron and group II intron. They are found in mitochondria and chloroplasts genomes of plants, fungi, protists and algae, as well as in bacterial and archaeobacterial (group II intron) genomes. All self-splicing introns can fold into their respective conserved secondary structures, although they are different in sequences. The typical secondary structure of a group I intron consists of approximately ten paired elements, and group II introns have a typical structure with six double-helical domains (DI-DVI). Self-splicing introns can self-splice from their pre-RNAs by two consecutive transesterification reactions joining the flanking exons and releasing the introns. Group I introns use an exogenous G to initiate the splicing reaction, but group II introns use an internal bulged adenosine in DVI. Their respective secondary structure and protein factors are all important for the splicing reaction. Group I intron and group II intron have been used in bioengineering and also have many potential applications in functional genome and gene therapy. Self-splicing introns have become a research focus with the development of their structure and function.

Key words: self-splicing introns; ORF; structure; splicing; maturase

自我剪接内含子指在没有任何酶或蛋白质因子存在的条件下, 可通过自我剪接反应将自身从前体中剪切去除, 并将相邻的5'-端和3'-端外显子(exon)连接成成熟mRNA的一类RNA分子。这类内含子又称为核糖核酸内切酶(ribozyme), 即核酶。

自我剪接内含子是重要的遗传元件, 可通过归巢过程在不同的基因间发生转移。自我剪接内含子可在RNA水平上发生自我剪接, 对宿主基因无任何

毒害作用, 成为研究RNA折叠和催化的模式分子, 对其结构与剪接关系的深入研究将有助于我们探索自我剪接内含子在生物体内的功能、起源和进化。

收稿日期: 2009-05-13; 修回日期: 2009-06-30

基金项目: 国家自然科学基金项目(30770463)

*通讯作者 Tel: 021-67792651; E-mail: mengqing@dhu.edu.cn

1 自我剪接内含子的种类和分布

根据结构特征, 自我剪接内含子主要分为 I 型内含子(group I intron)与 II 型内含子(group II intron)两大类。I 型内含子数量最多、分布范围最广, 主要分布于细胞核 rRNA 基因、真菌的线粒体和质体基因组、细菌^[1]以及原生物体内, 但目前尚未发现古细菌体内有 I 型内含子的存在。II 型内含子存在于植物和真菌的线粒体和叶绿体中^[2], 也存在于大约 25% 的真核生物基因组中。古细菌体内也发现了 II 型内含子的存在^[3-5], 但仍未发现自我剪接内含子在动物体内的存在^[6]。

内含子长期以来一直被认为是真核生物特有的一类遗传元件。然而, 自 10 年前在细菌中发现第一个 II 型内含子以来, 特别是最近几年, 随着基因组测序的迅速发展, 发现内含子存在于 G⁺ 和 G⁻ 细菌中也是一种极普遍的现象, 且绝大多数是 II 型内含子, I 型内含子较少。目前报道的 I 型内含子(<http://www.rna.icmb.utexas.edu>), 大多数位于 tRNA 基因中, 少数位于 rRNA 基因中, 仅在 *Bacillus anthracis* 的染色体 *recA* 基因^[7] 和蓝细菌 *Trichodesmium erythraeum* 的 *rir* (编码核糖核酸还原酶) 基因^[1] 中各发现了一个 I 型内含子。约 1/4 已测序的细菌基因组包含有一个或一个以上 II 型内含子, 甚至在古细菌中也发现 II 型内含子的踪迹^[3-5]。

细菌的基因密度极高, 但是与真核生物相比, 许多细菌的 I 型内含子和 II 型内含子被排除在保守的蛋白编码基因之外。例如蓝藻类细菌 *Thermosynechococcus elongatus* 拥有的 28 个 II 型内含子中^[8], 有 26 个内含子处在基因之外或移动元件内(插入序列等), 另外 1 个位于 tRNA 基因中, 1 个位于类似蛋白编码基因而非功能基因中^[5, 9, 10]。由此可见, 在细菌基因组中一定存在某种屏障(barrier)阻碍了内含子进入蛋白质编码基因。而真核生物中几乎所有的 I 型内含子和 II 型内含子都保守地存在于编码蛋白质的基因中。这一现象表明内含子插入基因具有选择性^[9, 11]。

借助 BLAST 和 FASTA 等软件, 对国际基因数据库中公布的 *T. erythraeum* 基因组序列进行初步的筛选、分析、比较和鉴定, 在 *rir* 和 *dnaM*^[12] (编码保守的 DNA polymerase III β 亚基) 这些保守的蛋白编码基因中分别发现了 3 个和 4 个 II 型内含子, 这与已经鉴定的许多细菌的 I 型内含子和 II 型内含子都存在于保守的蛋白质编码基因之外相悖。在 *rir* 基

因中还发现有 4 个蛋白质内含子(intein)与 3 个 II 型内含子共存。细菌中多个 II 型内含子与多个蛋白质内含子共存于一个基因中, 以前从未有过报道。超长的 *rir* 基因中拥有的 7 个间隔区已占超过 80% 的序列, 这只是以前在真核生物中见到过的一个特征^[12]。这一发现将有助于揭示细菌基因组内含子进入蛋白质编码基因存在的屏障。

通过进一步的分析鉴定, 在 *T. erythraeum* 基因组序列中又发现多个 II 型内含子, 且这些 II 型内含子多数都存在于保守的蛋白编码基因中。*T. erythraeum* 基因组中存在的 I 型内含子和多个 II 型内含子将为研究其结构与功能、起源和进化、剪接与转移提供多方面的实验材料。

2 I 型内含子

2.1 结构 武汉大学张翼等研究结果^[13, 14] 及 Comparative RNA 数据库收录的 I 型内含子序列^[15] 表明, 其长度为 140—4 200 nt, 序列保守性很小。I 型内含子的二级结构主要由 10 个臂状结构(P1—P10) 和 10 个环状结构(L1—L10) 及臂环间的连接结构构成(图 1)。有时还存在一些不稳定配对结构, Li 等^[16] 研究表明 IE 内含子含有 14 个保守臂环结构。臂环结构折叠形成催化核心, 外围结构自核心伸展出来^[17-20]。P1 的 3' 端为内含子序列, 5' 端为 5'—外显子序列, 即剪接反应中的底物链, P1 末端的 G·U 碱基对在剪接反应中可被识别为 5' 剪接位点, 剪接第一步完成之后形成 P10 中的 Ω G(图 2c)^[21] 在 3' 剪接位点的识别中具有重要作用。活性中心(阴影部分)位于 P3—P7 及 J4/5、J8/7 附近^[22, 23], 主要由 P4—6(包括 P4、P5 和 P6) 和 P3—8(包括 P3、P7 和 P8) 两大结构域组成。虽然催化核心的结构是保守的, 但除了位于活性位点的几个碱基外, 其核苷酸序列并无明显的保守性。

2.2 编码蛋白质 许多长度在 1 000 nt 以上的 I 型内含子具有编码蛋白质的开放阅读框(open reading frame, ORF), 如细胞状粘菌(*Dictyostelium discoideum*) 的线粒体 DNA 中编码细胞色素氧化酶^[24] 亚基 I 和 II 融合基因中含有 4 个 I 型内含子, 其中 3 个 I 型内含子含有 ORF 编码序列, 分别位于 L1、L2 及 L8(图 1)。这些冗长的编码序列通常编码具有成熟酶活性和核酸内切酶活性的蛋白质, 有人认为 I 型内含子编码蛋白质的成熟酶活性源于核酸内切酶活性^[25], 不同的核酸内切酶具有不同的氨基酸模体: LAGLIDADG 模体、GIY-YIG 模体、H-N-H 模体及

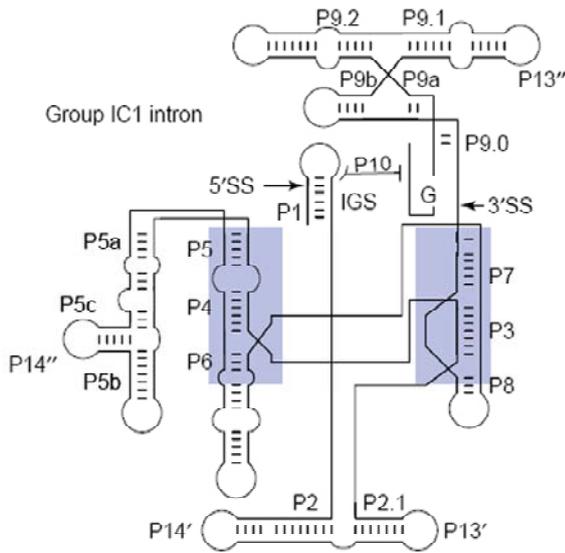


图1 I型内含子的剪接机制图^[6]

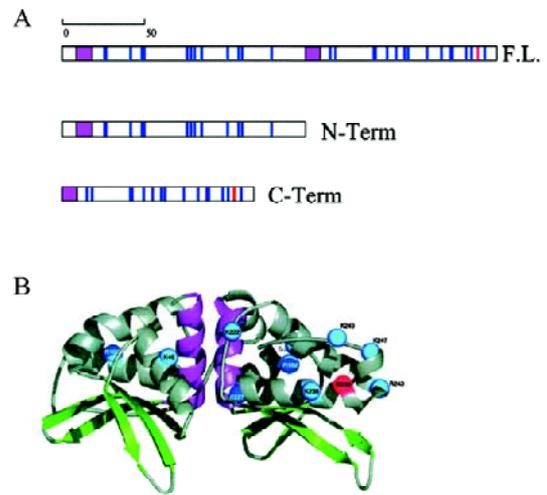


图3 I型内含子的二级结构示意图^[28]

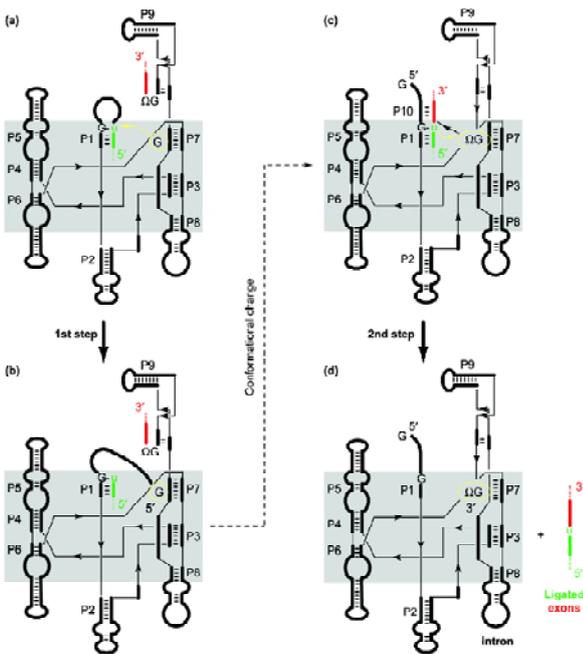


图2 I-AniI蛋白的结构示意图^[21]

His-Cys 模体^[26]。其中LAGLIDADG 模体最大且研究也最为清楚，H-N-H 模体最小最不清楚^[27]。Downing等^[28]对*Aspergillus nidulans*中I-AniI的研究表明，内含子编码的蛋白质分为N-端结构域和C-端结构域(图3)。I-AniI中含有两个不同的LAGLIDADG 模体，一个位于N-端结构域，一个位于C-端结构域，如图3A所示，红色 α 螺旋结构即为I-AniI蛋白的LAGLIDADG 模体；蓝色部分为蛋白质序列中保守的氨基酸残基，这些残基对蛋白发

挥功能至关重要；绿色部分是核酸结合部位，由图可知均为 β 折叠结构。实验表明C-端结构域具有帮助内含子剪接的活性，而N端结构域主要帮助内含子发生归巢。

2.3 结构与剪接的关系 I型内含子与II型内含子的剪接均包括两步转酯反应，但不同的结构导致其剪接机制也各不相同。

大多数I型内含子在体外能够独立的完成自剪接，少数I型内含子需要蛋白质因子(即成熟酶)帮助才能完成正确的折叠。I型内含子的剪接反应包括两步磷酸酯键的转移。首先，一个外源的鸟苷G结合到内含子催化中心的G结合位点，作为亲核试剂以其3'-OH攻击5'剪接位点的磷酸二酯键，与内含子的5'端第一个核苷酸形成3',5'-磷酸二酯键，并释放出5'-外显子；外源鸟苷即离开，G结合位点被5'剪接位点附近的 Ω G取代。而后，游离5'-外显子的3'-OH攻击3'剪接位点的磷酸二酯键，导致5'-外显子与3'-外显子的连接和内含子的释放^[29](图2)。I型内含子自我剪接只需要外源鸟苷G和镁离子即能自发进行，无需供给能量和酶。但是RNA必须折叠成为正确的构象即必须形成天然的功能结构，而在此过程中内含子各结构域均具有不同的作用。Xie等^[30]对细菌噬菌体的外壳蛋白MS2进行研究，得出此RNA结合蛋白结合于真核生物中I型内含子，主要与P5ab发生作用，使得L9-P5之间的相互作用更加稳定，从而使内含子的催化中心更加紧密而改变了3'水解活性。构成催化核心的保守结构元件对其发挥功能是必不可少的，但外围非保守的结构元件对内含子的功能发挥同样必不可少。Xie

等以内含子的自我剪接活性为依据对IE亚组I型内含子的各外围结构元件进行筛选, 最终发现假丝酵母*Candida* I型内含子中P2.1对催化活性结构折叠是必需的。他们还推测P2.1极有可能是通过与核心螺旋结构P3和P6形成三级螺旋作用参与催化核心的折叠^[31]。

Chauhan和Woodson^[32]通过位点特异性突变对细菌I型内含子的一个环状结构及其受体进行删除, 发现内含子的三级结构相互作用很不稳定, 表明这种三级结构相互作用在内含子的天然态和类似天然态的过渡中起协调作用。他们还对突变的I型内含子进行变性聚丙烯酰胺凝胶电泳和RNase T1酶切实验, 结果显示, 突变内含子在折叠初始即很难形成均一的碱基配对, 且他们还证实天然折叠的内含子在5-20 ms内即完成折叠过程, 而突变体中40%-60%需要30-40 ms, 其余则需30-200 ms。因此, 正确的三级结构相互作用能够保证内含子形成稳定的三维结构, 而只有内含子形成稳定的三维结构, 才能直接快速的折叠成天然的功能结构, 继而完成其剪接。I型内含子中含有GNRA(其中N为A、U、C或G; R为G或A)序列的环能够形成稳定而保守的发夹结构, GNRA高频率的存在于I型内含子中, 可形成环-受体结构参与三级结构相互作用。这种发夹结构帮助内含子折叠成正确的分子结构, 与蛋白质发生相互作用, 阻止逆转录酶的延伸等, 在内含子的剪接中发挥重要的生物学功能^[33]。图2中的阴影部分即为I型内含子的剪接活性结构区, 是由多个发夹结构和单链连接区构成, 其中多数发夹结构即含有GNRA, 如L2、L9等。

外源鸟苷G亲核攻击5'剪接位点, 是内含子剪接的发起者, 在整个剪接过程中发挥重要的作用, 近年来对其结构特点和生化特性研究颇多。Houglund等^[34]证明外源鸟苷2'-OH在四膜虫的I型内含子剪接反应中形成氢键, 是催化反应必需的。将此2'-OH删除, 内含子即不能折叠为催化活性结构, 生化分析与结构分析表明2'-OH为催化核心内部金属离子的配位基团, 原子突变循环法证实不论基态还是跃迁态2'-OH均为氢键供体, 在剪接辅助因子的识别及整个剪接反应中发挥重要作用。

*Tetrahymena*内含子P4-P6区的晶体结构解析结果显示^[35], P4-P6区RNA在P5部分有一个150°的弯折, 被称为P5abc的部分与P4、P5和P6并排靠在一起, 形成了一个紧密的发夹形状。

为这种形状提供稳定力量的是一个由2个镁离子介导的主链接触和一个核糖介导的主链接触。1998年, Golden等^[36]对构成*Tetrahymena*核酶催化核心的P4-P6与P3-P9结构域的晶体结构研究表明, 核酶RNA折叠成了规则紧密的形状, 非常类似蛋白质酶的构造。P4-P6与单独结晶时的构象相比没有大的改变, 仍旧呈发夹状。P3-P9绕在P4-P6的一边, 两个结构域一起形成了一个裂隙, 可以容纳包含5'剪接位点的短双螺旋。没有采取正常RNA双螺旋的A构象, 而是形成了一个适合结合鸟苷的口袋形状。这清楚地显示了一个在没有底物情况下已预先折叠好的催化结构。这些结构充分证明了大RNA分子可以形成紧密有序的空间构象来执行其催化功能。2004年, Adams等^[37,38]又相继报道了完整的I型内含子(包括5'-外显子和3'-外显子)的晶体结构及多个螺旋间的相互作用, 进一步揭示了结构与剪接之间的关系。

近期, 我们对鉴定自*Nostoc punctiforme*基因组的一个I型内含子结构预测表明, L1与P10、L1与J8/7两种相互作用是内含子剪接必需的, L1与P10之间、L1与J8/7之间的相互作用使得内含子形成正确的构象从而完成自身剪接。为了验证这一预测, 我们构建了此I型内含子的一系列突变体。实验先对其中一个结构(L1)进行突变, 发现突变体并不剪接, 再对突变体另一个结构(分别为P10和J8/7)进行突变, 表达结果证明发生一次突变后的克隆均不剪接, 而发生两次突变的克隆恢复了剪接能力, 这就表明L1与P10和L1与J8/7两种相互作用对其剪接是必需的。

3 II型内含子

3.1 结构 II型内含子的长度从几百个碱基到几千个碱基不等, 核苷酸一级序列保守性小, 但二级结构保守, 主要由6个茎环结构域(DI-DVI)组成, 呈车轮状放射分布。其中DV为催化中心^[39,40], 通常由34个核苷酸构成; DI最大, 包含帮助内含子进行特定的构象变化的外显子结合序列(EBS)(图4A); DVI含有一个突起的碱基A^[41], 是剪接反应的发起者, 在剪接后的产物中形成分支位点; DIV环最不保守, 某些能够编码蛋白质的内含子在DIV环含有ORF, 编码的蛋白质通常具有四个功能结构域(图4B): RT结构域(具有反转录酶的活性, 亚结构域0-7)、X结构域(具有成熟酶的活性, 帮助内含子剪接)、D结构域(非保守的DNA结合区域)、

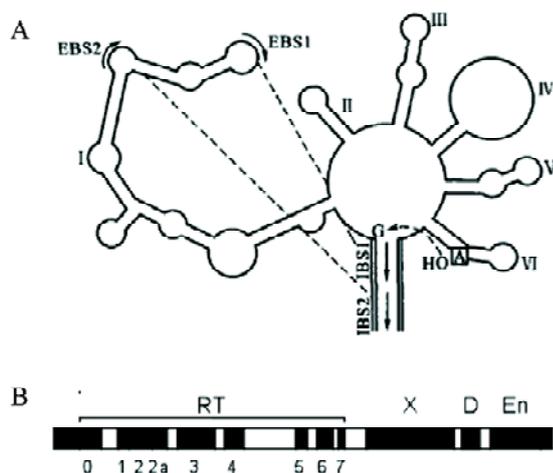


图4 II型内含子的二级结构图

En 结构域(具有核酸内切酶的活性)^[42-44]。这些不同的结构域具有不同的作用,内含子在体内的剪接依赖 X 结构域。内含子作为一个移动的遗传元件需要所有四个结构域,但并非所有 II 型内含子编码的蛋白质均具有上述四个功能区,如鉴定自 *T. erythraeum* 基因组中的 24 个 II 型内含子中,19 个具有蛋白质编码序列,4 个没有蛋白质编码序列,另外一个目前还不清楚;而其中仅 10 个 II 型内含子编码的蛋白质具有完整的四个功能区,9 个缺失其中的一个或两个功能区或某个功能区的一部分(部分研究结果待发表)^[12, 45]。II 型内含子结构的多样性、是否编码蛋白质以及编码蛋白质功能区的多样性,揭示 II 型内含子可能在生物功能基因的表达与调控过程中起重要作用。

3.2 编码蛋白质 II 型内含子可折叠成一个具有六个结构域的保守二级结构,内含子编码的成熟酶辅助相似 I 型内含子的剪接,那么这些成熟酶是结合在 II 型内含子六个结构域的哪个部位起作用呢;成熟酶在辅助内含子剪接过程中是帮助维持其正确构型,还是作为一个催化单位行使功能;细菌 II 型内含子的剪接功能区域究竟在 II 型内含子的哪个部位等一系列问题在国际上还都是空白。我们以克隆自 *T. erythraeum* 基因组的 pDN4^[45] II 型内含子作为研究对象,对 II 型内含子的 DIVa 区^[46]进行研究,检测 DIVa 区是否是剪接必需的。我们对 pDN4(自身并不编码剪接所需的成熟酶)的 DIVa 区进行不同的缺失突变,另外,将 pDN4 II 型内含子的 DIVa 区与另一个 II 型内含子的 DIVa 区进行互换,然后由成熟酶 pAM1 (*rir* 基因中第三个 II 型内含子编码的成

熟酶)协助其剪接。实验结果表明,我们所得到的 pDN4 II 型内含子四种不同的缺失突变体在成熟酶 pAM1 的帮助下均能发生剪接, pDN4 的 DIVa 区与另一 II 型内含子的 DIVa 区互换后,两个重组克隆在成熟酶 pAM1 的帮助下均发生剪接,但是剪接效率有所降低,这表明 DIVa 区这一结构对剪接并非必需,但是对其剪接效率是有一定影响的(研究结果另文发表)。这与 p62(reverse transcriptase-maturase protein)结合于 DVIa 区对 aI2 的体内剪接并非必需^[47]是相符的。

尽管某些 II 型内含子在体外非生理条件下可以催化自身的剪接,但它们在体内的剪接需要特异性的成熟酶或蛋白质帮助完成。真核生物细胞器的 II 型内含子募集细胞蛋白质帮助内含子剪接,而细菌的 II 型内含子能编码一个成熟酶蛋白特异性地帮助自身剪接和自身内含子的转移,所有先前分析的成熟酶都只能特异性地与编码它们的内含子起作用,但是密切相关的成熟酶之间可能存在交叉反应。

植物叶绿体只有一个编码成熟酶相关蛋白质的 II 型内含子即 trnK-I1, trnK-I1 编码的蛋白质称为 MatK, Vogel 等^[48]证明 MatK 可能帮助许多缺少 ORF 的 II 型内含子剪接。我们将几个不同的内含子编码的不同成熟酶与同一个内含子分别进行共表达,结果表明这些成熟酶均能帮助同一内含子完成剪接,孟清等^[45]也已得出一种内含子编码的成熟酶能够帮助多个不同基因中的结构相似的不同内含子完成剪接,以上结论均表明不同内含子编码的成熟酶之间存在交叉反应。

3.3 结构与剪接的关系 II 型内含子剪接也包含两步转酯反应。第一步通过 5' - 外显子中的内含子结合序列 (IBS) 与对应的 EBS 相互作用正确定位 5' 剪接位点,使其靠近 DVI 中突起的 A, A 中 2' -OH 进攻 5' 剪接位点的磷酸二酯键,释放 5' - 外显子,形成套索状中间产物。

第二步通过两个单一的碱基配对相互作用正确定位 3' - 外显子;这种相互作用随着 RNA 种类的不同而不尽相同^[41]。第二步转酯反应与第一步相似,随着 5' - 外显子释放,其 3' -OH 进攻 3' 剪接位点,导致内含子两侧外显子连接并释放套索状内含子。剪接反应中磷酸二酯键的断裂和连接是均等的,不需要外源能量。

体外自剪接反应需要相对极端的反应条件,如 Mg²⁺ 等金属离子。金属离子帮助内含子进行构象改

变使其折叠成为有功能的活性结构, 完成自身剪接。Steiner等^[49]对II型内含子折叠过程与 Mg^{2+} 的关系进行研究, 利用单分子荧光标记技术对 Mg^{2+} 不同浓度时II型内含子的折叠进行动态分析发现内含子折叠具有 Mg^{2+} 依赖性。在II型内含子的折叠过程中观察到三种不同的构象分子, 而只有在 Mg^{2+} 浓度高于20 mmol/L时内含子才折叠成可结合底物的构象, 随着 Mg^{2+} 浓度的增加, 内含子的构象不断变化直到形成活性结构。在体内, 内含子剪接需要蛋白质因子(成熟酶)帮助内含子折叠成催化活性结构, 且与其他区域的相互作用诱导其结构变化, 从而促进自剪接, 产生剪接的外显子和核糖核蛋白(ribonucleoprotein, RNP)。

许多生化研究已经证实II型内含子的大部分结构特点, 但是对其功能结构的组织空间和活性中心的分子排布仍无详细的描述。Toor等^[50]对*Oceanobacillus iheyensis*中一个完整II型内含子的结构进行晶体学分析, 发现多种三级结构相互作用形成网络促进内含子的亚结构域围绕活性中心(DV)有序的折叠, DV的突出部分在大沟处(催化中心)形成一个异常的螺旋结构与催化中心共同连接在两个二价金属离子上, 恰与“两个金属离子”这一催化反应机制相符。而结构与功能模拟实验也支持II型内含子与剪接体内含子源于同一祖先的假说^[2, 51]。

4 结语

自我剪接内含子可介导自身的剪切和两侧外显子连接; 同时, 作为独特的遗传元件, 归巢和反转录转座使得它们能够在基因内或基因间发生移动。在过去的几十年里, 关于自我剪接内含子的结构、结构特征与剪接的关系以及剪接机制的研究取得了许多重要成果, 但仍有许多关键问题有待解决。利用随机突变加上功能选择导致分子定向进化, 将有望进一步阐明自我剪接内含子结构与功能的关系; 克隆来自不同物种的自我剪接内含子(II型内含子)以及其编码的蛋白质, 深入研究不同II型内含子编码的成熟酶蛋白能否协助同种或多种细菌中、相同或不同基因中以及相似或不相似的内含子的剪接, 将有望进一步阐明II型内含子的起源和进化历史; 对自我剪接内含子不同结构域进行定点突变, 构建不同的内含子供体和受体, 将有望进一步阐明原核生物II型内含子的剪接机理、转移过程以及在基因表达过程中的调控作用。

[参 考 文 献]

- [1] Meng Q, Zhang Y, Liu XQ. Rare group I intron with insertion sequence element in a bacterial ribonucleotide reductase gene. *J Bacteriol*, 2007, 189(5):2150-4
- [2] Michel F, Umesono K, Ozeki H. Comparative and functional anatomy of group II catalytic introns—a review. *Gene*, 1989, 82(1):5-30
- [3] Dai L, Zimmerly S. ORF-less and reverse-transcriptase-encoding group II introns in archaeobacteria, with a pattern of homing into related group II intron ORFs. *RNA*, 2003, 9(1):14-9
- [4] Rest JS, Mindell DP. Retroids in archaea: phylogeny and lateral origins. *Mol Biol Evol*, 2003, 20(7):1134-42
- [5] Toro N. Bacteria and archaea group II introns: additional mobile genetic elements in the environment. *Environ Microbiol*, 2003, 5(3):143-51
- [6] Haugen P, Simon DM, Bhattacharya D. The natural history of group I introns. *Trends Genet*, 2005, 21(2):111-9
- [7] Ko M, Choi H, Park C. Group I self-splicing intron in the recA gene of *Bacillus anthracis*. *J Bacteriol*, 2002, 184(14):3917-22
- [8] Nakamura Y, Kaneko T, Sato S, et al. Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *DNA Res*, 2002, 9(4):123-30
- [9] Dai L, Zimmerly S. Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res*, 2002, 30(5):1091-102
- [10] Dai L, Toor N, Olson R, et al. Database for mobile group II introns. *Nucleic Acids Res*, 2003, 31(1):424-6
- [11] Edgell DR, Belfort M, Shub DA. Barriers to intron promiscuity in bacteria. *J Bacteriol*, 2000, 182(19):5281-9
- [12] Liu XQ, Yang J, Meng Q. Four inteins and three group II introns encoded in a bacterial ribonucleotide reductase gene. *J Biol Chem*, 2003, 278(47):46826-31
- [13] 李志杰, 张翼. I型内含子核酶研究进展. *生物化学与生物物理进展*, 2003, 30(3):363-9
- [14] Zhou Y, Lu C, Wu QJ, et al. GISSD: Group I intron sequence and structure database. *Nucleic Acids Res*, 2008, 36(Database issue):D31-7
- [15] Cannone JJ, Subramanian S, Scunare MN, et al. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, 2002, 3:2
- [16] Li ZJ, Zhang Y. Predicting the secondary structures and tertiary interactions of 211 group I introns in IE subgroup. *Nucleic Acids Res*, 2005, 33(7):2118-28
- [17] Davies RW, Waring RB, Ray JA, et al. Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature*, 1982, 300(5894):719-24
- [18] Michel F, Jacquier A, Dujon B. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie*, 1982, 64(10):867-81
- [19] Waring RB, Davies RW. Assessment of a model for intron RNA secondary structure relevant to RNA self-splicing—a

- review. *Gene*, 1984, 28(3):277-91
- [20] Lehnert V, Chytil M, Verdine GL, et al. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem Biol*, 1996, 3(12):993-1009
- [21] Vicens Q, Cech TR. Atomic level architecture of group I introns revealed. *Trends Biochem Sci*, 2006, 31(1):41-51
- [22] Ikawa Y, Shiraishi H, Inoue T. Minimal catalytic domain of a group I self-splicing intron RNA. *Nat Struct Biol*, 2000, 7(11):1032-5
- [23] Strobel SA, Ortoleva-Donnelly L. A hydrogen-bonding triad stabilizes the chemical transition state of a group I ribozyme. *Chem Biol*, 1999, 6(3):153-65
- [24] Ogawa S, Matsuo K, Angaa K, et al. Group-I introns in the cytochrome oxidase genes of *Dictyostelium discoideum*: two related ORFs in one loop of a group-I intron, a *cox1/2* hybrid gene and an unusually large *cox3* gene. *Curr Genet*, 1997, 31(1):80-8
- [25] Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res*, 2001, 29(18):3757-74
- [26] Belfort M, Roberts RJ. Homing endonucleases: keeping the house in order. *Nucleic Acids Res*, 1997, 25(17):3379-88
- [27] Guhan N, Muniyappa K. Structural and functional characteristics of homing endonucleases. *Crit Rev Biochem Mol Biol*, 2003, 38(3):199-248
- [28] Downing ME, Brady KL, Caprara MG. A C-terminal fragment of an intron-encoded maturase is sufficient for promoting group I intron splicing. *RNA*, 2005, 11(4):437-46
- [29] Cech TR. Self-splicing of group I introns. *Annu Rev Biochem*, 1990, 59:543-68
- [30] Xie MH, Wu QJ, Jiang YF, et al. A phage RNA-binding protein binds to a non-cognate structured RNA and stabilizes its core structure. *Biochem Biophys Res Commun*, 2009, 378(2):168-73
- [31] Xiao M, Li TT, Yuan XY, et al. A peripheral element assembles the compact core structure essential for group I intron self-splicing. *Nucleic Acids Res*, 2005, 33(14):4602-11
- [32] Chauhan S, Woodson SA. Tertiary interactions determine the accuracy of RNA folding. *J Am Chem Soc*, 2008, 130(4):1296-303
- [33] Prathiba J, Malathi R. Group I introns and GNRA tetraloops: remnants of 'The RNA world'? *Mol Biol Rep*, 2008, 35(2):239-49
- [34] Houglund JL, Sengupta RN, Dai Q, et al. The 2'-hydroxyl group of the guanosine nucleophile donates a functionally important hydrogen bond in the tetrahymena ribozyme reaction. *Biochemistry*, 2008, 47(29):7684-94
- [35] Cate JH, Gooding AR, Podell ER, et al. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, 1996, 273(5282):1678-85
- [36] Golden BL, Gooding AR, Podell ER, et al. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science*, 1998, 282(5387):259-64
- [37] Adams PL, Stahley MR, Gill M, et al. Crystal structure of a group I intron splicing intermediate. *RNA*, 2004, 10(12):1867-87
- [38] Adams PL, Stahley MR, Kosek AB, et al. Crystal structure of a self-splicing group I intron with both exons. *Nature*, 2004, 430(6995):45-50
- [39] Qin PZ, Pyle AM. The architectural organization and mechanistic function of group II intron structural elements. *Curr Opin Struct Biol*, 1998, 8(3):301-8
- [40] Michel F, Ferat JL. Structure and activities of group II introns. *Annu Rev Biochem*, 1995, 64:435-61
- [41] Jacquier A, Michel F. Multiple exon-binding sites in class II self-splicing introns. *Cell*, 1987, 50(1):17-29
- [42] Zimmerly S, Hausner G, Wu XC. Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res*, 2001, 29(5):1238-50
- [43] Craig NL, Craigie R, Gellert M. *Mobile DNA II* [M]. Washington DC: ASM Press, 2002
- [44] San Filippo J, Lambowitz AM. Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein. *J Mol Biol*, 2002, 324(5):933-51
- [45] Meng Q, Wang YF, Liu XQ. An intron-encoded protein assists RNA splicing of multiple similar introns of different bacterial genes. *J Biol Chem*, 2005, 280(42):35085-8
- [46] Singh RN, Saldanha RJ, D'souza LM, et al. Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J Mol Biol*, 2002, 318(2):287-303
- [47] Huang HR, Chao MY, Armstrong B, et al. The DIVa maturase binding site in the yeast group II intron aI2 is essential for intron homing but not for *in vivo* splicing. *Mol Cell Biol*, 2003, 23(23):8809-19
- [48] Vogel J, Borner T, Hess WR. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res*, 1999, 27(19):3866-74
- [49] Steiner M, Karanatilaka KS, Sigel RKO, et al. Single-molecule studies of group II intron ribozymes. *Proc Natl Acad Sci USA*, 2008, 105(37):13853-8
- [50] Toor N, Keating KS, Taylor SD, et al. Crystal structure of a self-spliced group II intron. *Science*, 2008, 320(5872):77-82
- [51] Saldanha R, Mohr G, Belfort M, et al. Group I and group II introns. *FASEB J*, 1993, 7(1):15-24