核内不均一核糖核蛋白在前体 mRNA 加工中的功能

黄 嘉 陈献华 徐 平*

（复旦大学生命科学学院基因生理学研究室，脑研究中心，上海 200433）

摘要 核内不均一核糖核蛋白(hnRNP)是一类存在于真核生物体内具有类似结构特征的高丰度RNA结合蛋白，一般均匀分布在核内。多种hnRNP具有多样的功能，参与从转录调节，前体mRNA剪接，mRNA输出到mRNA降解等多种生物过程，从而进行基因表达调控。现着重介绍hnRNP在前体mRNA加工过程(加帽，剪接，加尾，输出，选择性降解)中的功能及研究进展。

关键词 核内不均一核糖核蛋白；前体mRNA加工；RNA结合蛋白

编码蛋白质的DNA序列首先经过RNA聚合酶II的作用产生原始的转录物。这些原始转录物经过一系列复杂的过程而被加工成成熟的mRNA并输出到核外进行翻译。这一系列过程被称为前体mRNA加工过程，而这一个过程中的中间阶段转录物(包括原始转录物)长短不一，且分布在核内，所以被称为核内不均一性RNA(heterogeneous nuclear RNA, hnRNA)。

前体mRNA加工过程需要多种蛋白质的参与，在加工过程的不同状态，会由多种不同的蛋白质与RNA构成核糖核蛋白复合体(ribonucleoprotein complex, RNP复合体)。RNP复合体按照结合RNA类型可分为核内不均一性核糖核蛋白复合体(heterogeneous nuclear ribonucleoprotein complex，hnRNP复合体)，小核核糖核蛋白复合体 (small nuclear ribonucleoprotein complex，snRNP复合体)和mRNA-蛋白复合体(mRNA-protein complex, mRNP复合体)这3类。这些复合体的组成蛋白质中，最主要的是RNA结合蛋白。在组成hnRNP复合体的RNA结合蛋白中有一类蛋白质，它们具有类似的结构特征和胞内分布，但又不同于其他类别核蛋白复合体(如snRNP复合体)的固定组成蛋白质，被称为核内不均一性核糖核蛋白(heterogeneous nuclear ribonucleoprotein, hnRNP)[1,2]。研究表明，多种hnRNP活跃于各种不同的RNP复合体中，在转录调节、前体mRNA剪接、mRNA输出、mRNA降解等多种生物过程中担当重要角色。本文将着重于介绍hnRNP在前体mRNA加工过程中的功能及研究进展。

1 hnRNP的分布与结构特征

蛋白质的功能与其分布以及结构特征是密不可分的。hnRNP大都均匀地分布在核内，种类很多，经单克隆抗体沉淀分离并通过双向电泳分析检测出的高丰度的hnRNP多达30种[3]。这些蛋白质都有独立的结合活性以及特异的RNA序列选择性[1]。其中一些hnRNP还通过可变剪接或转录后修饰可产生多种低丰度亚型[3]。

作为RNA结合蛋白，hnRNP都拥有至少一个RNA结合区域，有很多还有一个辅助区域。hnRNP的RNA结合区域常见的有两类：RNA识别模式(RNA-recognition motif, RRM)和KH区(K-homology domain)。两种结合域的三维晶体结构是不同的，这说明它们识别的RNA结合特征序列也是不一样的。

RRM也叫做保守RNA结合区(conserved sequence-RNA binding domain, CS-RBD)，是一段由约80个氨基酸组成序列，中间加插着两段保守序列RNP1 (K/R) G (F/Y) (G/A) F V X (F/Y)和RNP2 (L/I) (F/Y) (V/I) (G/K) (N/G) (L/M)，RNA结合蛋白一般含有1~4个RRM[4]。

KH区首先在hnRNP-K蛋白中被发现，是一段由含有一定疏水性残基的约60个氨基酸组成的序列，序列中央有一段高度保守的VIGXXGXXI序列。目前发现的RBP最多能含有15个拷贝的KH域[4,5]。

hnRNP常见的辅助区域是RGG区，该结构一般存在于蛋白质的C端，是一段重复出现的“精氨酸-甘氨酸-甘氨酸”顺序[6]。RGG区介导蛋白质-蛋白质相互作用，从而影响RNA与蛋白质结合的活性[7,8]。目前已知有RGG区的hnRNP有A1, A2/B1, D1/D2,G, K, P2, R, U。有证据显示RGG区能够影响蛋白质的胞内定位[9]。

收稿日期：2003-11-20 接受日期：2004-01-07

* 通讯作者。Tel/Fax: 021-65642903, E-mail: matibuck@hotmail.com

表1 几种主要的hnRNP的结构与功能简表[2]

蛋白质	域结构	分子量(kDa)	可能功能	穿梭
A1	2XRBD, RGG	34	mRNA 剪接 mRNA 输出 端粒酶合成	+
A2/B1	2XRBD, RGG	36, 38	mRNA 剪接 mRNA 定位	+
C1/C2	1XRBD	41, 43	mRNA 剪接 mRNA 稳定性	-
D(AUF1)	2XRBD, RGG	44~48	端粒酶合成 mRNA 稳定性 重组	+
E1/E2	3XKH	38, 39	mRNA 稳定性 翻译调控	+
F	3XRBD	53	mRNA 剪接	未知
H/H'(DSEF-1)	3XRBD	56	mRNA 剪接 加 poly(A)	未知
I(PTB)	4XRBD	59	mRNA 剪接 mRNA 定位 加 poly(A)	+
K	3XKH, RGG	62	转录 翻译调控	+
L	4XRBD	68	mRNA 输出 mRNA 稳定性	未知
Q	3XRBD, RGG	55~70	mRNA 剪接 mRNA 输出 mRNA 编辑	+
R	3XRBD, RGG	82	未知	未知
U	RGG	120	核内定位	-
HuR	3XRBD	36	mRNA 稳定性 mRNA 输出	+

hnRNP 活跃于各种不同的 RNP 复合体中，参与转录调节，前体 mRNA 剪接，mRNA 输出，mRNA 降解等多种生物过程。表 1 中简述了主要的几种 hnRNP 结构及功能(摘自文献[2]，略有修改)。

2 前体 mRNA 加工过程和 hnRNP 在其中的功能

真核细胞核内前体 mRNA 加工通过 5' 加帽，剪接(移除内含子)，3' 末端切割加尾从而形成成熟的 mRNA，成熟的 mRNA 和 hnRNP 及其他蛋白质形成复合体输出核外再经过选择性降解参与翻译。这些步骤并不是简单的线性顺序，而是在转录物延伸期和转录同时发生的，从而形成一个大型的“生产链”(图 1)[10]。这些步骤大都需要 hnRNP 的参与。

2.1 hnRNP 在前体 mRNA 剪接中的功能

前体mRNA剪接中的主要作用分子是RNA催化分子，主要有 5 种(U1, U2, U4, U5 和 U6)，我们称之为小核 RNA(snRNA)。每一种 snRNAs 都会和至少 7 个蛋白质亚单位形成复合体 snRNP。剪接小体是细胞内进行前体mRNA剪接的大型RNA-蛋白质复合体，而这些 snRNP 则构成了剪接小体的核心[11]。前体 mRNA 剪接存在着两种模式：固定剪接与可变剪接。固定剪接，指在细胞体内以固定的模式发生，剪接方式不受到调节，不会造成外显子略过。可变剪接，与固定剪接不同，通过不同的途径，不同的调节方式，能从同一条基因序列产生一套不相同的 mRNA 序列从而合成功能不同的蛋白质，这使基因编码更有效率。基因的外显子和内含子都存在着一些特异性序列，起着增强或者减弱外显子剪接的作用。这些序列我们分别称为外显子剪接增强子，外显子剪接抑制子，内含子剪接增强子和内含子剪接抑制子。增强子和抑制子分别具有刺激和抑制剪接位点选择的作用。而这种调节作用是通过 RNA 结合蛋白来实现的[12]。

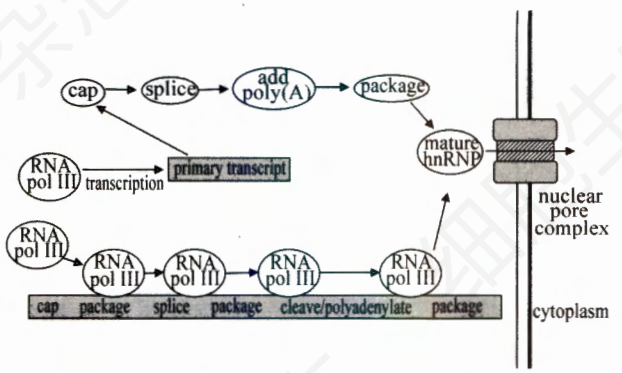

图1 前体 mRNA 处理过程示意图[10]

前体 mRNA 处理过程并非线性，其加帽、剪接、加尾等多个过程是与 RNA 多聚酶 II 的转录并发的。

hnRNP 通过影响增强子或抑制子序列，刺激或抑制剪接位点的选择。最早被发现作用于前体 mRNA 剪接过程中的是 hnRNP-A1。hnRNP-A1 通过浓度调节剪接方式，对传统剪接因子(SF2/ASF)起抑制作用。随着研究的深入，越来越多的 hnRNP 被发现参与剪接过程，包括 hnRNP A/B，C，F，H，I，Q[2,3]。这些 hnRNP 对剪接大都起抑制作用，这很可能是因为大部分 hnRNP 形成 hnRNP 复合体结合于前体 mRNA 内含子上，只有个别 hnRNP 参与形成剪接小体(图 2)[11]，hnRNP 复合体通过与剪接小体竞争性结合 hnRNA 来造成持续性地抑制作用。剪接抑制序列已被证实存在于可变剪接的前体 mRNA 序列中，另外也很可能存在于固定剪接前体 mRNA 序列中[12]。此外，对神经细胞 c-src 基因的研究发现其 N1 外显子的可变剪接能被 PTB/hnRNP-I 抑制，而 hnRNP-F、hnRNP-H 和 KH 类剪接调节蛋白 (KH-homology splicing regulation protein, KSRP) 则对这种剪接表现出增强作用[3]。hnRNP-H 等蛋白质参与的增强作用说明了 hnRNP 能够通过多种途径调节剪接过程。

2.2 hnRNP 在 mRNA 输出中的功能

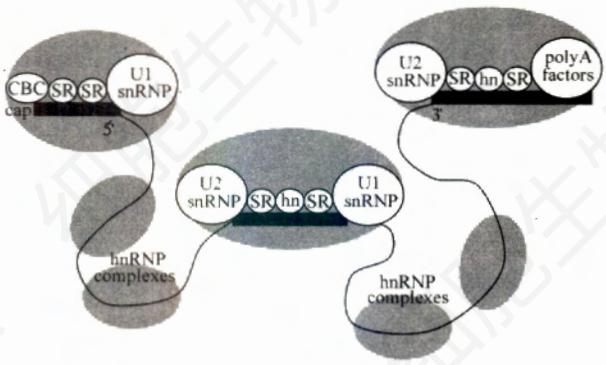

图 2　前体 mRNA 剪接过程示意图[11]

真核生物前体 mRNA 一般具有多个内含子和外显子。SR 蛋白识别外显子上的剪接增强序列，与剪接因子聚集在前体 mRNA 的 5' 或 3' 剪接位点形成剪接小体。除个别的 hnRNP 蛋白(hn)会结合在外显子上，大部分 hnRNP 形成 hnRNP 复合体结合于内含子上。

蛋白质是在核外合成的，因此成熟的 mRNA 分子必须输出到核外，在核糖体的作用下合成蛋白质。mRNA 输出是一个高度选择性的过程。只有经过加帽，剪接，加尾等处理的成熟的 mRNA 分子才会被运输通过。细胞核是双层膜结构，唯一的通道是核孔复合体(NPC)[13]。mRNA 分子通过 NPC 是耗能的，它们无法单独输出核外，而必须形成 mRNA- 蛋白质复合体，通过 NPC 对复合体中核定位信号或核输出信号的识别来完成输出过程[1]。

hnRNP 能够穿梭于核质之间[14,15]。最初的实验发现并证明hnRNP-A1能在融合细胞异种细胞核间穿梭[14,16]。随后又发现很多 hnRNP，如 D, E, I 和 K，也能够在核质之间穿梭。hnRNP-A1 在核内和胞质都被发现与 poly(A)+RNA 结合，这说明穿梭类 hnRNP 很可能同 mRNA 输出相关。对摇蚊(*Chironomus tentans*)的巴氏小体(Balbiani ring, BR)的 RNP 形成和输出的研究发现，人 hnRNP-A/B 的高度同源性蛋白质 C-hrp36 在转录时与 BR 前体 mRNA 结合，参与形成 BR-RNP，伴随 BR 前体 mRNA 一同穿过 NPC 输出到胞外[17]。研究发现 hnRNP-A1 上的一段特殊的氨基酸序列是决定核定位的辅助区域，与经典的核定位序列相似性较低，被称为 M9[18]。M9 未发现具有 RNA 结合特性，但对 hnRNP-A1 的输入和输出均有影响[19,20]。hnRNP-A2 和 B1 都具有与 M9 类似的氨基酸序列。而 hnRNP-K 中的输出信号序列与 M9 不一样，说明不同的 hnRNP 输出途径不尽相同。这些穿梭类 hnRNP 拥有的不同的信号序列很可能就是决定 mRNA 输出的关键。另外一些非穿梭类 hnRNP(比如 hnRNP-C1/C2 和 U)在 mRNA 输出核外之前会与 mRNA 解离。hnRNP-C1/C2 存在一段核停留信号，插入该段信号后，穿梭类 hnRNP 无法输出核外[18,21]。该蛋白质结合内含子 3' 末端和加尾信号下游区域[1]，这说明这类 hnRNP 很可能选择性

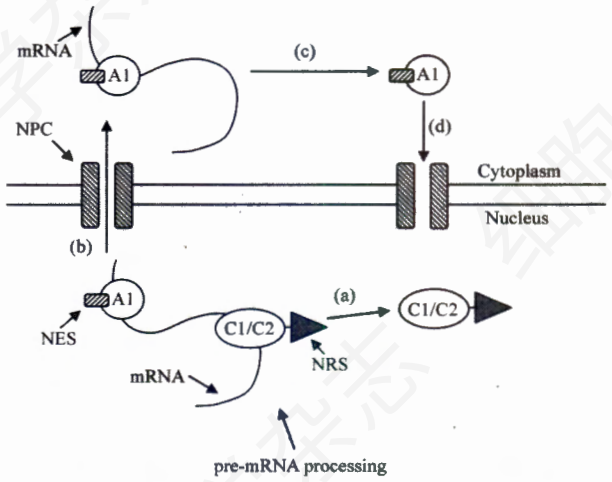

图 3　hnRNP 蛋白参与 mRNA 输出过程的模式图[22]

为了简化，只提到了 hnRNP-C1/C2 与 hnRNP-A1。成熟的 mRNA 输出之前，带有 NRS 的 hnRNP-C1/C2 会与之解离(a)。解离后的 mRNA 通过 hnRNP-A1 NES/M9 的作用通过 NPC(b)。通过后，hnRNP-A1 随即解离(c)。通过新的穿梭途径重新输入到核内参与下一次输出(d)。NES：核输出信号，NRS：核停留信号。

地使未成熟的 mRNA 留在核内。关于 hnRNP 作用于 mRNA 输出的模型见图 3[22]。成熟的 mRNA 输出之前，带有 NRS 的 hnRNP-C1/C2 会与之解离。解离后的 mRNA 通过 hnRNP-A1 NES/M9 的作用通过 NPC 通过后，hnRNP-A1 随即解离通过新的穿梭途径重新输入到核内参与下一次输出。目前仍无确实证据说明 hnRNP 是 mRNA 输出的充分决定条件，但是 hnRNP 在 mRNA 输出中所担当的角色是相当重要的。

2.3　hnRNP 在 mRNA 降解中的功能

各种 mRNA 有着不同的稳定性，它们在体内的生存周期也相差很大(从几分钟到超过 24 小时)[23]。这种差异最终大大地影响了编码蛋白的表达量，因此 mRNA 降解是基因表达调节很重要的一环。已知的 mRNA 降解途径有 4 种：去腺苷化依赖途径、内源性核蛋白酶降解途径、无义介导降解途径和不停顿降解途径[24]。mRNA 分子中包含一些特征序列能被 RNA 结合蛋白识别并结合从而影响并调节 mRNA 的稳定性，选择性地通过上述几种途径降解。

mRNA 分子中包含一些顺式元件能被 RNA 结合蛋白识别并结合从而影响 mRNA 的稳定性。许多类型的 RNA 结合蛋白参与了 mRNA 稳定性的调节，其中有相当一部分是 hnRNP(C, AUF1/hnRNP-D, L, PCBP/hnRNP-E，HuA/HuR，hnRNP-Q 等)[3,23,24]。在众多顺式元件中，对于 3'-UTR 部分的研究较为深入。AU 富集序列(AU-rich element, ARE)是 3'-UTR 影响 mRNA 稳定性最主要的顺式元件[24]。在果蝇中的实验发现 HuR 能结合 ARE，影响含 ARE 的 mRNA 稳定性。HuR 是穿梭类 hnRNP，这说明 ARE 介导的 mRNA 降解能够被 RNA 结合蛋白在胞内的不同分

布所调节。Nab4p(Hrp1p)是酵母中存在的穿梭类hnRNP，被发现具有类似的作用，因此这种作用机制是较保守的。hnRNP-D也能影响mRNA稳定性，研究证明它通过识别3'-UTR一段C富集区(C-rich element, CRE)，与PCBP/αCP等共同作用，影响α-珠蛋白mRNA稳定性[3]。有些mRNA的编码区存在顺式元件影响稳定性，如*c-fos*, *c-myc*和β-微管蛋白。对*c-fos*序列中两个特征序列CRD-1和CRD-2的研究发现，ARE和AUF1的p37亚型以及hnRNP-Q参与形成与这两段序列作用的核蛋白复合体，具体机制尚不清楚[25]。

2.4 hnRNP在前体mRNA 3'末端切割和加尾中的其他功能

mRNA的3'末端加尾是由核酸序列上一段保守的加尾信号控制的。而加尾信号的识别是通过两种RNA结合蛋白CPSF和CstF完成的。CPSF会结合切口上游的AAUAAA序列，而CstF则识别下游的U富集序列。当CstF和CPSF结合到位后，另外两个切割因子CF-1和CF-2结合上来将RNA序列从切割位点切开。然后多聚腺苷聚合酶加入并一次性在切口的3'末端加入200 nt左右多聚腺苷序列。多聚腺苷尾部加好后，多聚腺苷结合蛋白II与之结合，通过某种未知机制决定尾巴的最后长度。加工完成的mRNA于是被输出到胞质[26]。一些hnRNP能够影响mRNA的3'末端形成。前面曾提到过的PTB/hnRNP-I能识别并结合加尾信号上游序列，并可能参与CF-2多聚腺苷位点激活。另外，hnRNP-H'DSEF-1也可能具有切割/加尾因子的功能[3]。而酵母中的一类结构上类似hnRNP A1的蛋白质Hrp1p也被证实是3'端进行切割和加尾的必需因子[27]，这一功能被证实与其精氨酸残基甲基化有关[28]。

3 小结

综上所述，hnRNP在前体mRNA加工过程中具有很重要的功能。这本质上是因为其能与hnRNA结合从而影响hnRNA的结构、通过反式调节影响hnRNA与其他蛋白质的结合的特性。不同的hnRNP协同作用识别不同的hnRNA序列，这决定了其功能的多样性。对hnRNP的研究已经进行了近20年，然而仍有很多问题等待科学工作者们去解决。

参考文献 (References)

[1] Dreyfuss G *et al. Ann Rev Biochem*, 1993, **62**: 289
[2] Dreyfuss G *et al. Nat Rev Mol Cell Biol*, 2002, **3**: 195
[3] Krecic AM *et al. Curr Opin Cell Biol*, 1999, **11**: 363
[4] Burd CG *et al. Science*, 1994, **265**: 615
[5] Siomi H *et al. Nucleic Acids Res*, 1993, **21**: 1193
[6] Kiledjian M *et al. EMBO J*, 1992, **11**: 2655
[7] Pontius BW *et al. Proc Natl Acad Sci USA*, 1990, **87**: 8403
[8] Cartegni L *et al. J Mol Biol*, 1996, **259**: 337
[9] Nichols RC *et al. Exp Cell Res*, 2000, **256**: 522
[10] Bentley D. *Curr Opin Cell Biol*, 2002, **14**: 336
[11] Reed R. *Curr Opin Cell Biol*, 2000, **12**:340
[12] Maniatis T *et al. Nature*, 2002, **418**: 236
[13] Mehlin H *et al. Cell*, 1992, **69**:605
[14] Pinol-Roma S *et al. Nature*, 1992, **355**: 730
[15] Pinol-Roma S *et al. Trends Cell Biol*, 1993, **3**: 151
[16] Borer RA *et al. Cell*, 1989, **56**: 379
[17] Visa N *et al. J Cell Biol*, 1996, **133**: 5
[18] Zenklusen D *et al. FEBS Lett*, 2001, **498**: 150
[19] Weighardt F *et al. J Cell Sci*, 1995, 108: 545
[20] Michael WM *et al. Cell* 1995, **83**: 415
[21] Mayrand SH *et al. Proc Natl Acad Sci USA*, 1993, **90**: 7764
[22] Pinol-Roma S. *Semin Cell Dev Biol*, 1997, **8**: 57
[23] Tourriere H *et al. Biochimie*, 2002, **84**: 821
[24] Brewer G. *Ageing Res Rev*, 2002, **1**: 607
[25] Guhaniyogi J *et al. Gene*, 2001, **265**: 11
[26] Shatkin AJ *et al. Nat Struct Biol*, 2000, **7**: 838
[27] Kessler MM *et al. Genes Dev*, 1997, **11**: 2545
[28] Valentini SR *et al. RNA*, 1999, **5**: 272

The Function of Heterogeneous Nuclear Ribonucleoproteins in pre-mRNA Processing

Jia Huang, Xian-Hua Chen, Ping Xu*

(*Laboratory of Genomic Phyisology and Brain Research Center, Fudan University, Shanghai 200433, China*)

Abstract Heterogeneous nuclear ribonucleoprotein (hnRNP) is a family of proteins that are highly expressed in eukaryotic cells and diffusely distributed in the nucleus. Various hnRNP proteins share similar structural characteristics and play a variety of functions in regulation of gene expression, including transcription, pre-mRNA splicing, mRNA export and mRNA decay. In this review, we briefly summarize the recent progresses regarding the function of hnRNP in pre-mRNA processing.

Key words hnRNP; pre-mRNA processing; RNA-binding protein

Received: November 20, 2003 Accepted: January 7, 2004

*Corresponding author. Tel/Fax: 86-21-65642903, E-mail: matibuck@hotmail.com