International Journal of Bioinformatics Research
ISSN: 0975–3087, E-ISSN: 0975–9115, Vol. 3, Issue 1, 2011, pp-194-199
Available online at http://www.bioinfo.in/contents.php?id=21

PSEUDOKNOTS IN HUMAN SNRNPS

SANJAY KUMAR DEY, SAYAK GANGULI*, PAUSHALI ROY, PROTIP BASU, HIRAK JYOTI CHAKRABORTY AND ABHIJIT DATTA[1]
*DBT-Centre for Bioinformatics, Presidency University, Kolkata
[1]Post Graduate Department of Botany, Bethune College, Kolkata
*Corresponding author. E-mail: sayakbif@yahoo.com

Received: April 05, 2011; Accepted: May 20, 2011

Abstract- Snurps or *small nuclear ribonucleoproteins (*snRNPs*)*, are RNA-protein complexes that combine with unmodified pre-mRNA and various other proteins to form a Spliceosome, comprising of five small nuclear RNAs (snRNAs)—U1, U2, U4, U5, and U6 snRNA—as well as many protein factors, upon which splicing of pre-mRNA occurs. While, RNA pseudoknots play crucial role in protein synthesis by helping in internal ribosome entry, frameshifting, stop codon readthrough in many viral species and the 3'NCR pseudoknots helps viral RNAs to replicate, has been reported by a number of investigators, its presence in human snurps has not yet been done. The present in silico study reveals the presence of pseudoknots in the mRNAs of the proteins associated with human Spliceosome. It not only emphasizes their significance as catalytic RNA world relics but also opens the scope of research in the functional and structural associations of RNA pseudoknots in eukaryotic gene regulation.
Keywords: snRNP, splicing, pseudoknot, gene regulation, catalytic RNA world.

Introduction

Small nuclear ribonucleoproteins (snRNPs) are known to assemble in a stepwise manner onto the pre-mRNA to form the spliceosome. The spliceosome is comprised of five small nuclear RNAs (snRNAs)—U1, U2, U4, U5, and U6 snRNA and many protein factors [1, 2]. The coding regions (exons) of the gene are separated by non-coding DNA (introns) that are not involved in gene expression. These introns are removed by the process known as splicing, resulting in the finally processed mRNA. An RNA pseudoknot is a tertiary structural element formed when bases of a single- stranded loop pair interact with complementary bases outside the loop within the RNA molecule [3, 4]. Though the presences of 14 topologically possible distinct types of pseudoknots have been hypothesized, most commonly occurring pseudoknots are of the H-type, where H signifies a hairpin loop [3]. The other important types are HH type, LL type, HL_{OUT} type and HL_{IN} type etc. [3]. In the case of H-type pseudoknots, bases in a hairpin loop pair with complementary bases outside the loop (Fig.1). HH-type pseudoknot is formed by base pairing of a hairpin loop with another hairpin loop and HL type pseudoknot is formed due to base pairing of a hairpin loop with a single stranded part of a bulge or of an internal or multiple loops (Fig.1). Base pairing of single stranded bulge (B), interior (I) and multibranched (M) loops with complementary regions elsewhere in the RNA also makes pseudoknots. The B, M and I nomenclatures as well as H-H type are not used extensively; rather most structures are referred to as H-type pseudoknots [4].

In eukaryotes, the pseudoknots in 80S ribosomal complex correspond to an L shaped density located at the mRNA exit channel of ribosomal proteins S5 (rsS5). These pseudoknots are important for the formation of the ribosomal A (amino-acyl), P (peptidyl) and E (exit) sites by the interaction of the IRES (internal ribosome entry site) with many key protein component and the 80S complex during protein synthesis [4].

Programmed-ribosomal Frameshifting is a translational mechanism used by many viruses to coordinately express proteins from an mRNA [5]. RNA pseudoknots within the coding regions are mainly responsible for this frame shifting activity. Another important function of the pseudoknot is to facilitate read through of the stop codon, in this regard H-type pseudoknot present at eight nucleotides downstream of the stop codon has been reported to enhance the process [6]. Pseudoknots are also used as binding sites for proteins or single stranded loops of RNA. Indeed pseudoknotted RNAs are often generated during *in vitro* selection of RNA aptamers that bind various biomolecules [4]. Pseudoknotting is also being the most efficient way of folding RNAs in an active conformation (for example- ribozymes) [4].

The role of RNA pseudoknots in translation has been reported in detail, but its presence in human snurps has yet not been revealed. Thus identification of RNA pseudoknots in the mRNAs of the protein part of the human snRNPs will be of excellent biological significance to establish the structural and functional

relationship between the RNA pseudoknots and gene expression strategies in eukaryotes.

Materials and Method

- The DNA sequences for the human snRNPs were downloaded from the Genbank of NCBI website. The sequences used in this study are listed in Table1.
- The DNA sequences were then converted into RNA sequences using an in house tool.
- The RNAs were then analyzed by Pseudoviewer package.

Results

The results of the pseudoknot analysis showed four types of pseudoknots, namely, HL_{OUT}, LL, classical H and HL_{IN} types. The HL_{OUT} type pseudoknot was found to be the most predominant type occurring in 41 of the cases. While LL, classical H and HL_{IN} types of pseudoknots were found 21, 14 and 12 times respectively (Table2, Fig.1).

Discussion and Conclusion

As is evident from the results, several pseudoknot motifs were identified in the human snurps (Table2, Fig.1). The presence of this important structural element is a significant finding in RNA biology as it has been known over the years that pseudoknots are important ubiquitous structures in the mRNA which guides the process of translation through its twists and turns. Furthermore Cech [7] has emphasized that pseudoknot and associated structures formed the core of RNA based gene regulation in the ancient RNA world before the adaptive radiation of proteins in their diverse biological functions. The basics of RNA based gene regulation that exists in the modern RNA world comprises of the riboswitches and the various attenuator and repressor systems found in bacteria and few eukaryotes. The identification and characterization of pseudoknots in the proteins associated with the spliceosome provides further evidences that these pseudoknots are RNA world relics and probably originate from an ancestral stock of RNA regulons. It is important to note that the identification of RNA based regulation of important mammalian gene regulatory systems promotes the fact that we are still under the control of RNA, though there has been the evolution of DNA based genomes.

Acknowledgement

The authors acknowledge the support from Department of Biotechnology, Govt. of India for this work.

References

[1] Kruger K., Grabowski P.J., Zaug A.J., Sands J., Gottschling D.E. and Cech T.R. (1982) C*ell*, 31(1),147-57.

[2] Guerrier-Takada C., Gardiner K., Marsh T., Pace N. and Altman S. (1983) *Cell,* 35, 849−857.

[3] Brierley I., Pennell S. and Gilbert R. J. C. (2007) *Nature Reviews Microbiology,* 5**,** 598-610.

[4] Namy O., Moran S. J., Stuart D. I., Gilbert R. J. C. and Brierley I. (2006) *Nature,* 44, 244-247.

[5] Gluick T. C., Wills N. M, Gesteland R. F. and Draper D. E. (1997) *Biochemistry*, *36* (51),16173–16186.

[6] Gesteland R.F., Cech T.R., Atkins J.F. (eds). (2006) *The RNA world* (Cold Spring Harbor Laboratory, Cold Spring Harbor) 3rd edition, 309.

Table 1- The accession numbers for DNA sequences of the human snRNPs mRNA under study

Serial Number	Accession Number	Description
1	>gi|40217846	ref|NM_014014.2| Homo sapiens small nuclear ribonucleoprotein 200kDa (U5) (SNRNP200), mRNA
2	>gi|31077202	ref|NM_022717.2| Homo sapiens small nuclear ribonucleoprotein 35kDa (U11/U12) (SNRNP35), transcript variant 2, mRNA
3	>gi|212549559	ref|NM_180703.2| Homo sapiens small nuclear ribonucleoprotein 35kDa (U11/U12) (SNRNP35), transcript variant 4, mRNA
4	>gi|57634537	ref|NM_003089.4| Homo sapiens small nuclear ribonucleoprotein 70kDa (U1) (SNRNP70), mRNA
5	>gi|38149990	ref|NM_003091.3| Homo sapiens small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB), transcript variant 2, mRNA
6	>gi|61098048	ref|NM_003094.2| Homo sapiens small nuclear ribonucleoprotein polypeptide E (SNRPE), mRNA
7	>gi|29540552	ref|NM_022805.2| Homo sapiens small nuclear ribonucleoprotein polypeptide N (SNRPN), transcript variant 2, mRNA
8	>gi|29540554	ref|NM_022807.2| Homo sapiens small nuclear ribonucleoprotein polypeptide N (SNRPN), transcript variant 4, mRNA
9	>gi|4507126	ref|NM_003093.1| Homo sapiens small nuclear ribonucleoprotein polypeptide C (SNRPC), mRNA
10	>gi|24474790	emb|AJ505017.1| Homo sapiens mRNA for U5 small nuclear ribonucleoprotein component (SNRP116 gene), trancript variant 2
11	>gi|38149980	ref|NM_198220.1| Homo sapiens small nuclear ribonucleoprotein polypeptide B'' (SNRPB2), transcript variant 2, mRNA
12	>gi|71143122	ref|NM_152551.3| Homo sapiens small nuclear ribonucleoprotein 48kDa (U11/U12) (SNRNP48), mRNA
13	>gi|164565439	ref|NM_024571.3| Homo sapiens small nuclear ribonucleoprotein 25kDa (U11/U12) (SNRNP25), mRNA
14	>gi|38150006	ref|NM_198216.1| Homo sapiens small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB), transcript variant 1, mRNA
15	>gi|29540556	ref|NM_003097.3| Homo sapiens small nuclear ribonucleoprotein polypeptide N (SNRPN), transcript variant 1, mRNA
16	>gi|29540553	ref|NM_022806.2| Homo sapiens small nuclear ribonucleoprotein polypeptide N (SNRPN), transcript variant 3, mRNA
17	>gi|50593001	ref|NM_003090.2| Homo sapiens small nuclear ribonucleoprotein polypeptide A' (SNRPA1), mRNA
18	>gi|29294623	ref|NM_177542.1| Homo sapiens small nuclear ribonucleoprotein D2 polypeptide 16.5kDa (SNRPD2), transcript variant 2, mRNA
19	>gi|29294622	ref|NM_004597.4| Homo sapiens small nuclear ribonucleoprotein D2 polypeptide 16.5kDa (SNRPD2), transcript variant 1, mRNA
20	>gi|28416941	ref|NM_006938.2| Homo sapiens small nuclear ribonucleoprotein D1 polypeptide 16kDa (SNRPD1), mRNA
21	>gi|39725715	ref|NM_004596.3| Homo sapiens small nuclear ribonucleoprotein polypeptide A (SNRPA), mRNA
22	>gi|83776586	ref|NM_003095.2| Homo sapiens small nuclear ribonucleoprotein polypeptide F (SNRPF), mRNA
23	>gi|36497	emb|X52979.1| Human gene for small nuclear ribonucleoproteins SmB and SmB'
24	>gi|53690153	gb|AY742712.1| Homo sapiens small nuclear ribonucleoprotein polypeptide B mRNA, complete cds
25	>gi|705365	gb|AH003282.1|SEG_HUMSNRP0 Homo sapiens small nuclear ribonucleoprotein (SNRPD1) gene, complete cds
26	>gi|37539	emb|X07401.1| Human mRNA for hU1-70K snRNP protein (RNP6)
27	>gi|5870128	gb|AH008174.1|SEG_HSSNRPB Homo sapiens small nuclear ribonucleoprotein (SNRPB) gene, alternative splice products, complete cds
28	>gi|15217058	gb|AF400432.1|AF400432 Homo sapiens small nuclear ribonucleoprotein polypeptide N (SNRPN) mRNA, complete cds
29	>gi|2708306	gb|AF016370.1|AF016370 Homo sapiens U4/U6 small nuclear ribonucleoprotein hPrp3 mRNA, complete cds
30	>gi|2708304	gb|AF016369.1|AF016369 Homo sapiens U4/U6 small nuclear ribonucleoprotein hPrp4 mRNA, complete cds
31	>gi|338249	gb|M21251.1|HUMSNRNP01 Human Alu repeats in the region 5' to the small nuclear ribonucleoprotein E gene
32	>gi|37536	emb|X06818.1| Human mRNA for hU1-70K snRNP protein (RNPH2, part 1)
33	>gi|37533	emb|X06817.1| Human mRNA for hU1-70k snRNP protein (RNPH1, part 1)
34	>gi|37532	emb|X06816.1| Human mRNA for hU1-70K small nuclear RNP protein (RNP3)
35	>gi|37530	emb|X06815.1| Human mRNA for hU1-70K small nuclear RNP protein (RNP FL1.7) (nonproductive)
36	>gi|37528	emb|X06814.1| Human mRNA for hU1-70K small nuclear RNP protein (RNP12)
37	>gi|37524	emb|X06812.1| Human mRNA for hU1-70K snRNP protein (RNP8)
38	>gi|37522	emb|X06811.1| Human mRNA for U1-70K snRNP protein (RNP11)
39	>gi|337442	gb|AH001532.1|SEG_HUMRNP70 Human small nuclear ribonucleoprotein (U1-70K) gene
40	>gi|337446	gb|M22636.1|HUMRNP70A Human U1 small nuclear ribonucleoprotein 70 kd protein mRNA, complete cds

Table 2- Accession numbers and description of predominant types of Pseudoknots found in Human snurps

Sl. No.	Accession No	Nt. Region	Type of pseudoknot
1.1	>gi|40217846|ref|NM_014014.2|	1-1711	None
1.2	Do	1712-1955	HL_{IN}-1 HL_{OUT}-1
1.3	Do	1956-1276	HL_{OUT}-1
1.4	Do	1277-1560	HL_{OUT}-1
1.5	Do	1561-2979	None
1.6	Do	2980-3264	HL_{OUT}-1
1.7	Do	3265-3547	HL_{OUT}-1
1.8	Do	3548-3827	LL-1
1.9	Do	3828-4111	None
1.10	Do	4112-4394	H type-2 HL_{IN}-1
1.11	Do	4395-4675	None
1.12	Do	4676-4958	LL-1
1.13	Do	4959-5243	HL_{IN}-1
1.14	Do	5244-5528	HL_{IN}-1 HL_{OUT}-2
1.15	Do	5529-6196	None
2.1	>gi|31077202|ref|NM_022717.2|	1-568	None
2.2	Do	569-997	HL_{IN}-1
3.1	>gi|212549559|ref|NM_180703.2|	1-285	HL_{OUT}-1
3.2	Do	286-1343	None
4.1	>gi|57634537|ref|NM_003089.4|	1-529	None
5.1	>gi|38149990|ref|NM_003091.3|	1-284	H-1 HL_{IN}-1
6.1	>gi|61098048|ref|NM_003094.2|	1-284	HL_{OUT}-1
6.2	Do	285-566	LL-1 HL_{OUT}-1
6.3	Do	567-1134	None
6.4	Do	1135-1417	HL_{OUT}-1
6.5	Do	1418-1611	LL-1
7.1	>gi|29540552|ref|NM_022805.2|	1-569	None
8.1	>gi|29540554|ref|NM_022807.2|	1-284	LL-1
8.2	Do	285-565	None
8.3	Do	566-849	HL_{OUT}-1
8.4	Do	850-1420	None
8.5	Do	1421-1780	H type-1
9.1	>gi|4507126|ref|NM_003093.1|	1-279	LL-1
9.2	Do	280-564	HL_{IN}-1
9.3	Do	565-753	LL-1
10.1	>gi|24474790|emb|AJ505017.1|	1-281	None
10.2	Do	282-566	HL_{OUT}-1
10.3	Do	567-847	H-1
10.4	Do	848-1844	None
10.5	Do	1845-2202	LL-1
10.6	Do	2203-2951	None
11.1	>gi|38149980|ref|NM_198220.1|	1-353	LL-1
11.2	Do	354-701	None
11.3	Do	702-986	H-1 HL_{OUT}-1
11.4	Do	987-1208	None
12.1	>gi|71143122|ref|NM_152551.3|	1-1063	None
12.2	Do	1064-1415	HL_{OUT}-1
12.3	Do	1416-1766	None
12.4	Do	1767-2119	LL-1
12.5	Do	2120-3536	None
12.6	Do	3537-4188	HL_{OUT}-1

12.7	Do	4189-4494	HL_{OUT}-1
13.1	>gi|164565439|ref|NM_024571.3|	1-358	LL-1
13.2	Do	359-712	None
13.3	Do	713-1124	HL_{OUT}-1
14.1	>gi|38150006|ref|NM_198216.1|	1-354	LL-1
14.2	Do	355-711	H-1 LL-1
14.3	Do	712-1036	LL-1
15.1	>gi|29540556|ref|NM_003097.3|	1-354	LL-1
15.2	Do	355-1715	None
16.1	>gi|29540553|ref|NM_022806.2|	1-1044	None
16.2	Do	1045-1399	H-1
16.3	Do	1400-1626	None
17.1	>gi|50593001|ref|NM_003090.2|	1-357	None
17.2	Do	358-711	HL_{OUT}-1
17.3	Do	712-1088	None
18.1	>gi|29294623|ref|NM_177542.1|	1-354	HL_{OUT}-1
18.2	Do	355-707	HL_{OUT}-1
18.3	Do	708-804	None
19.1	>gi|29294622|ref|NM_004597.4|	1-425	LL-1
19.2	Do	426-790	HL_{OUT}-1
20.1	>gi|28416941|ref|NM_006938.2|	1-711	None
20.2	Do	712-1066	HL_{OUT}-1
20.3	Do	1167-1633	None
21.1	>gi|39725715|ref|NM_004596.3|	1-780	None
21.2	Do	781-1133	LL-1
21.3	Do	1134-1646	None
22.1	>gi|83776586|ref|NM_003095.2|	1-427	H-1
22.2	Do	428-824	None
23.1	>gi|36497|emb|X52979.1|	1-423	HL_{OUT}-1
23.2	Do	424-959	None
24.1	>gi|53690153|gb|AY742712.1|	1-355	HL_{OUT}-1
24.2	Do	356-717	None
25.1	>gi|705365|gb|AH003282.1|	1-427	None
25.2	Do	428-855	HL_{IN}-1 HL_{OUT}-1
25.3	Do	856-2131	None
25.4	Do	2132-2548	HL_{OUT}-1
26.1	>gi|37539|emb|X07401.1|	1-717	None
26.2	Do	718-1027	HL_{OUT}-1
27.1	>gi|5870128|gb|AH008174.1|	1-423	H-1 HL_{OUT}-1
27.2	Do	424-1717	None
27.3	Do	1718-2145	HL_{OUT}-1
27.4	Do	2146-3857	None
27.5	Do	3858-4285	HL_{OUT}-1
27.6	Do	4286-5137	None
27.7	Do	5138-5510	H-1
28.1	>gi|15217058|gb|AF400432.1|AF400432	1-835	None
29.1	>gi|2708306|gb|AF016370.1|AF016370	1-428	H-1
29.2	Do	429-1271	None
30.1	>gi|2708304|gb|AF016369.1|AF016369	1-1259	None
30.2	Do	1260-1683	HL_{OUT}-1
30.3	Do	1684-2188	None
31.1	>gi|338249|gb|M21251.1|HUMSNRNP01	1-426	HL_{OUT}-1
31.2	Do	427-758	None
32.1	>gi|37536|emb|X06818.1|	1-140	None
33.1	>gi|37533|emb|X06817.1|	1-257	None
34.1	>gi|37532|emb|X06816.1|	1-531	H-1 HL_{OUT}-1

35.1	>gi|37530|emb|X06815.1|	1-851	None
35.2	Do	852-1184	HL_{OUT}-1
35.3		1185-1593	None
36.1	>gi|37528|emb|X06814.1|	1-358	None
36.2	Do	359-718	HL_{IN}-1
37.1	>gi|37524|emb|X06812.1|	1-423	LL-1
37.2	Do	422-710	None
37.3	Do	711-918	HL_{OUT}-1
38.1	>gi|37522|emb|X06811.1|	1-400	LL-1
39.1	>gi|337442|gb|AH001532.1|SEG_HUMRNP70	1-429	None
39.2	Do	430-858	H-1 HL_{IN}-1
39.3	Do	859-1283	None
39.4	Do	1284-1639	LL-1
39.5	Do	1640-2063	LL-1
39.6	Do	2064-2417	None
39.7	Do	2418-2772	HL_{OUT}-1
39.8	Do	2773-3836	None
39.9	Do	3486-3837	HL_{OUT}-1
39.10	Do	3838-4194	HL_{OUT}-1
39.11	Do	4195-4556	HL_{OUT}-1
39.12	Do	4557-4985	HL_{OUT}-1
39.13	Do	4986-6267	None
39.14	Do	6268-6667	HL_{IN}-1
40.1	>gi|337446|gb|M22636.1|	1-496	HL_{OUT}-1
40.2	Do	497-995	None
40.3	Do	996-1358	HL_{IN}-1
40.4	Do	1359-1690	None

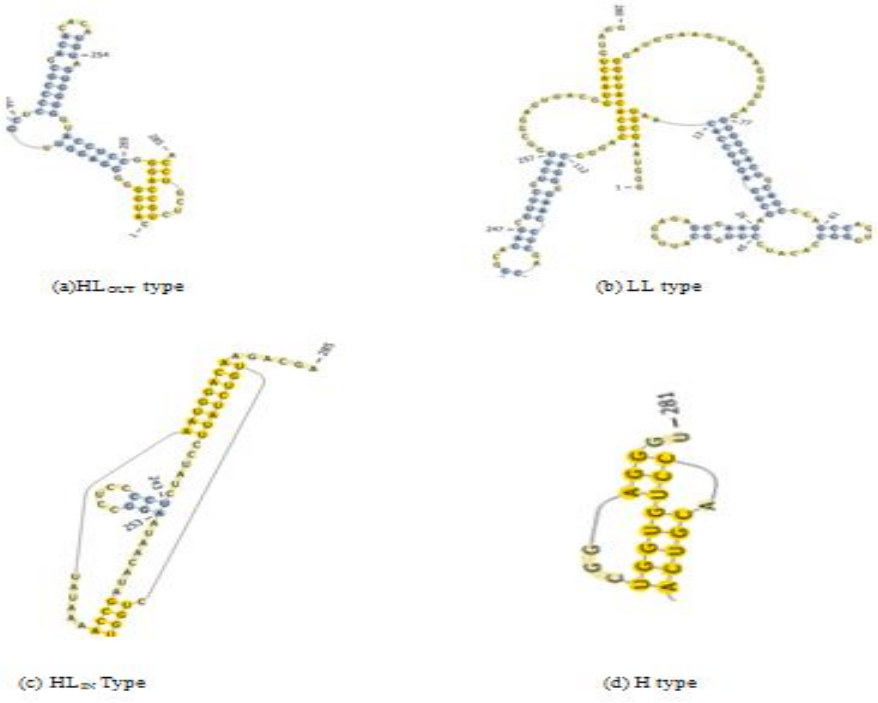

Fig. 1 (a-d): Four major types of pseudoknots found in Human snRNPs.