ISSN 1682-296X (Print)
ISSN 1682-2978 (Online)

Bio Technology

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Biotechnology 4 (3): 194-201, 2005
ISSN 1682-296X
© 2005 Asian Network for Scientific Information

Two Pseudoknots from Human C2H2 Zinc Finger Genes Discovered by Computational Structure Analysis

Shaowu Meng and Zhuo Zhang
Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613

Abstract: Pseudoknots have been extensively identified in viruses. However, very few pseudoknots in Human genome have been previously reported. In this study, we identify two pseudoknots in mRNAs of two separate C2H2 zinc finger (C2H2-ZNF) genes after studying the entire 226 C2H2-ZNF genes in Human Chromosome 19 through using standard tools in an original manner, such as searching Pfam database and analyzing the primary structure of these genes and the secondary structure of their putative mRNAs. The two pseudoknots are different in structure and one of them is very likely to stimulate +1 programmed ribosomal frameshifting. We also found some interesting characteristics of the two genes: each gene has significant C2H2-ZNF domains (e-2 or better) in two ORFs of different reading frames; the two ORFs each have a promoter, a transcription start site, a start codon, a Kozak pattern, a Poly(A) site and a Poly(A) signal and thus each of them can be viewed as a gene. Moreover, we compare these two pseudoknots with the structures of the reported pseudoknots in mammals. Present discovery would be useful in the deeper understanding of structural characteristics and functions of pseudoknots.

Key words: Pseudoknot, structure analysis, structure comparison, function inference, human C2H2-ZNF genes, human chromosome 19

INTRODUCTION

A hairpin is an antiparallel duplex structure that forms by pairing of inverted repeat sequences within a single-stranded RNA (or DNA). In a hairpin, the helical section(s) is called the stem and the unpaired base segment(s) at the end of the structure is called the loop. A pseudoknot is a triple-stranded RNA structure formed when the loop at the top of a hairpin has sequences complementary to an unpaired segment near (or at) the base of the stem.

Therefore, such a pseudoknot consists of two stems (usually named Stem 1 and Stem 2) and two loops (usually named Loop 1 and Loop 2) from the 5′ to the 3′ directions. Stringent topological rules are often used to determine the probability of pseudoknots formation[1]. Pseudo-knotted RNA structures have been extensively identified in viruses, including all types of natural RNAs, such as genomic RNAs[2] structural RNAs in ribonucleoprotein complexes[3], rRNAs[4], mRNAs[5], tmRNA[6] and catalytic and self-splicing RNAs[7].

Pseudoknots can play many roles in cells, such as structural components[3] and functional regulation[7,8], but many pseudoknots in known cases serve as stimulators to cause a programmed ribosomal frameshifting by forcing ribosomes to impede and then slip into another reading frame[9-11]. A programmed ribosomal frameshifting is that a ribosome shifts a reading frame at a particular mRNA site to yield a new fused protein encoded by two overlapped open reading frames[12].

The known cases about the ribosomal frameshifting are either +1 programmed ribosomal frameshifting or –1 programmed ribosomal frameshifting, in which a ribosome reads into the Reading Frame +1 or the Reading Frame –1 by slipping one base toward the 3′ direction or toward the 5′ direction, respectively[9,13,14]. The reported patterns of frameshifting sites are X XXY YYZ (Here, X, Y and Z represent any nucleotide) mainly in viruses and retroviruses[15,16] UUU Ynn (Here, Y is pyrimidine, such as C, T and U; -n- represents any nucleotide) in both eubacteria and mammalian cells[17] or -CUU-ANN-N- (Here, -N- represents any nucleotide) in the yeast *Saccharomyces cerevisiae*[18]. Note that some mutant pseudoknots cannot efficiently stimulate frameshifting due to the deviation of a normal conformation of a functional pseudoknot[19].

Corresponding Author: Dr. Shaowu Meng, Department of Biology, University of Toronto at Mississauga, 3359 Mississauga Road North, Mississauga, Ontario L5L 1C6, Canada
Tel: 01 905 569 4400

However, only a few previous studies report on the discovery of pseudoknots in Mammals. Wills[20] proposed that a functional RNA pseudoknot is found in mRNAs of Human prion genes (Fig. 1A) and inferred his pseudoknot stimulated programmed ribosomal frameshifting. Usually, a prion gene contains five copies of a 24-nucleotide repeat that is highly conserved among species. Barrette *et al.*[21] supported this discovery as they found pseudoknots in 76 mammalian mRNAs of prion genes after comparative sequence analysis and pattern searching (Fig. 1B). However, based on the fact that all strains of the scrapie form of the prion protein consist of the same polypeptide sequence, Barrette *et al.*[21] further inferred that the prion pseudoknots cannot stimulate programmed ribosomal frameshifting, but could function as a site to promote nucleation and thus may interfere with proper translation and production of the prion protein[22].

Matsufuji *et al.*[23] reported a pseudoknot in the mRNA of Rat Antizyme 1 genes and thought this pseudoknot stimulates a +1 programmed ribosomal frameshifting (Fig. 2A). Ivanov *et al.*[24] also proposed a pseudoknot in the mRNA of Human Antizyme 2 genes (Fig. 2B and C) and inferred his pseudoknot stimulates a +1 ribosomal frameshifting, too. Shigemoto *et al.*[14] inferred there may be a pseudoknot in an mRNA of Mouse Edr (embryonal carcinoma differentiation regulated) genes and considered that this pseudoknot causes a -1 programmed ribosomal frameshifting (Fig. 3). An Edr gene is a single copy mammalian gene and bears a CCHC zinc finger motif in the first reading frame and a putative aspartyl protease catalytic site motif in the second reading frame. Importantly, both of the two motifs are retroviral-like[14]. A sequence pattern of a CCHC zinc finger motif is CX2CX4HX4C where, X can be any amino acid. As the last example, Ben-Asouli *et al.*[8] reported an interesting pseudoknot located in upstream of the AUG initiation codon in Human interferon-γ mRNA and proposed that this pseudoknot acts to adjust translation of the interferon-γ mRNA by activating the interferon-inducible protein kinase, PKR.

C2H2 zinc finger (C2H2-ZNF) genes make up one of the largest Human gene families with over 700 members in the Human genome[25]. In Human chromosome 19, 226 repetitive C2H2-ZNF genes aggregate into 11 clusters[26]. An interesting side is that a C2H2-ZNF protein has many repetitive C2H2-ZNF domains in a tandem array[26]. In a zinc finger domain of a C2H2-ZNF protein, the two conserved cysteines (C) and the two conserved histidines (H) bind one atom of Zn in a tetrahedral array to yield a finger-like projection. A sequence pattern of a C2H2-ZNF domain is #-X-C-X(1-5)-C-X3-#-X5-#-X2-H-X(3-6)-H/C. (Here, X represents any amino acid, # denotes hydrophobic residues and the numbers indicate the numbers of residues. H/C says this position's amino acid is either histidines or cysteines). C2H2-ZNF proteins can bind both RNA and DNA and may be the original nucleic acid-binding proteins[27] usually playing a critical role as transcriptional regulators to regulate the expression of other genes by mediating interactions between DNAs and proteins[28].

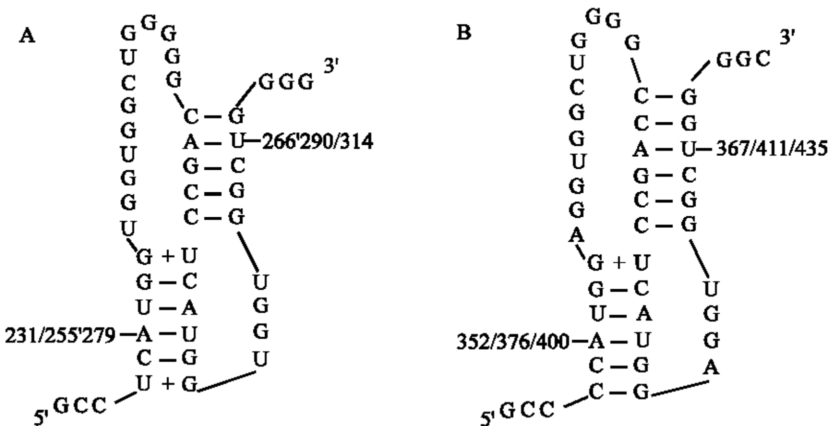

Fig. 1: Prion pseudoknot secondary structures cited from Barrette *et al.*[21] (A) The Human pseudoknot described by Wills[20], (B) The pseudoknot found in cattle

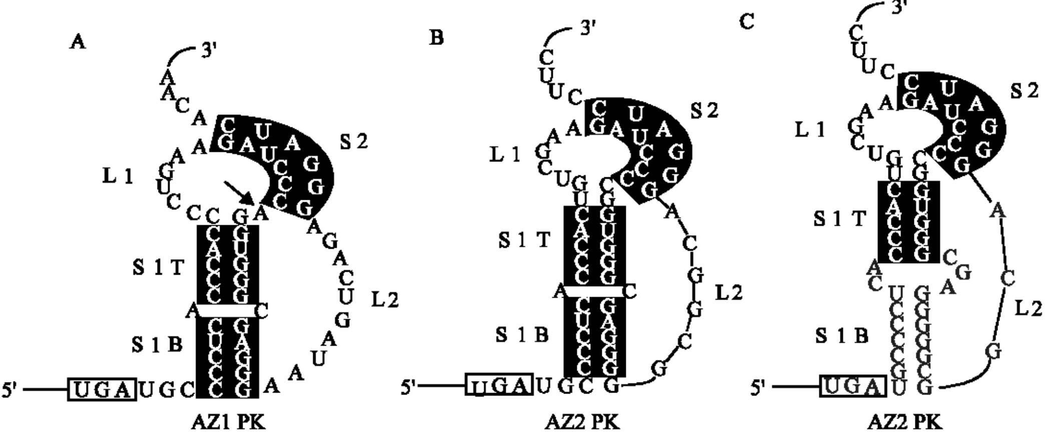

Fig. 2: Antizyme pseudoknot secondary structures cited from Ivanov *et al.*[24] S1B is the bottom of Stem 1; S1T is the top of Stem 1; L1 is Loop 1; L2 is Loop 2; AZ1PK is the Antizyme 1 pseudoknot; AZ2PK is the Antizyme 2 pseudoknot. Shaded nucleotides indicate potential base pairing completely conserved between Antizyme 1 and Antizyme 2. (A) A pseudoknot structure of the Rat Antizyme 1 defined by Matsufuji *et al.*[23] (B) A potential pseudoknot structure of the Human Antizyme 2. (C) An alternative pseudoknot structure of the Human Antizyme 2

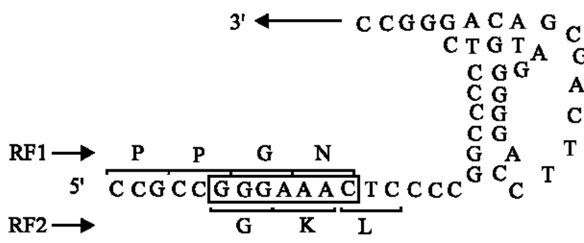

Fig. 3: The Edr pseudoknot secondary structure cited from Shigemoto *et al.*[14]. The heptameric slippage sequence is denoted in box. RF1 indicates the original sequence and RF2 indicates the frameshiftted sequence

The C2H2-ZNF gene with above 700 copies in Human genome is more repetitive than the prion gene with only five copies. Besides, both C2H2-ZNF motif and CCHC-ZNF motif are short zinc finger domains. Therefore, it is interesting to investigate whether there is a pseudoknot in the mRNA of a C2H2- ZNF gene since pseudoknots can play many important roles in cells, such as structural components[3] and functional regulation[7,8]. In this study, we use similarity search and structure analysis to study the secondary structures of putative mRNAs of all C2H2-ZNF genes in Human Chromosome 19, aiming to identify pseudoknots. Present discovery is helpful in the deeper understanding of structural characteristics and functions of pseudoknots.

MATERIALS AND METHODS

As the initial step, we downloaded the DNA sequence of the total 226 C2H2-ZNF genes in Human chromosome 19 from the website http://bahama.jgi-psf.org/pub/ch19/. For each of the 226 genes, then we used sequential Steps from A to E to test whether there is a pseudoknot in its putative mRNA and used Steps F and Step G to test whether it is possible for the pseudoknot to stimulate programmed ribosomal frameshifting.

Step A: To find out ORFs in different reading frames (that is, Reading Frame 0, +1, +2, -0, -1, -2; Reading Frame 0 is an original sequence, Reading Frame +1 is a sequence after deleting one nucleotide from Reading Frame 0, Reading Frame +2 is a sequence after deleting two nucleotides from Reading Frame 0 and indicates the reverse and complementary of a sequence) of a C2H2-ZNF gene, we use ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) to detect the ORF in each reading frame.

Step B: To detect C2H2-ZNF domains in each reading frame, we translate the six reading frames of this gene into six amino acid sequences using a program Translate (http://tw.expasy.org/tools/dna.html) and then search Pfam (http://www.sanger.ac.uk/Software/Pfam/) using HMMER (http://hmmer.wustl.edu/) as searching tool and the six amino acid sequences as query sequences. The cutoff threshold is set as GA.

Step C: Then, we identify long ORFs (with more than 23 amino acids) that sit in different reading frames and that also contain significant C2H2-ZNF domains. If a gene cannot meet this condition, we return to Step A to examine another gene. Otherwise we proceed to Step D.

Step D: We analyse the 5' and 3' non-coding regions of the long ORFs to determine whether these ORFs each have a promoter, a transcription start site, a Poly(A) site and a Poly(A) signal, a start codon and a Kozak pattern. The first four elements are decided by using software in the public website http://research.i2r.a-star.edu.sg/promoter/DGE_main.htm[29]; the last two elements are decided by using ATG^{pr} [30] in the public website http://www.hri.co.jp/atgpr/. If these elements exist, the ORFs can be structurally considered as genes and we proceed to Step E. If such structural elements for these ORFs do not exist, we return to Step A to test next gene.

Step E: We infer mRNAs by deleting introns (the sequences between domains) separately from the two long ORFs (derived from Step C) in different reading frames of this gene[31]. Subsequently, we analyze the secondary structure of a putative mRNA, aiming to find a pseudoknot in the mRNA of this gene. We use GeneBee-NET[32] (http://www.genebee.msu.su/services/rna2_reduced.html) to accomplish this analysis. The main use of this tool is to analyze the secondary structure of the first sequence in an alignment (other sequences in the alignment function as comparisons), then this tool can find out significant pseudoknots[32]. We use this tool and analyze the secondary structure of the mRNA only in Reading Frame 0 (original sequence). If no pseudoknot is found, we return to Step A to examine next gene. If a pseudoknot exists, we continue to next step.

Step F: This step is to check whether those ORFs (in one gene) from Step C are overlapping by locating these ORFs. If the ORFs are overlapping, we proceed to next step. If the ORFs are not overlapping, we check next gene from Step E.

Step G: To check whether there is a slippery site around the pseudoknot from Step E, we search these patterns X XXY YYZ, UUU Ynn and CUU-ANN-N in the same mRNA where the pseudoknot is found.

Finally, we compare the pseudoknot identified in this study with those pseudoknots found in mammals by other researchers so that we can further explore the structure and function of a pseudoknot.

RESULTS

Only two genes can pass Step E. This indicates that the mRNAs of the two genes have significant pseudoknots. These two genes are denoted Z3-2 and Z3-7, which are the second gene and the seventh gene, respectively, in the third cluster of the C2H2-ZNF genes in Human chromosome 19. Z3-2 has 1338 nucleotides and Z3-7 has 2769 nucleotides. Both of them are located at clone 21743738. However, only Z3-7 goes through the Step G. Z3-2 passes Step F, but cannot pass Step G. That is, both Z3-2 and Z3-7 have overlapping ORFs in different reading frames. A likely frameshifting site, UUUCCU, is found five upstream nucleotides before the pseudoknot of Z3-7. We have not found a frameshifting site around the pseudoknot of Z3-2.

Each of the two genes has two ORFs that are in different reading frames but that contain significant same C2H2-ZNF domains with e-values of e-2 or better (Table 1). Here, we denote the ORF in Reading Frame 0 as ORF1 and the ORF in Reading Frame +1 as ORF2.

After detecting the 5' and 3' non-coding regions of the two ORFs as per Step D, we found:

- In gene Z3-2, the ORF1 spans from position 91 to position 1233, with Kozak pattern CXXATGG, Poly(A) signal GATATAAAAA and with transcription start site at position –285 (i.e., the 285th nucleotide upstream the first position of Z3-2), Poly(A) site at position 1322; the ORF2 spans from position 470 to position 1402, with Kozak pattern GXXATGG, Poly(A) signal GATATAAAAA and with transcription start site at position -285, Poly(A) site at position 2115.

Table 1: C2H2-ZNF domains within ORFs in distinct reading frames of Z3-2 and Z3-7

Gene	Reading frame	ORF	ORF-from	ORF-To	Score	E-value
Z3-2	0	ORF1	307	375	6.2	1.3
	0	ORF1	475	543	29.3	5.9e-06
	0	ORF1	559	627	30.8	2.1e-06
	0	ORF1	643	711	25.3	9.3e-05
	0	ORF1	727	795	31.0	1.7e-06
	0	ORF1	811	879	27.7	1.7e-05
	0	ORF1	895	963	30.4	2.6e-06
	0	ORF1	979	1047	35.5	8.1e-08
	0	ORF1	1063	1131	33.2	4.0e-07
	0	ORF1	1147	1215	29.4	5.5e-06
	+1	ORF2	1211	1279	32.5	6.4e-07
Z3-7	0	ORF1	1513	1581	14.0	0.22
	0	ORF1	1597	1665	19.7	0.0046
	0	ORF1	1681	1746	13.7	0.23
	0	ORF1	1762	1830	27.4	2.2e-05
	+1	ORF2	2066	2134	25.9	6.0e-05
	+1	ORF2	2150	2218	8.9	0.71
	+1	ORF2	2234	2302	19.0	0.0075

Reading frame 0 is the original sequence; Reading frame +1 is the sequence after deleting one nucleotide from Reading Frame 0. ORF-from means the start position of an open reading frame; ORF-To indicates the end position

- In gene Z3-7, the ORF1 spans from position 1465 to position 1869, with Kozak pattern AXXATGc, Poly(A) signal ACATGAAAAG and with transcription start site at position 1249, Poly(A) site at position 1961; the ORF2 spans from position 1541 to position 2371, with Kozak pattern tXXATGa, Poly(A) signal CTAACA AATG and with transcription start site at position 1249, Poly(A) site at position 2424.
- For each of the two genes, its two ORFs share a promoter and a transcription start site, but have different start codons, different Kozak patterns, different Poly(A) sites and different Poly(A) signals. In view of their primary structures, each of the two ORFs can be considered as a gene.

After studying the secondary structure of putative mRNAs of the two genes as per Step F, we found a significant pseudoknot in Z3-2 with an energy -1.200000 Kkal/mol, which consists of two stems and two loops. Its Stem 1 has five base pairs, Stem 2 has three base pairs, Loop 1 has one base and Loop 2 has nine bases (Fig. 4); we also found a significant pseudoknot in Z3-7 with an energy -2.000000 Kkal/mol, which consists of two stems and two loops, too. Its Stem 1 has seven base pairs, Stem 2 has four base pairs, Loop 1 has four bases and Loop 2 has six bases (Fig. 5).

We also compare the two pseudoknots in this study with those pseudoknots reported by others. The results are in Table 2.

DISCUSSION

An e-value is considered to be significant for domain searching if it has a value of 5e-2 or better[27]. Although one domain in ORF1 of Z3-2 with e-value 1.3, two domains in ORF1 of Z3-7 with e-values 0.22 and 0.23 and one domain in ORF2 of Z3-7 with e-value 0.71, are not significant, the other domains, including nine C2H2-ZNF domains within ORF1 and one such domain within ORF2 in Z3-2 and two C2H2-ZNF domains in OFR1 and two such domains in ORF2 in Z3-7, are greatly significant with e-values much less than the standard 5e-2 (Table 1). It indicates that these data in our paper are reliable.

In general, a gene contains a 5' flanking region (including a promoter and a transcription start site), an ORF and a 3' flanking region (including a Poly(A) signal and a Poly(A) site). Especially in eukaryotic organisms, a gene is called a split gene since its exons are separated by the noncoding introns[33]. Usually, a mature mRNA of a gene of a eukaryotic organism is mainly composed of exons[31]. As seen in the previous section, each ORF of

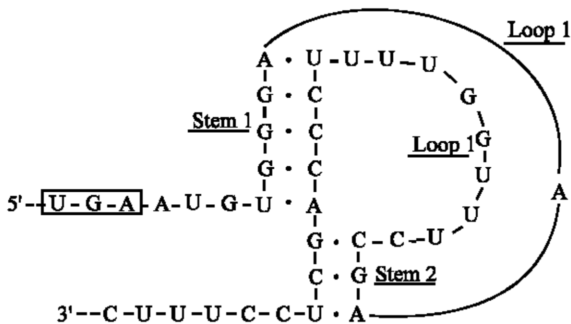

Fig. 4: The secondary structure of an mRNA of the C2H2-ZNF gene Z3-2

the two genes (Z3-2 and Z3-7) can be viewed as a gene because it has a complete primary structure of a gene, including a promoter, a transcription start site, a Kozak pattern, a Poly(A) signal and a Poly(A) site. And thus, it is very likely to deduce a mature mRNA from such an ORF by deleting those introns. Moreover, GeneBee-NET (used in Step F) is a program to analyze the secondary structure of RNAs[32]. This program can also identify significant pseudoknots. These facts suggest that our method is reasonable.

The structure of a pseudoknot is a key factor to determine its function. Different structures may result in quite different functions of pseudoknots. According to a study on the pseudoknots in mRNAs of bacteriophage T2, the nucleotide number in Stem 2 is relatively stable in pseudoknots of close organisms and the nucleotide number in Loop1 and Loop 2 is more variable[34]. Present study (Table 2) also indicates that in mammals, the nucleotide number in Stem 2 is relatively stable in pseudoknots and the nucleotide number in Loop1 and Loop 2 is more variable.

Moreover, Chen *et al.*[35] studied retroviral RNA pseudoknots and had an idea that a structural feature-one unpaired nucleotide at the junction between Stem 1 and Stem 2-is required for an efficient pseudoknot. Kang *et al.*[19] supported this idea through studying mouse mammary tumor virus and further reported that the pseudoknot cannot stimulate frameshifting if there are two unpaired nucleotides at its junction. However, within reported pseudoknots in mammals, only the pseudoknot in mRNA of antizyme 1 has one unpaired nucleotide at the junction between Stem 1 and Stem 2 (Table 2). It indicates the pseudoknots reported in mammals are slightly different with virus pseudoknots in structure.

As we has mentioned above, the structures of pseudoknots in Z3-2 and Z3-7 are different. The pseudoknot in the mRNA of Z3-2 has five base pairs in Stem 1, three base pairs in Stem 2, one base in Loop 1 and

Table 2: Comparison of the reported pseudoknots in Mammals

Researchers	Stem 1	Stem 2 (P)	Junction (S)	Loop 1 (S)	Loop 2 (S)	Organisms	Genes
Wills	6P	5	0	13	4	Human	Prion
Barrette *et al.*[21]	6P	6	0	12	4	Cattle	Prion
Matsufuji *et al.*[23]	12P, 2S	6	1	6	9	Rat	Antizyme 1
Ivanov *et al.*[24]	13P, 2S	6	0	5	6	Human	Antizyme 2
Ivanov *et al.*[18]	14P, 5S	6	0	5	3	Human	Antizyme 2
Shigemoto *et al.*[14]	5P, 2S	4	2	1	8	Mouse	Edr
Ben-Asouli *et al.*[8]	7P	4	4	41	7	Human	Interferon-γ
Z3-2, this paper	5P	3	0	1	9	Human	C2H2-ZNF
Z3-7, this paper	7P	4	0	4	6	Human	C2H2-ZNF

Junction is nucleotides to connect Stem 1 with Stem 2; P indicates base pairs and S denotes single base.

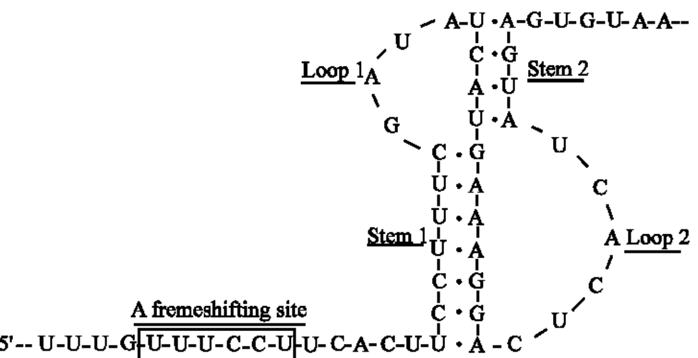

Fig. 5: The secondary structure of an mRNA of the C2H2-ZNF gene Z3-7

nine bases in Loop 2 (Fig. 4 and Table 2); the pseudoknot in the mRNA of Z3-7 has seven base pairs in Stem 1, four base pairs in Stem 2, four bases in Loop 1 and six bases in Loop 2 (Fig. 5 and Table 2). However, the common feature of the two pseudoknots is no unpaired nucleotide at the junction between Stem 1 and Stem 2.

This difference in structure of pseudoknots of Z3-2 and Z3-7 possibly causes their divergence in pseudoknots' function. Z3-7 has two overlapping ORFs separately in Reading Frame 0 and Reading Frame 1 (Table 1) and a likely frameshifting site close to its pseudoknot. These facts indicate that this pseudoknot of Z3-7 may be functional to stimulate +1 programmed ribosomal frameshifting since a programmed ribosomal frameshifting is determined mainly by a frameshifting-prone site, which often is stimulated by a downstream pseudoknot[13,15]. However, we have not found a possible frameshifting site around the pseudoknot of Z3-2. This indicates that the pseudoknot of Z3-2 may not stimulate a programmed ribosomal frameshifting. We will study the function of the pseudoknot of Z3-2 in further work.

Interestingly, the prion pseudoknots proposed by Wills[20] and by Barrette *et al.*[21] are also slightly different (Fig. 1 and Table 2). The pseudoknot from Wills[20] has six base pairs in Stem 1, five base pairs in Stem 2 (Fig. 1A), but the pseudoknot from Barrette *et al.*[21] has six base pairs in both Stem 1 and Stem 2 (Fig. 1B). There is no unpaired nucleotide at the junction between Stem 1 and Stem 2 in pseudoknot models from both Wills[20] and Barrette *et al.*[21]. The functions of these two prion pseudoknots are different. Wills[20] inferred his pseudoknot stimulated programmed ribosomal frameshifting, but Barrette *et al.*[21] thought that his prion pseudoknots could function as a site to promote nucleation and thus may interfere with proper translation and production of the prion protein[22].

The pseudoknots reported by Matsufuji *et al.*[23] in Antizyme 1 and Ivanov *et al.*[24] in Antizyme 2 are also different (Fig. 2 and Table 2). Matsufuji *et al.*[23] reported that the pseudoknot in Rat Antizyme 1 has 12 base pairs and two single unpaired bases in Stem 1, six base pairs in Stem 2, six bases in Loop 1 and nine bases in Loop 2 (Fig. 2A). Ivanov *et al.*[24] reported that the pseudoknot in Human Antizyme 2 has 13 base pairs and two single unpaired bases in Stem 1, six base pairs in Stem 2, five bases in Loop 1 and six bases in Loop 2 (Fig. 2B). In the same article, Ivanov *et al.*[24] also put forward an alternative structure for his pseudoknot, with 14 base pairs and five single unpaired bases in Stem 1, six base pairs in Stem 2, five bases in Loop 1 and three bases in Loop 2 (Fig. 2C). Note that there is an unpaired adenine at the junction between Stem 1 and Stem 2 in the pseudoknot model from Matsufuji *et al.*[23] but no unpaired nucleotide at the junction between Stem 1 and Stem 2 in the pseudoknot model from Ivanov *et al.*[24].

Though these pseudoknot models about Antizymes are different, Ivanov *et al.*[24] found some conserved base pairs between Antizyme 1 and Antizyme 2 (Fig. 2). That is,

five base pairs in Stem 1 of both Antizyme 1 and Antizyme 2 are the same and the Stem 2 in both Antizyme 1 and Antizyme 2 is the same, too. Moreover, both Matsufuji *et al.*[23] and Ivanov *et al.*[24] considered their pseudoknots stimulate +1 ribosomal frameshifting. These facts make us believe that the same function of these two pseudoknots is from those conserved base pairs.

Also, the pseudoknot from Shigemoto *et al.*[14] is different with others' models discussed here (Fig. 3 and Table 2), with five base pairs and two unpaired nucleotides in Stem 1, four base pairs in Stem 2, one nucleotide in Loop 1, eight nucleotides in Loop 2 and two unpaired nucleotides in the junction between Stem 1 and Stem 2. Shigemoto *et al.*[14] also reported a different function: her pseudoknot stimulated –1 frameshifting.

The pseudoknot from Ben-Asouli *et al.*[8] is quite different with others' models, with seven base pairs in Stem 1, four base pairs in Stem 2, 41 nucleotide in Loop 1, seven nucleotides in Loop 2 and four unpaired nucleotides in the junction between Stem 1 and Stem 2. Moreover, Ben-Asouli *et al.*[8] also reported a quite different function, his pseudoknot acts to adjust translation of the interferon-γ mRNA by activating the interferon-inducible protein kinase, PKR.

In summary, these reports and discoveries tell us that the structure of a pseudoknot decides its function. More specifically, the conserved region of a pseudoknot decides its function. Besides, we also see the pseudoknots reported in mammals are slightly different with virus pseudoknots in structure. Meanwhile, this study also brings us two questions: What is the most minimal structure for a pseudoknot to maintain its efficient function in mammals? And what is the most optimal structure for a pseudoknot to maintain its function in mammals? As a future research direction, we will try to find out some solutions to these questions especially for the pseudoknots from Human C2H2-ZNF genes.

ACKNOWLEDGMENTS

We are very grateful to Limsoon Wong, See Kiong Ng, Chua Nam Hai, Charlie Hodgman and Ellen Mackay for their helpful comments on this work and to Jinyan Li for carefully reading a draft of the manuscript.

REFERENCES

1. http://virology.wisc.edu/acp/pages/Classes/ClassMaterials/TableofWords_711_02.html
2. ten Dam, E., C.W. Pleij, and D. Draper, 1992. Structural and functional aspects of RNA pseudoknots. Biochemistry, 31: 11665-11676.
3. Gilley, D. and E.H. Blackburn, 1991. The telomerase RNA pseudoknot is critical for the stable assembly of the catalytically active ribonucleoprotein. Proc. Natl. Acad. Sci. USA., 96: 6621-6625.
4. Powers, T. and H.F. Noller, 1991. A functional pseudoknot in 16S ribosomal-RNA. EMBO J., 10: 2203–2214.
5. Feng, Y.X., H. Yuan, A. Rein and J.G. Levin, 1992. Bipartite signal for read-through suppression in murine leukemia virus mRNA: An eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot. J. Virol., 66: 5127-32.
6. Nameki, N., B. Felden, J.F. Atkins, R.F. Gesteland, H. Himeno and A. Muto, 1999. Functional and structural analysis of a pseudoknot upstream of the tag- encoded sequence in *E. coli* tmRNA. J. Mol. Biol., 286: 733-744.
7. Ferré-D´Amaré, A.R., K. Zhou and J. A. Doudna, 1998. Crystal structure of a hepatitis delta viurs ribozyme. Nature, 395: 567-574.
8. Ben-Asouli, Y., Y. Banai, Y. Pel-Or, A. Shir and R. Kaempfer, 2002. Human interferon-γ mRNA autoregulates it's translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. Cell, 108: 221-232.
9. Farabaugh, P.J., 1996. Programmed translational frameshifting. Annu. Rev. Genet., 30: 507-528.
10. Gesteland, R.F. and J. F. Atkins, 1996. Recoding: Dynamic reprogramming of translation. Annu. Rev. Biochem., 65: 741-768.
11. Harger, J.W., A. Meskauskas and J.D. Dinman, 2002. An 'integrated model' of programmed ribosomal frameshifting. Trends in Biochemical Sci., 27: 448-454.
12. Baranov, P.V., O.L. Gurvich, A.W. Hammer, R.F. Gesteland and J.F. Atkins, 2003. Recode 2003, Nucleic Acids Res., 31: 87-89.
13. Atkins, J.F. and R.F. Gesteland, 1999. Intricacies of ribosomal frameshifting. Nature Structural Biology, 6: 206-207.
14. Shigemoto, K., J. Brennan, E. Walls, C.J. Watson, D. Stott, P.W.J. Rigby and A.D. Reith, 2001. Identification and characterization of a developmentally regulated mammalian gene that utilizes-1 programmed ribosomal frameshifting. Nucleic Acids Res., 29: 4079-4088.
15. Alam, S.L., J.F. Atkins and R.F. Gesteland, 1999. Programmed ribosomal frameshifting: Much ado about knotting. Proc. Natl. Acad. Sci. USA., 96: 14177-14179.
16. Dinman, J.D., M.J. Ruiz-Echevarria and S.W. Peltz, 1998. Translating old drugs into new treatments: Ribosomal frameshifting as a target for antiviral agents. Trends Biotechnol., 16: 190-196.

17. Atkins, J.F., P.V. Baranov, O. Fayet, A.J. Herr, M.T. Howard, I.P. Ivanov, S. Matsufuji, W.A. Miller, M.B. Moore, M.F. Prère, N.M. Wills, J. Zhou and R.F. Gesteland, 2001. Overriding standard decoding: Implications of recoding for ribosome function and enrichment of gene expression. Cold Spring Harbor Symposia on Quantitative Biol., LXVI: 217-232.
18. Shah, A.A., M.C. Giddings, J.B. Parvaz, R.F. Gesteland, J.F. Atkins and I.P. Ivanov, Computational identification of putative programmed translational frameshift sites. Bioinformatics, 18: 1046-1053.
19. Kang, H., J.V. Hines and I.T. Jr., 1996. Conformation of a non-frameshifting RNA pseudoknot from mouse mammary tumor virus. J. Mol. Biol., 259: 135-147.
20. Wills, P.R., 1992. Potential pseudoknots in the PrP-encoding mRNA. J. Theor. Biol., 159: 523-527.
21. Barrette, I., G. Poisson, P. Gendron and F. Major, 2001. Pseudoknots in prion protein mRNAs confirmed by comparative sequence analysis and pattern searching. Nucleic Acids Res., 29: 753-758.
22. Safar, I., H. Wille, V. Itri, D. Groth, H. Serban, M. Torchia, F.E. Cohen and S.B. Prusiner, 1998. Eight prion strains have PrP(Sc) molecules with different conformations. Nature Med., 4: 1157-1165.
23. Matsufuji, S., T. Matsufuji, Y. Miyazaki, Y. Murakami, J.F. Atkins, R.F. Gesteland and S. Hayashi, 1995. Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. Cell, 80: 51-60.
24. Ivanov, I.P., R.F. Gesteland and J.F. Atkins, 1998. A second mammalian antizyme: Conservation of programmed ribosomal frameshifting. Genomics, 52: 119-129.
25. Venter, I.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans and R.A. Holt *et al*., 2001. The sequence of the human genome. Science, 291: 1304-1351.
26. Dehal, P., P. Predki, A.S. Olsen, A. Kobayashi, P. Folta, S. Lucas, M. Land, A. Terry, C.L.E. Zhou, S. Rash, Q. Zhang, L. Gordon, J. Kim, C. Elkin, M.J. Pollard, P. Richardson, D. Rokhsar, E. Uberbacher, T. Hawkins, E. Branscomb and L. Stubbs, 2001. Human chromosome 19 and related regions in mouse: Conservative and lineage-specific evolution. Science, 293: 104-111.
27. Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall and E.L.L. Sonnhammer, 2002. The Pfam contribution to the annual NAR database issue. Nucleic Acids Res., 30: 276-280.
28. Shannon, M. and L. Stubbs, 1998. Analysis of homologous XRCC1-linked zinc-finger gene families in human and mouse: evidence for orthologous genes. Genomics, 49: 112-121.
29. Bajic, V.B. and S.H. Seah, 2003. Dragon gene start finder: An advanced system for finding approximate locations of the start of gene transcriptional units. Genome Res., 13: 1923-1929.
30. Nishikawa, T., T. Ota and T. Isogai, 2000. Prediction whether a human cDNA sequence contains initiation codon by combining statistical information and similarity with protein sequences. Bioinformatics, 16: 960-967.
31. Weaver, R.F., 1998. Molecular Biology, WCB/McGraw-Hill, Boston.
32. Brodsky, L.I., V.V. Ivanov, dzidis Ya. L. Kalai, A.M. Leontovich, V.K. Nikolaev, S.I. Feranchuk and V.A. Drachev, 1995. GeneBee-NET: Internet-based server for analyzing biopolymers structure. Biochemistry, 60: 923-928.
33. Portin, P., 2002. Historical development of the concept of the gene. J. Medicine and Philosophy, 27: 257-286.
34. Du, Z. and D.W. Hoffman, 1997. An NMR and mutational study of the pseudoknot within the gene 32 mRNA of bacteriophage T2: Insights into a family of structurally related RNA pseudoknots. Nucleic Acids Res., 25: 1130-1135.
35. Chen, X., M. Chamorro, S.I. Lee, L.X. Shen, J.V. Hines, I.Jr. Tinoco and H.E. Varmus, 1995. Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: Nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. EMBO J., 14: 842-852.