Stable RNA hairpins in 88 coding regions of human mRNA*

PAN Min[1]**, WANG Chuanming[1]** and LIU Ciquan[1,2,3]***

(1. Modern Biological Center, Yunnan University, Kunming 650091, China; 2. Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China; 3. Center for Theoretical Biology, Peking University, Beijing 100871, China)

Received March 1, 2004; revised March 19, 2004

Abstract RNA hairpins containing UNCG, GNRA, CUUG (N=A, U, C or G, R=G or A) loops are unusually thermodynamic stable and conserved structures. The structural features of these hairpin loops are very special, and they play very important roles *in vivo*. They are prevalent in rRNA, catalytic RNA and non-coding mRNA. However, the 5′ C(UUCG)G 3′ hairpin is not found in the folding structure of 88 human mRNA coding regions. It is also different from rRNA in that there is no preference for certain sequences among tetraloops in these 88 mRNA folding structures.

Keywords: thermodynamic stability, hairpin, UNCG, GNRA, RNA structure.

Biology's "central dogma" states that genetic information flows from DNA to RNA, then to protein. Since the initial discovery of catalytic RNA, numerous studies have shown that RNA does more than simply serve this intermediary function[1]. RNA plays a very important role *in vivo*, and RNA's function is intimately linked with RNA folding structure. Many RNA structures have been determined by NMR and X-ray diffraction. RNA structures deposited in Protein Data Bank (PDB) and Nucleic Acid Database (NDB) are accumulating at an ever increasing rate. Computer simulations have also been widely used on RNA structural research. People are rethinking the role of RNA[1].

RNA hairpins are the common structural characteristics of all RNAs. Hairpin loop containing four nucleotides is called tetraloop. Some tetraloops are characterized by exceptionally high thermodynamic stability. The great majority of these tetraloops belong to three classes: UNCG, GNRA, CUUG (N=A, U, C or G; R=G or A)[2~6]. The structures of several stable RNA hairpins have revealed networks of stabilizing interactions within the hairpin loop: non-Watson-Crick base pairs and base-phosphate and base-sugar contacts[6]. Stable RNA hairpins define nucleation sites for folding[7], determine tertiary interactions in RNAs[8,9], and are recognized by RNA-binding proteins[6].

UNCG and GNRA are often found with an exceptional frequency in rRNA, catalytic RNA and mRNA non-coding region[6,10]. However, whether these stable tetraloops are prevalent in all RNAs is still not clear, especially in mRNA coding region.

For the important role of mRNA folding structure in eukaryotic gene expression we studied the folding structures of 88 human mRNA coding regions. The UUCG tetraloop, and the ultra stable hairpin loop, are not found within these mRNA coding regions.

1 Materials and methods

A total of 88 human mRNA sequences were chosen from the Integrated Sequence-Structure Database (ISSD) (http://www.protein.bio.msu.su/issd/) which comprises the coding sequences of mRNAs, amino acid sequences and structural parameters of the corresponding proteins. The ISSD version 1.0 and 2.0 collect 118 records totally. The records whose structural resolutions are worse than 2.8 Å (≥2.8 Å) were discarded. The records whose amino acid sequences are different from those in PDB and homologous records were also removed. At last, 88 mRNA coding sequences were chosen as our samples (Table 1).

* Supported by the National Natural Science Foundation of China (Grant Nos. 90208018, 30160036) and the Tian Yuan Fund of Mathematics (Grant No. A0324101)

** The authors contribute equally to this work

*** To whom correspondence should be addressed. E-mail: liucq@ynu.edu.cn

Table 1. 88 samples chosen from ISSD

	PBD_ID		PBD_ID		PBD_ID		PBD_ID		PBD_ID		PBD_ID		PBD_ID		PBD_ID
1	1AAP_A	12	1DLH_A	23	1HCN_A	34	1HUR_A	45	1LIT	56	1PSN	67	1TTA_A	78	2HHB_A
2	1ABM_A	13	1DYN_A	24	1HCQ_A	35	1ICE_A	46	1LPB_B	57	1RBP	68	1UBQ	79	2HHM_A
3	1AII	14	1ERT	25	1HDR	36	1ILK	47	1LPE	58	1REX	69	1ULA	80	2HMB
4	1ALD	15	1ESL	26	1HDX_A	37	1ILR_1	48	1LYA_A	59	1RHP_A	70	1VCA_A	81	2HNP
5	1ANG	16	1FIB	27	1HFC	38	1IOB	49	1MHL_A	60	1RTG	71	1VHR_B	82	2ILK
6	1APY_A	17	1FIL	28	1HIK	39	1JKW	50	1MIL	61	1SAC_A	72	2ACH_A	83	2KNT
7	1BHS	18	1FNA	29	1HMP_A	40	1JLM	51	1NUE_A	62	1SPD_A	73	2ALR	84	2TGI
8	1CDW_A	19	1FUJ_A	30	1HSA_A	41	1KPA_A	52	1PBW_A	63	1SRA	74	2CBA	85	3CD4
9	1CKS_B	20	1GIF_A	31	1HSB_B	42	1KRN	53	1PHT	64	1TNF_A	75	2CPL	86	3GRS
10	1CLL	21	1GUH_A	32	1HUL_A	43	1LCF	54	1POD	65	1TPK_A	76	2FKE	87	4FGF
11	1CSB_A	22	1HCL	33	1HUP	44	1LCL	55	1PPB_H	66	1TSR_A	77	2GMF_B	88	4I1B

UNCG and GNRA tetraloops occur very often in rRNA, catalytic RNA and the non-coding region of mRNA. In order to know whether these RNA hairpins are also distributed within the coding region of mRNA, we searched UNCG and GNRA sequences in 88 coding regions of human mRNAs, and used the software RNAstructure 3.71 on these mRNA coding sequences. There are two ways to study the folding structures of these mRNAs. The first way is dividing the mRNA coding sequence with sequential increasing length, which starts with 90 nucleotides (nt), the number of nucleotides going up each time by 30 (1 ~ 90 nt, 1 ~ 120 nt, 1 ~ 150 nt, 1 ~ 180 nt ……) until the full-length of each mRNA coding sequence was reached. Then we regarded each subsection as a folding unit and used the RNAstructure 3.71 to fold these subsections. The second way is dividing each mRNA coding region with 90 nt length, exon length, and full length respectively. Each subsection was folded as a folding unit by RNAstructure 3.71.

2 Results

2.1 The distribution of UNCG and GNRA sequences within mRNA coding sequences

In our 88 mRNA sequences (42819 nt), there were 44 samples containing UUCG sequences. The sequence UUCG is only found 75 times totally. The sequences of other members of UNCG family are also seldom found (Table 2). However, the sequence GNRA is found much more times than the sequence UNCG, especially the sequences GAAA, GAGA and GGAA (Table 2).

In all 42819 nt sequences, there are only 19 UUCG sequences flanked by at least one potential base-pair (Table 3). Most of these 19 sequences are flanked by only one potential base-pair. Obviously, these segments cannot form RNA hairpin with the sequence UUCG in the loop. Other four UUCG sequences within 1NSK_R, 1SAC_A, 1SRA have three potential base-pairs at two terminals (Table 3).

Table 2. UNCG, GNRA occurrence times in coding mRNA sequences and hairpin structures

Family	Sequence	Occurrence times in sequence	Occurrence times in hairpin structure			
			Common hairpin	90 nt divided	Full-length folding	Exon divided
UNCG	UUCG	75	0	0	1	0
	UACG	60	0	2	1	0
	UCCG	88	0	2	2	1
	UGCG	75	0	0	0	0
GNRA	GAAA	305	3	7	10	5
	GAGA	275	1	9	3	4
	GGAA	289	1	5	1	0
	GCAA	186	1	2	5	1
	GUAA	57	0	1	1	0
	GGGA	221	1	2	2	3
	GCGA	55	0	2	0	1
	GUGA	206	1	1	1	0
Number of tetraloop hairpins			57	319	190	162

Sequential searches suggested that all 88 mRNA sequences (42817 nt) have only four segments containing UUCG nucleotides that may form 3-base-pair stem hairpins with the UUCG in the loop.

Table 3. UUCG mRNA sequences which have at least one possible base-pair at two terminals

PDB-ID	Location	Sequence
1BHS	478	CGCCAGCAAGUUCGCGCUCGAAGG
1CLL	412	CUAUGAAGAAUUCGUACAGAUGAU
1CSB_A	94	CUGCUGGGCCUUCGGGCUGUGGA
1FIL	401	GCCUCCCACCUUCGGCGUUCCCAG
1GIF_A	148	GCUCAUGGCCUUCGGCGGCUCCAG
1HDX_A	107	GCUUAUGAAGUUCGCAUUAAGAUG
1HUR_A	367	CCUCCUGGUGUUCGCCAACAAGCA
1ICE_A	234	AUCUCACUGCUUCGGACAUGACUA
1JLM	542	CAGAACCAGCUUCGGGAGAAGAUC
1NSK_R	308	CCAGGCACCAUUCGUGGGACUUC
1PSN	43	UAUGGAGUACUUCGGCACUAUCGG
1PSN	301	CAAUCAGAUCUUCGGCCUGAGCGA
1SAC_A	110	ACCUUGUGUUUUCGAGCCUAUAGU
1SPD_A	58	CAUCAUCAAUUUCGAGCAGAAGGA
1SRA	352	CACCCGCUUUUUCGAGACCUGUGA
1SRA	412	GGCCGGCUGCUUCGGCAUCAAGCA
1TSR_A	356	AGAAACACUUUUCGACAUAGUGUG
2HHB_A	272	GCGCACAAGCUUCGGUGGACCCG
2HMB	46	CAGCAAGAAUUUCGAUGACUACAU

The sequence is given in the 5′ to 3′ direction of mRNA. The potential base-pairs are indicated by underlines. The PDB-IDs containing UUCG sequences which have three potential base-pairs at two terminals are shown in bold.

2.2 The distribution of some tetraloops within the coding region of mRNA

Each sample with sequential increasing length (1 ~ 90 nt, 1 ~ 120 nt, 1 ~ 150 nt, 1 ~ 180 nt ……until the full-length) was folded by using RNAstructure 3.71. We got 1428 folding subsections and 1953 hairpins from all samples totally. There are some hairpins that can exist stably in all sequential increasing subsections. We named these hairpins "common hairpins". Tetraloops arise 57 times in all common hairpins, while they belong to 46 kinds of tetraloops. GNRA family is found 8 times, but UNCG family is not found (Table 2). There is no preference for certain sequences among these tetraloops in common hairpin.

Besides in the "common hairpin", there are many tetraloops in mRNA folding structures which are folded by using the two ways mentioned before. However, the UGCG tetraloop is not found, and the UACG and UCCG tetraloops that are found a few times cannot exist stably (Table 2). The UUCG tetraloop arises only one time in full-length folding structure of 1NSK_R. This UUCG hairpin is closed by AU base-pair, and its stem contains only three base-pairs (Fig. 1). Full length of the 1NSK_R mRNA is 450 nucleotides, and UUCG is located from 308 to 311 nt. A hairpin with the UUCG loop should form in 1 ~ 330 nt folding structure, but it does not arise. With the increasing of sequence (1 ~ 360 nt, 1 ~ 390 nt, 1 ~ 420 nt), a UUCG hairpin does not form until to the 1 ~ 450 nt folding structure (full length). It is suggested that this UUCG hairpin cannot exist stably in the folding structure of 1NSK_R. Although sequential searches suggested that the three sequences within 1SAC_A and 1SRA mRNA also form potential UUCG hairpins (Table 3), the UUCG hairpin is also not found in their sequential folding process.

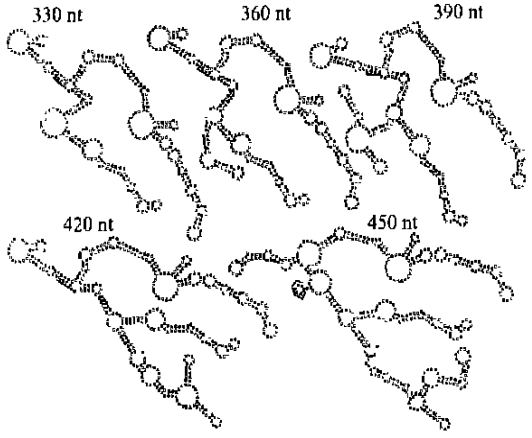

Fig. 1. Folding structure of 1NSK_R sequential folding (from 330 nt to full-length 450 nt). The location of UUCG sequences is indicated with black lines.

There are several GNRA tetraloops in 88 mRNA folding structures which were folded by using the two ways, but GNRA is no more prevalent than other tetraloops (Table 2).

3 Discussion

3.1 Stable RNA hairpins

Thermodynamic studies of RNA and DNA loop showed that the stability of hairpin depends on the size and sequence of the loop[2,6].

Between RNA tetraloops of identical size but different sequence, the thermodynamic parameters showed that certain RNA hairpins containing the loop sequence, UNCG and GNRA tetraloop, are more stable than those containing UUUG, GCUU, UUUU, AAAA, CCCC tetraloops and so on[3~5]. The GNRA sequence confers an additional stability of $\Delta G^{o} \approx -1$ kcal/mol over the stability of homopyrimidine loops[6]. However, there is evidence indicating that the UNCG loops are more stable than the GNRA

loops[4, 5]. Stabilities of these RNA molecules are affected by the closing base pair. Substitution of a G C for a C G closing base-pair decreases the stability of RNA hairpins. The C(UUCG)G loop is most stable in tetraloops and the C(UUCG)G hairpin loop (with a stem of three additional base pairs) has a melting temperature approximately 20 ℃ higher than a hairpin with purportedly normal thermodynamic stability, such as the G(UUUU)C hairpin[3~6].

The thermodynamic stabilities of UNCG loops and GNRA loops are based on their unusual conformations.

The structure of UUCG loop is very special. There are peculiar interactions between nucleotides in the UUCG loop (Fig. 2). These features include: two C2′-*endo* ribose puckers at U2 and C3; a *syn* conformation about the glycosidic bond of G4; the exposure of the U2 base to the solvent; hydrogen bonding between the U_1pU_2 phosphate oxygen atom and the C3 exocyclic amino group[11]; a U1-G4 pair defined by a hydrogen bond between U1 ribose 2′-OH and the carbonyl oxygen atom of G8 and a bifurcated hydrogen bond between the U1 O2 carbonyl oxygen atoms and both the imino and amino hydrogen atoms of G8[12]. There are only two U2 C3 unpaired nucleotides in the loop. The 2′-hydroxyl groups are responsible for the observed extra stability characterizing this UUCG loop[13, 14].

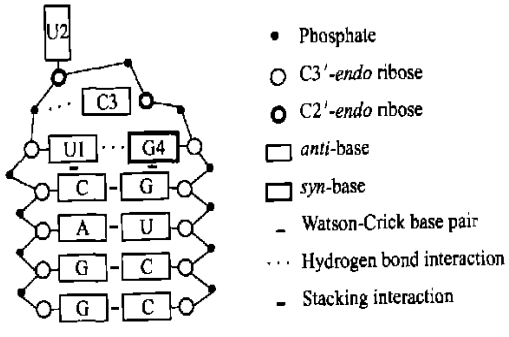

Fig. 2. Schematic secondary structure diagram of the UUCG tetraloop showing nucleotide conformations and critical interactions determined by NMR[16].

In prokaryotic ribosomal RNAs, most UUCG tetraloops are closed by a C G base-pair[10]. However, this preference is greatly reduced in eukaryotic rRNA species where many UUCG tetraloops are closed by G C base-pairs[15]. Although CG-closed UUCG hairpin is more stable than GC-closed[3~5], essential loop structural features of them are similar[15]. NMR results suggest the presence of significant conformational flexibility in the G(UUCG)C loop structure at temperatures far below the global melting of the molecule, in contrast to the behavior observed for the C(UUCG)G molecule[15].

In 1991, Heus and Pardi[17] determined the structures of GCAA and GAAA loops. From then on, some other GNRA tetraloops were also determined by NMR and X-ray diffraction. Structural features of GNRA family members are similar. The G of the GNRA is stacked on the 5′ side of the stem and the NRA is stacked on the 3′ side of the stem[18]. The glycosidic angles for all residues are in an *anti* conformation, and the loop is additionally stabilized by a G A base-pair between the first and the last loop nucleotide, which is similar to UNCG loop[17]. In addition to G A base-pair (hydrogen bond interactions are G_1 H_2 to A_4 N_7 and G_1 N_3 to A_4 H_6), the stability of the loop is also dependent on a network of hydrogen bonds surrounding the G A base-pair. In GAGA, GCAA and GAAA hairpins, the same pattern of potential hydrogen bonds are observed and are schematically drawn in Fig. 3. They are, G_1 H_2 to A_4 OP, G_1 H_1 to A_4 OP, G_1 O_2' to A_4 H_6, G_1 2′OH to the six and seven positions of nucleotide 3. The same overall structural motif in all three hairpins is found. Small sequence-dependent changes in structure are observed between these three-loops (GAGA, GCAA, GAAA); however, with slightly different hydrogen bonding patterns and stacking[18]. These changes do not destroy the GNRA structural motif.

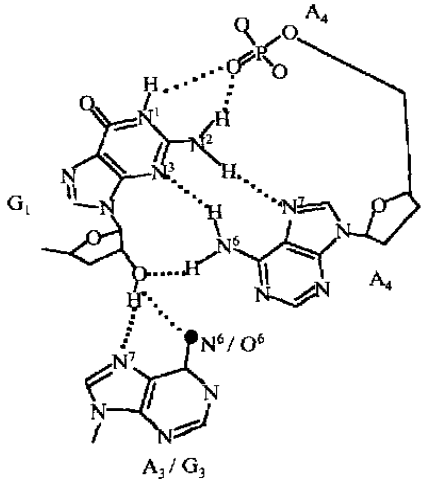

Fig. 3. Schematic drawing of the hydrogen bonding network proposed for $G_1N_2R_3A_4$ tetraloops[18]. N2 is omitted. The hydrogen-bonding interactions in loop mainly happen between G1, R3(A3/G3), A4. Dot lines indicate hydrogen bond.

3.2 Biological functions of stable RNA hairpins

Stable RNA hairpins play very important roles *in vivo*. Stable GNRA and UNCG tetraloops always act as nucleation sites during RNA folding and help to suppress alternative secondary structures[7]. GNRA tetraloops and their receptors which are particularly common in natural RNAs and interactions between them have been identified in the structures of group I and II introns and RNAase P[9]. GNRA and UNCG loops can interact with proteins[6, 15]. C(UUCG)G hairpin stability can cause reverse transcriptase to terminate[3]. In addition, the high thermodynamic stability of C(UUCG)G hairpin makes them ideal tools for stabilizing nucleic acid structures[6].

3.3 Ubiquity of stable hairpins in rRNA, catalytic RNA and the non-coding region of mRNA

RNA tetraloops are particularly common *in vivo*, especially in rRNAs. The phylogenetic analysis of hairpin loop sequences in rRNA of archaea and bacteria revealed tetra-loops account for about 55% of all hairpin loops in 16SrRNA and 38% in 23SrRNA[10]. Although 256 (4^4) different tetraloop sequences are possible, the three classes: UNCG, GNRA and (more rarely) CUUG frequently occur in RNA hairpin tetraloops, and 70% of these three loops are UNCG and GNRA[10].

Besides within rRNA, these RNA tetraloop sequences are also found abundantly in other RNAs. The C(UUCG)G hairpins are prevalent in the intercistronic regions (non-coding regions) of bacteriophage T4 mRNA[3]. One of two major hairpins in human U6 small nuclear RNA is a C(UUCG)G hairpin[3]. The hairpins IV of chicken, rat and human U1 snRNAs are G(UUCG)G hairpin[19]. Chicken, rat and human U4 snRNAs contain UACG tetraloop hairpins[20]. Most of the *E. coli* rho-independent transcription terminators are tetraloops and the most abundant loop sequences are UUCG and GAAA[21]. M1 RNA, the catalytic subunit of *E. coli* RNase P, contains five tetraloops, of which one has the sequence UNCG and four have GNRA[22]. The 4.5S RNA of eukaryote contains a GAAA hairpin[23]. Self-splicing RNA Group I and II introns also contain UNCG and GNRA hairpins[17, 24].

3.4 UUCG hairpin is not found in the 88 coding regions of human mRNA

Although these stable hairpins are found abundantly in many kinds of RNAs, our results indicated that the stable UUCG hairpin may not be universal in the coding regions of human mRNA.

The distributions of UNCG and GNRA sequences in mRNA coding sequences have revealed that the possibility of forming UUCG tetraloop in mRNA coding regions is remote. Furthermore, the analysis of the folding structures of 88 human mRNA coding regions also showed that UUCG hairpin loop forms only one time within the 1953 hairpins which were folded by using the first sequential folding way. Other members of UNCG family arise a few times (Table 2), but these UNCG tetraloops do not exist stably in sequential folding process, and could not exist stably *in vivo*, either.

The analysis of sequences and tetraloops in 88 human coding mRNA showed that the GNRA sequence and tetraloop are more prevalent than UNCG sequence and tetraloop.

Tetraloops arise 57 times in common hairpins totally, and there are 46 kinds of tetraloops. GNRA and UNCG tetraloops account for about 14% of all tetraloops in "common hairpin", which is much smaller than the proportion (70%) in rRNA. There is no strong preference for certain tetraloop sequences in 90 nt and exon length-divided folding structure. All of these phenomena led to the suggestion that the folding structures of 88 human mRNAs have no preference for certain tetraloop sequences among the tetraloops, which is different from rRNA.

3.5 The possible mechanism

In 1988 Tuerk et al.[3] first discovered the ultra thermodynamic stability of C(UUCG)G hairpin, and this hairpin can stop DNA synthesis by reverse transcriptase. They suggested that the termination of reverse transcriptase results from its inability to denature C(UUCG)G hairpin, and then the process of DNA synthesis was terminated[3]. In addition, Gene 60 of bacteriophage T4 contains a nucleotide sequence that interrupts the coding sequence and is not translated into gene product. This sequence is not spliced out of the message but is folded into a secondary structure, containing a C(UUCG)G hairpin, that allows the ribosome to skip ≈50 nt bases and continue translation[25]. This C(UUCG)G hairpin is skipped, and is not read by ribosome. These phenomena may explain why the UUCG hairpin is absent in the 88

coding regions of human mRNA.

So we propose that the stable UUCG hairpin, if it existed, would hinder the movement of ribosome along mRNA, and most of mRNA fragments which contain UUCG hairpins or sequences could mainly be distributed in 3′-UTR, 5′-UTR and intron.

4 Results and outlook

Since mRNA molecules are large, and always combined with proteins tightly *in vivo*, mRNAs are very difficult to extract, purify and crystallize under current experimental conditions. Furthermore, because people still do not realize the importance of mRNA coding region's structure, the distribution of stable hairpin in mRNA coding region is still an unsolved question. Based on our initial study on the folding structure of mRNA coding regions, we found that there are not favorite tetraloop hairpins in 88 mRNA coding regions, and the C(UUCG)G hairpin is not found, either.

UUCG hairpin is prevalent in rRNA, catalytic RNA and mRNA non-coding region, but we can not find any stably existent UUCG hairpin in these coding regions of mRNA. It is a very interesting finding. We are trying to find the reason of this phenomenon, and we hope that our study will be helpful to explore the function of mRNA folding structure in the coding region.

References

1 Dennis C. The brave new world of RNA. Nature, 2002, 418: 122.
2 Chastain, M. et al. Structural elements in RNA. Progress in Nucleic Acid Research and Molecular Biology, 1991, 41: 131.
3 Tuerk, C. et al. CUUCGG hairpins: Extraordinarily stable RNA secondary structures associated with various biochemical processes. Porc Natl Acad Sci USA, 1988, 85: 1364.
4 Antao, V. P. et al. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. Nucleic Acids Research, 1992, 20(4): 819.
5 Antao, V. P. et al. A thermodynamic strudy of unusually stable RNA and DNA hairpins. Nucleic Acids Research, 1991, 19(20): 5901.
6 Varani, G. Exceptionally stable nucleic acid hairpins. Annual Review Biophys Biomol Struct. 1992, 24: 379.
7 Molinaro, M. et al. Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications. Nucleic Acids Research, 1995, 23(15): 3056.
8 Conn, G. L. et al. RNA structure. Current Opinion in Structural Biology, 1998, 8: 278.
9 Leontis, N. B. et al. Analysis of RNA motifs. Current Opinion in Structural Biology, 2003, 13: 300.
10 Woese, C. R. et al. Architecture of ribosomal RNA: Constraints on the sequence of "tetra-loops". Porc Natl Acad Sci USA, 1990, 87: 8467.
11 Chong, C. et al. Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. Nature, 1990, 346(16): 680.
12 Allain, F. H. T. et al. Structure of the P1 helix from Group I self-splicing introns. Journal of Molecular Biology, 1995, 250: 333.
13 Williams, D. J. et al. Unrestrained stochastic dynamics simulations of the UUCG tetraloop using an implicit solvation model. Biophysical Journal, 1999, 76: 3192.
14 Williams, D J, et al. Experimental and theoretical studies of the effects of deoxyribose substitutions on the stability of the UUCG tetraloop. Journal of Molecular Biology, 2000, 297: 251.
15 Williams, D. J. et al. Experimental and computational studies of the G[UUCG]C RNA tetraloop. Journal of Molecular Biology, 2000, 297: 1045.
16 Miller, J. L. et al. Theoretical studies of an exceptionally stable RNA tetraloop: observation of convergence from an incorrect NMR structure to the correct one using unrestrained molecular dynamics. Journal of Molecular Biology, 1997, 270: 436.
17 Heus, H. A. et al. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. Science, 1991, 253: 191.
18 Jucker, F. M. et al. A network of heterogeneous hydrogen bonds in GNRA tetraloops. Journal of Molecular Biology, 1996, 264: 968.
19 Branlant, C. et al. The conformation of chicken, rat and human U1A RNAs in solution. Nucleic Acids Research, 1981, 9(4): 841.
20 Comolli, L. R. et al. NMR structure of the 3′ stem-loop from human U4 snRNA. Nucleic Acids Research, 2002, 30(20): 4371.
21 Carafa, Y A, et al. Prediction of Rho-independent *Escherichia coli* transcription terminators: A statistical analysis of their RNA stem-loop structures. Journal of Molecular Biology, 1990, 216: 835.
22 Krummel, D. A. P. et al Verification of phylogenetic predictions *in vivo* and the importance of the tetraloop motif in a catalytic RNA. Porc Natl Acad Sci USA, 1999, 96: 11200~11205.
23 Ikawa, Y. et al. A comparative study on two GNRA-tetraloop receptors: 11-nt and IC3 motifs. J Biochem, 2001, 130(2): 251.
24 Uhlenbeck, O. C. Tetraloops and RNA folding. Nature, 1990, 364(16): 613.
25 Huang, W. M. et al. A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. Science, 1988, 239: 1005.