Non-Watson Crick base pairs might stabilize RNA structural motifs in ribozymes – A comparative study of group-I intron structures

K CHANDRASEKHAR and R MALATHI*

Department of Genetics, Dr ALM PG Institute of Basic Medical Sciences, University of Madras, Taramani Campus, Chennai 600 113, India

**Corresponding author (Fax, 91-44-24926709; Email, r_malathi@hotmail.com)*

In recent decades studies on RNA structure and function have gained significance due to discoveries on diversified functions of RNA. A common element for RNA secondary structure formed by series of non-Watson/Watson Crick base pairs, internal loops and pseudoknots have been the highlighting feature of recent structural determination of RNAs. The recent crystal structure of group-I introns has demonstrated that these might constitute RNA structural motifs in ribozymes, playing a crucial role in their enzymatic activity. To understand the functional significance of these non-canonical base pairs in catalytic RNA, we analysed the sequences of group-I introns from nuclear genes. The results suggest that they might form the building blocks of folded RNA motifs which are crucial to the catalytic activity of the ribozyme. The conservation of these, as observed from divergent organisms, argues for the presence of non-canonical base pairs as an important requisite for the structure and enzymatic property of ribozymes by enabling them to carry out functions such as replication, polymerase activity etc. in primordial conditions in the absence of proteins.

[Chandrasekhar K and Malathi R 2003 Non-Watson Crick base pairs might stabilize RNA structural motifs in ribozymes – A comparative study of group-I intron structures; *J. Biosci*. **28** 547–555]

1. Introduction

Base pair mismatches have been often found in RNAs ribosomal RNA (Gautheret *et al* 1995), ribozyme, viroids (Pieczenik 1994) etc. These base pair mismatches are recognition sites for proteins, metal ions and small molecules (Jeffrey *et al* 1999; Blanchard *et al* 1998). Indeed, structural model of natural or synthetic RNA's revealed the existence of a number of different non-Watson Crick arrangements occurring as a single, tandem or consecutive base pairs amongst which GU, GA, AA, UU are often encountered (Pley *et al* 1994; Baeyens and BhHolbrook 1995; Battisle *et al* 1996; Pan *et al* 1999). These non-standard base pairs could be ascribed a role: (i) GU base pairs in RNA provide a complex array of hydrogen bond donors and acceptors creating a surface area for binding of proteins and metal ions. (ii) Provide metal-binding site in a GA of a RNA dodecamer (Baeyens *et al* 1996), also in GU, GA of Tetrahymena (Golden *et al* 1998) and in hammerhead ribozyme (Pan *et al* 1999). (iii) GA/AG tandem is functionally important in mediating seleno protein translation (Walczak *et al* 1998). (iv) GNRA loop in large subunit of ribosomal RNA of sacrin/ricin act as a specific site for binding of Pokeweed antiviral protein (PAP) (Francis *et al* 2001). Studies on a number of RNA oligomers and ribozymes have shown that these mis-pairs can form folded RNA structural motifs or a part of an internal loop, leading to distortion in the RNA double helical structure or distinctive structural features that might enable intermolecular recognition. It has to be mentioned here that a GU base pair within the acceptor stem of tRNA plays a vital role in aminoacylation process (Ramos and Varanu 1997). Group-I intron ribozymes are structurally related class of functional RNA elements that are capable of catalyzing their

Keywords. Group-I intron; non-canonical base pair; ribozyme; RNA structure; *Tetrahymena thermophila*; wobble base pair

Abbreviations used: A, Adenine; C, cytosine; G, guanosine; IGS, internal guide sequences; J, linkers; L, loop; P, stem; U, uracil.

own excision and exon ligation in a well-characterized two-step process. Catalytic RNAs of this type encompass a wide variety of specific RNA sequences that share a common secondary and tertiary structure for the self-splicing process. The group-I intron has a well-organized structure with P1–P10 stems and loops having highly conserved base pairs (both Watson and non-Watson type) involved in forming domains, loops and linkers, during the active core formation (Michel and Westhof 1990). Group-I intron ribozyme is another important RNA (Walczak *et al* 1998) wherein occurrence of non-Watson Crick base pairs facilitate forming a multi-layer core complex leading to autocatalytic activity of the ribozyme. The group-I intron sequences (nuclear genes) from protozoa to fungal, to algae have been analysed for the precise location of different non-Watson Crick base pairs, to assess the functional and evolutionary significance of these interactions. Clearly, non-canonical base pairs in group-I intron occur at specific region of helices or stems. These non-canonical base pairs in group-I intron are functionally important in: (i) recognition of guanosine cofactor, (ii) binding to metal ions, thereby reducing the electrostatic repulsive interactions due to polynucleotides folding, and (iii) facilitating domain–domain interaction involving tertiary hydrogen bonds. Furthermore, the wobble pair GU is the most conserved in the group-I intron contributing to the folded structure and base pairs of the type GA and AA provides additional stabilization to the active core structure. A high degree of conservation of these in all the sequences analysed hints at the possible role of these base pairs in structure, function and evolution of RNA.

2. Materials and methods

Sequences of group-I intron corresponding to ribosomal RNA of various organisms representing myxomycota, rhizopoda, ciliophora, rhodophyta, gamophyta, chlorophyta, ascomycota and basidiomycota were obtained from Gene Bank (http://pundit.icmb.utexas.edu/RNA/GRPI/introns.html). The sequences were aligned and fitted into the secondary structure for comparisons. The occurrence of non-canonical base pairs in the P1 to P10 stem-loop structures were analysed.

3. Results and discussion

Recent reports on crystal structure of group-I intron ribozyme from *Tetrahymena* (Golden *et al* 1998) have proposed that non-Watson Crick base pairs contribute to a major extent in stabilizing and maintaining the tertiary folding leading to the self splicing process. With this in view a comparative sequences analysis was carried out to examine the conservation of these base pairs in group-I intron of nearly 48 nuclear genes from divergent species (especially from cynobacteria to protozoans species) to study their importance in relation to their structure–function, evolution and the results are presented in figure 1 and table 1.

Wobble pairs of GU type (figure 1 and table 1) assume a special significance among the non-canonical base pairs. Particularly, a GU base pair in P1 duplex plays a key role in the 5′ splice site cleavage-ligation process. It provides the necessary orientation for the formation of IGS during the docking of P1 with P10 helix. Additionally, this base pair also offers a distinct electrostatic potential (Walczak *et al* 1998) for facilitating close interaction of P1 with that of P7 helix which binds to a exogenous guanosine nucleotide, which, in turn, is required for initiating the splicing reaction. It has been suggested that (Alexander *et al* 1998) it is the uridine of GU base pairs which is perhaps involved in binding to divalent metal ions with high affinity due to localization of their specific structure that is necessary for the exogenous guanosine to initiate the 5′ splice site cleavage. Their conservation suggests their importance in catalytic activity of group-I intron ribozyme. It is highly interesting to see the extensive conservation of GU base pairs at that position in the entire organism suggesting that it should have a key role in the autocatalytic activity. During the core organization, the L2 of P2 stem are found to interact with L5c of P5c thereby creating a tertiary contact at the junction of catalytic core. For undergoing this type of loop–loop interaction, the P2 stem is provided with an unusual base pair of type UG which provides flexibility for the domain without creating any helix distortion in addition to the Watson base pairs. In few species instead of UG, GU is found. Whether it might be UG/GU, it prefers an unusual base pair of this type, which again gains more importance in the group-I intron structure as in P1 helix. The presence of this unusual base pair provides flexibility for the P2 stem to interact with near by domain/loop that protects them from hydroxyl radical cleavage and hence this base pair is extensively conserved.

Group-I intron ribozyme has a major P4–P6 domain and P3–P9 sub-domain, both of which are constituted by coaxially stacked helices. P4–P6 is comprised of a number of non-canonical base pairs with an extension of P5abc by a hinge. In P5 stem unusual base pair of type GU has been conserved only protozoans species while in other it was absent and instead occurrence of Watson type has been analysed. Recent reports (Walczak *et al* 1998; Colmenarejo and Tinoco 1999) indicated that GU base pair in P5 stem maintains the helix stability of the major domain during docking of L9 (of P9 stem) over P5 at the time of folding. However, absence of this base pair in other species indicates that though they are not evolutionarily important for the catalytic activity but might be

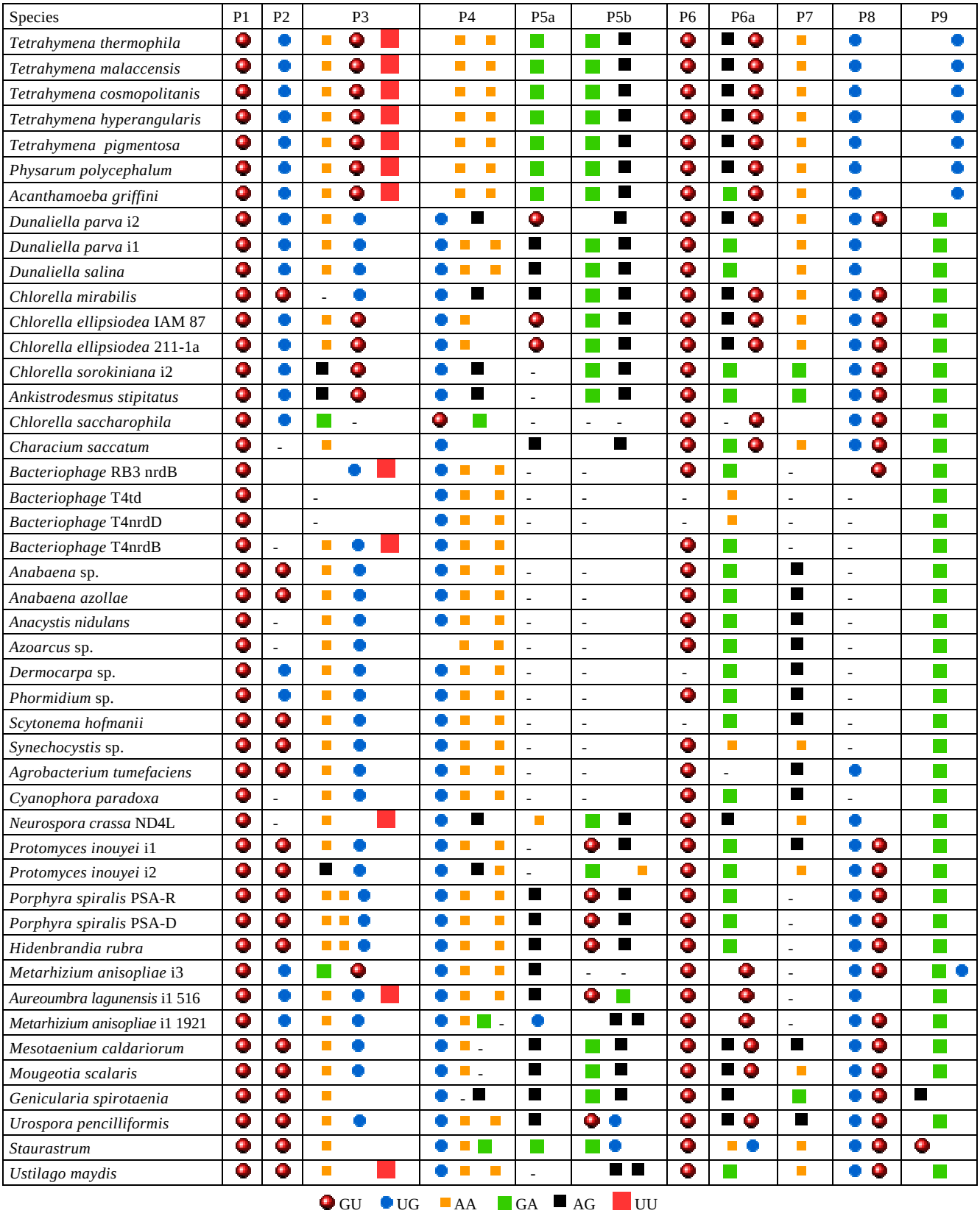

Figure 1. Conservation of non-Watson Crick base pair in the various domains of group-I intron structure. The different types of non-Watson Crick base pair are highlighted in different colours.

in species wise. A high score of GU conservation was analysed in P6 stem except in phage bacteria and in one *tRNA* genes it was absent. Their conservation throughout the table suggests that these are functionally crucial in maintaining the hairpin-shape of P4–P6 major domain and also being an evidence for their ancestral gene in possessing this base pair at the primordial stage. Reports also (Michel *et al* 1990) indicated that the U of GU is involved in forming base triples with U residue of J6/J7 linker that facilitates close brought of the sub-domain to the major domain during three-dimensional folding period. Additionally, the P4 stem of the major domain

Table 1. Species that contain group-I intron gene in nuclear genes of precursor ribosomal RNA and their accession number.

S.No.	Class	Species	Accession number
1.	Protozoa	*Tetrahymena thermophila*	J01235
2.		*Tetrahymena cosmopolitanis*	X03107
3.		*Tetrahymena hyperangularis*	X03106
4.		*Tetrahymena pigmentosa*	J01210
5.		*Tetrahymena malaccensis*	X03105
6.		*Physarum polycephalum*	L03183
7.		*Acanthamoeba griffini*	U02540
8.	Fungi	*Pneumocystis carinii*	X13687
9.		*Pneumocystis carinii* Pc1	M86760
10.		*Pneumocystis carinii* Pc2	L13614
11.		*Pneumocystis carinii* Pc3	L13615
12.	Gamophyta	*Staurastrum* sp.	X77452
13.		*Urospora penicilliformis*	#NA
14.	Basidiomycota	*Ustilago maydis*	X62396
15.	Chlorophyta	*Characium saccatum*	M84319
16.		*Chlorella saccharophila*	X73991
17.		*Dunaliella parva* i2	M62998
18.		*Dunaliella parva* i1	M62998
19.		*Dunaliella salina*	M84320
20.		*Chlorella mirablilis*	X74000
21.		*Chlorella ellipsoidea* 118	D13324
22.		*Chlorella ellipsoidea* 211-1a	X63520
23.		*Ankistrodesmus stipitatus*	X56100
24.		*Chlorella sorokiniana* i2	X73993
25.	Rhodophyta	*Porphyra spiralis* PSA-R	L26177
26.		*Porphyra spiralis* PSA-D	L26176
27.		*Hildenbrandia rubra*	L19345
28.	Bacteria	Bacteriophage T4td	M12742
29.		Bacteria RB3nrdB	#Gene Bank
30.		Bacteriophage RB3nrdB	X59078
31.		Bacteriophage T4nrdD	Y00122
32.	Ascomycota	*Protomyces inouyei* i1	D11377
33.		*Protomyces inouyei* i2	D11377
34.		*Neurospora crassa* ND4L	X05115
35.	tRNA containing species	Phormidium tRNA	M61163
36.		Dermocarpa sp. tRNA	U10480
37.		Synechocystis tRNA	U10482
38.		*Scytonema hofmanii* tRNA	U10481
39.		*Nicotiana tabacum* tRNA	M16898
40.		*Cyanophora paradox* tRNA	M22563
41.		*Anabaena* sp.	M38692
42.		*Anabena azollae*	M38691
43.		*Agrobacterium tumefaciens*	X66220
44.		*Azoarcus* sp. tRNA	X6622′
45.		*Anacystis nidulans* tRNA	M61158
46.		*Metarhizium anisopliae* i1	AF 197120
47.		*Metarhizium anisopliae* i3	AF 197124
48.		*Aureoumbra lagunensis*	U40258

gains more importance in the group-I intron ribozyme since they are known to be involved in many minor–major groove contact at the time of folding as reported recently (Michel and Westhof 1990). Moreover being in close proximity to the P1–P10 splicing site junction, distortion of helix might occur due to the backbones which is stabilized by few non-Watson base pairs, by providing a different array of geometry to the stem that reflects in the structure and also act as a site for metal ions to interact that indirectly aids in maintaining the three-dimensional folding in order to undergo the autocatalytic process. While analysing the conservation of non-Watson type base pairs in P4 stem, we find UG is highly conserved in all species except for in protozoans. Along with this on the other hand self-base pairing type of sheared AA was analysed. The conservation of AA is remarkable except for one or two species it was where AG/GA was found (figure 1). Conservation of UG/AA is suggestive for their presence in the primordial stage. However, in due course of evolution/geographical factors, other type of mis-pair AG/GA would have occurred but from the very fact it is understood that this domain requires a non-Watson pair to maintain the coaxial arrangement of P4–P6 domain during the close brought of the sub-domain and interactions.

In the intron, the P3–P7–P9 domain wraps around the P4–P6 major helix at the time of folding together forming the active ribozyme core. The P3–P9 domain is comprised of four helices P8, P3, P7 and P9·0 coaxially stacked over the other, each comprised non-Watson Crick base pairs. In P3 stem self-pairs of type AA occur. From the point of conservation their even distribution in the gene sequence except one or two species where AG/GA was analysed. The other type of mis-pairs like UU (Baeyens and BhHolbrook 1995) type (indicated as rectangular orange colour) (figure 1) in protocistista, myomycota, RB3nrdB bacteria and *Neurospora crassa* ND4L were observed whose conservation is less in other species as analysed from the table 1. The self base pairings could be evolved in recent species and their absence in other species clearly indicates that they are not evolutionary important for the structure since the progenitors miss this base pair. Analysis of conservation pattern revealed new type of base pairs, i.e. UC, AC, and especially UG in many sequences together with AA. In their absence, GU was observed in the stem. The presence of UG or GU along with iso-steric base pair AA in the early progenitors and even in the recently evolved species hints at their functional significance during the folding facilitating the RNA to act as an enzyme in the absence of protein. Functionally the P3 perhaps requires unusual pairs of any type for providing flexibility to the stem while folding, as well as for maintaining the co-axial arrangement of their sub-domain helix during the close brought of P9–P3–P8 over the major helix to allow for certain minor/major groove interaction with the base pair/bases at that junction required for self splicing process.

The P9 stem forming a part of sub-domain is important as P1 duplex. The residues in L9 of this stem undergo a minor groove–tetraloop interaction with P5 whose orientation thereby facilitates the alignment of 3′ splice site close to the P1 duplex during the multi core complex formation. This is facilitated by an unusual base pair GA (green colour) which is conserved in all the species except in protozoans and in one or two chlorophyta spp. where the base pair is UG/GU. G appears to be important based on the experimental reports (Golden *et al* 1998) metal ion near the G nucleotides of the adjacent G¥U base pairs, where hydrogen bonds with the N7 and carbonyl groups of both guanines while the carbonyl groups of uracil residues add to the negative surface of the binding pocket, but do not form hydrogen bonds with the hexammine. In fact by this it is clearly understood that interaction of P5–L9 tetraloop is stabilized by a metal ion and at the same time the repulsive force excreted at this junction where reunion of P1–P10 and other sub-domain/major domain remains in contact is also stabilizes leading to the higher-order structure formation for the autocatalytic activity. Evolutionarily their conservation in the analysed gene sequences of group-I intron (figure 1, shown in blue colour) can be considered as an "Universally conserved G" since functionally they play a major role in the group-I intron structure.

A stem with many non-Watson base pairs is P8, a part of the sub-domain whose function has been less highlighted in the reports. From the sequence analysis (table 1, figure 1) we find an extreme conservation of UG and simultaneously alone with GU type was observed. In few absence of both types were analysed, particularly in *tRNA* genes. Based upon the conservation and recent crystal structure of group-I intron it can be proposed that the unusual base pairs in P8 stem provide flexibility to the stem during a shift of P8 from one side of the helix to the other side by maintaining the orientation of P8–P3–P7–P9 helix. However, this orientation facilitates the adenyl residue at 306′ of J8/7 linker to form one hydrogen bond interaction with exogenous guanosine which is important for the 5′ splicing site initiation. Apart from this, other adenyl residues of this linker might involve in coordination with metal ions thereby indirectly stabilizing the active core structure from electrostatic force. Besides these, AG bp in *Tetrahymena* sp. alone. This indicates that the species lacking AG type can be considered as recently evolved species and absence of the UG/GU type of base pairs in P8 stem of bacteria and in *tRNA* genes supports tRNA containing group-I intron

genes evolved earlier but after the LSU/SSU ribosomal RNA of the nuclear genes.

The P7 helix whose structure is highly suitable for guanosine binding is stabilized by a number of interactions. It consists of a unique 'A' bulge which is stacked over GC base pair and a non-canonical base pair of the type AA seen in the beginning of the stem, in addition to Watson Crick base pairs. The AA base pair is mainly involved in maintaining the orientation of P7 helix as undisturbed during a shift of the P8 stem from one side to other side of the helix (P5abc), which indirectly facilitates the interaction with exogenous G with P7 helix and positions of 3′–OH exactly at the junction of 5′ splice site. Additionally, it perhaps maintains the continuity of the strand during three-dimensional folding. Other types of mis-pairs like AG/GA, GU, AC and GG types were seen in P7 helix during the analysis of group-I intron gene sequences. The self-base paring interaction involves A306 of J8/J7 and A269 of P7 helix and these sheared base pairs are conserved only in very few while in others it varies. It could be seen that this sheared AA is conserved in protozoans and only in few species of chlorophyta. While in tRNA it was AG and very rarely GA was observed in 3 species of chlorophyta as from the table 1. Adenyl residue at 269 of P7 helix is conserved in all the gene sequences, except in few species 'G' residue was observed that results in AG, while in other few species the residue at 306 of J8/J7 linker varies that in fact results in GA and other types of mis-pairs. Recent reports (Baeyens and De Bondt 1996) have indicated that the adenyl residue at 306 is implicated in forming a gateway for the exogenous guanosine binding that involves certain tertiary minor groove contact between other domains, which indirectly facilitates the initiation of splicing in *Tetrahymena*. However, our analysis as shown in figure 1 and table 1 it is not the adenyl alone at 306 position which is important for the G-binding but it should be purine at that position for the gateway formation. Thus it can be inferred that the AA base pairs containing species perhaps originated recently while the progenitors perhaps possessed this GA/AG types of unusual base pairs.

A leaf like extension of the group-I intron structure is the P5abc domain, which is separated from the major P4–P6 domain by a hinge and possesses unusual base pairs of type AG/GA in P5a, GU tandem in P5b and GA at the end of stem and beginning of the L5b loop. According to the recent reports (Alexander *et al* 1998), the AG/GA type in P5a provides flexibility to the helix during loop–loop or loop–domain contact at the time of initiation of folding, while the GU tandem in P5b off which N7 of G residue involves in co-ordination with the metal ions thereby stabilize the structure from electrostatic repulsive force during close interaction of backbones and nitrogenous bases. But, conservation of this GU is very less – it was analysed only in protozoans – which indicate that the progenitor of group-I introns lack this and in recent evolved group-I introns this would have been acquired during the course of evolution. Another unusual type AG was conserved in all species except in tRNA and bacteria species. Based on earlier reports, this unusual type is implicated in interaction with residues of J6/J6a linker, namely "Tetraloop" contact. This is in fact one of the iron core of the structure, since their interaction forms site for metal ions to interact while the interaction with other domain provides distinct array of geometry to the helix which indirectly reduces the backbone and nitrogenous repulsive in the core for the self splicing process. Phosphorothioate studies (Donnelly *et al* 1998; Murphy and Cech 1994) have shown that divalent ions, like magnesium, associate with adenyl residue of non-standard base pairs and bring about the splicing activity. Analysis of conservation pattern (figure 1 and table 1) shows that GA type in P5a is conserved in protozoans and in few species of chlorophyta, while in other few species of chlorophyta AG was analysed, and very rarely GU was analysed. In P5b stem, a high score of GA at the beginning of the stem and AG at the end and loop beginning of the L5b of P5b domain was analysed from the table and variation are seen only in radhophyta where GU/AG was analysed and in one species of gamophyta, *Urospora penicilliformis* where GU/UG was observed. Though the type of unusual base pairs varies here, nevertheless it involves G which plays a major counterpart role for this domain.

From the point of evolution, the absence of this type of base pairs in bacteria, *tRNA* genes and their occurrence in early chlorophyta, gamophyta, ascomycota species indicate that this might have acquired in early course of evolution and hence conserved in recent evolved species of protozoans, basidiomycota etc, but they lack in progenitor of group-I introns at the primitive stage where in the absence of proteins biological functions might be carried out by this catalytic RNA that are known to act as enzyme as like proteins.

Conclusion

From our analysis we can consider that these conserved mis-pairs occur only in functional domain largely in loops and few in stem. Though the characteristic property of these base pairs are well known but, their importance in the group-I ribozyme varies. However, it has to be highlighted that many of these non-canonical base pairs are found to be distributed evenly in the analysis and from the point of conservation in P1 and P9 they can be considered as an "Universal Conserved base pairs"

(figure 1). They play a major role in the core formation leading to self splicing process. Other base pairs also play a critical role in the three-dimensional structure formation. On comparing sequence analysis, the group-I intron structure filled with Watson base pairs require these mis-pairs for folding, stabilization of tertiary core, helix-interaction and in few cases, some of these base pairs, GU, AA and AG/GA act as recognition site for co-factors and metal ions binding. The GU type of base pair occur only in the stem, while GA/AG type are known to occur at the beginning of the loop or end of the stem and AA type are in-between inter-domains. Clearly, the occurrence of these unusual types is unique at certain regions forming specific structures required to carry out the autocatalytic activity. While analysing these conserved (table 2) features irrespectively, certain new types of base pairs were observed in nuclear and mitochondrial genome (table 2). This type of mis-pairs occur highly in mitochondrial genes while in nuclear genes it is very less, however recent reports indicates that these mispairs are keen in enhancing the activity of the splicing in mRNA by providing strong interaction within the core structure of RNA and so further studies should be carried out to analyse the role of the base pairs from the point function and structure in autocatalytic RNA.

Evolutionarily, the conservation analysis pattern indicates that the group-I intron ribozyme might have evolved from bacterial/fungal progenitors. Thus a hypothetical evolutionary tree (figure 2) has been generated from these conserved non-Watson base pairs of the catalytic RNA intron evolution. In due course of evolution, degeneration/ generation of certain base pairs might have occurred, but the important base pairs in the helix – required for the entire catalytic activity leading to the self-splicing reaction – are conserved, as shown in figure 1 and table 1.

Table 2. The occurrence of other types of non-canonical base pairs that are seen in nuclear and also in mitochondrial gene sequences of group-I intron ribozyme.

Nuclear gene sequences

Species	A–C	C–A	C–C	G–G	C–U	U–C
Metarhizium anisopliae i1	P3, P9	P8, P6	P5b	P5c	–	–
Metahizium anisopliae i3	P5b, P9, P7	P 2.1	–	–	P5, P5b	–
Bacteria T4td	–	–	–	–	–	P3
Bacteria T4nrdD	–	–	–	–	–	P3
Genicularia spirotaenia	P3	–	–	–	–	–
Mougeotia scalaris	P3	–	–	–	–	–
Dunaliella parava i2	–	P5b	–	–	–	P5b
Characium sacctum	–	–	–	–	–	–
Chorella saccharophila	P7	P8	–	–	–	–
Pneumocystis carini Pc1	P3	–	–	–	–	–
Pneumocystis carini Pc3	P3	–	–	–	P5b	–
Aureoumbus lagunensis	P9.3	–	–	–	–	–
Anabanea sp.	P3	–	–	–	–	–
Anabena azollae	P3	–	–	–	–	–
Azoarcus	P3	–	–	–	–	–
Phormidium	P3	–	–	–	–	–

Mitochondria and chloroplast gene sequence

	A–C	C–A	C–C	G–G	C–U	U–C	U–U
Sclerotina sclerotiorum	P3			P7			P6a
Podospora anseria ATP6	P3	–	–	P7	–		P6a
Podospora anseria NDi2	P3	–	–	P7,P9	–		P6a
Podospora anseria ND 3	P3	–	–	P7	–		P6a
Podospora anseria ND4	P3	–	–	–	–		P6a
Willipsis mrakii i1	–	–	–	–	–	P3	P6a
Chlamydomonas eugametos	–	–	–	–	–	P3	–
Chlamydomonas reinhardii	–	–	–	–	–	P3	–

Note: The P stands for stem and the numerical digits for the location of domains.

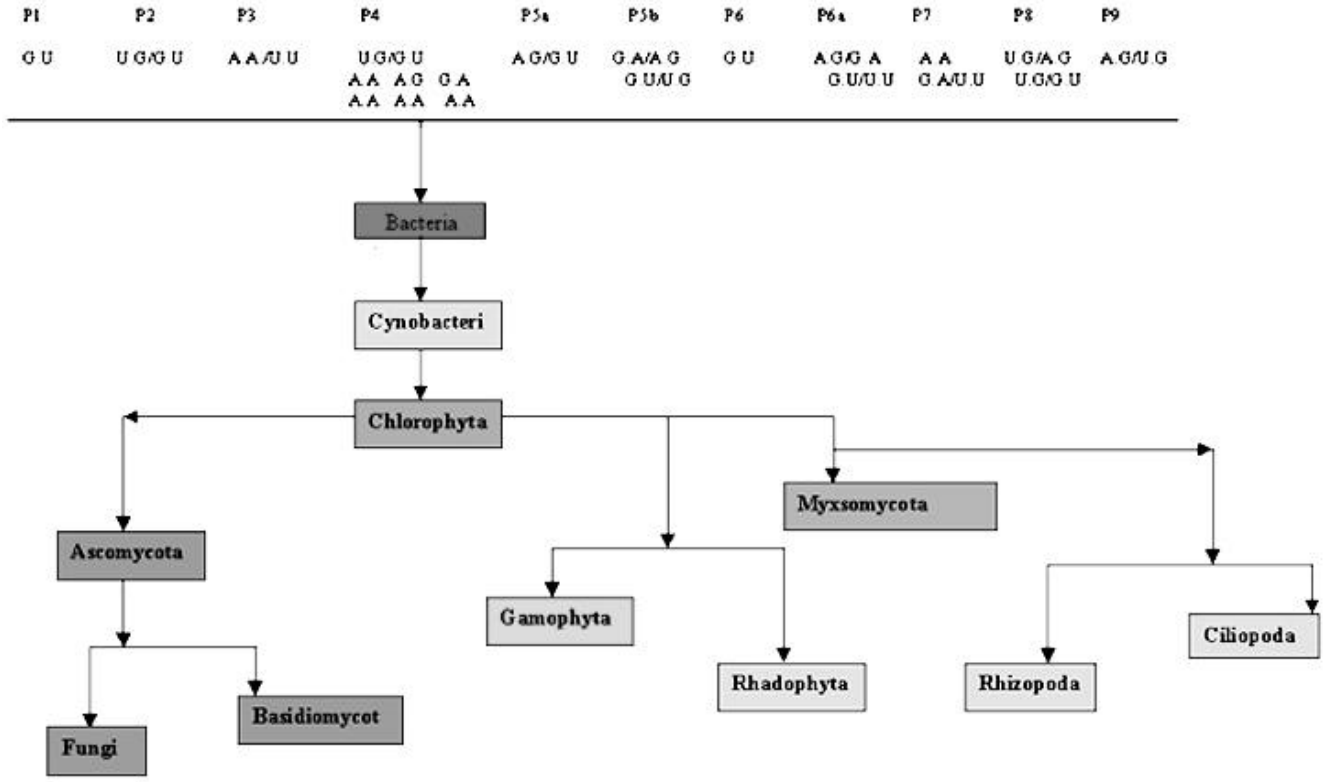

Figure 2. An hypothetical evolutionary tree was shown for the origin of group-I intron ribozymes in different species based on their conserved elements (non-canonical base pairs).

Thus, our findings centre around presence of different types of non-Watson base pairs and their conservation in intron evolution.

Acknowledgements

KC thanks Lady Tata Memorial Trust, Mumbai for the fellowship. We also thank Institute of Mathematical Sciences, Chennai, for providing the facilities to obtain the gene sequences of group-I intron.

References

Alexander A, Szewczak L O D, Sean P R, Eileen M and Strobel S A 1998 *Nature Struct. Biol.* **5** 60–66

Baeyens K J and De BhHolbrook Sr 1995 Structure of an RNA double helix including uracil–uracil base pairs in a loop; *Nature Struct. Biol.* **2** 56–62

Baeyens K J and De Bondt H L 1996 A curved RNA helix incorporating an internal loop with GA and AA; *Proc. Natl. Acad. Sci. USA* **93** 12851–12855

Battisle J L, Mao H, Rao N S, Tan R, Muhandriam D R, Kay L E, Frankel A D and Williamson J R 1996 Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE-RNA complex; *Science* **273** 1547–1551

Blanchard S C, Fourmy D, Eason R G and Puglisi D J 1998 rRNA chemical groups required for aminoglycoside binding; *Biochemistry* **37** 7716–7724

Colmenarejo C and Tinoco I Jr 1999 Structure and thermodynamics of metal binding in the P5 helix of a group-I intron ribozyme; *J. Mol. Biol.* **290** 119–135

Donnelly L O, Szewczak A A, Gutell R R and Strobel S A 1998 The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme; *RNA* **4** 498–519

Francis R, Chen Mao and Uckun M F 2001 Binding interactions between the active center cleft of recombinant pokeweed antiviral protein and the *a*-sarcin/ricin stem loop of ribosomal RNA; *J. Biol. Chem.* **276** 24075–24081

Gautheret D, Konings D and Gutell R R 1995 GU base pairing motifs in ribosomal RNA; *RNA* **8** 807–814

Golden B L, Gooding A R, Podell E R and Cech T R 1998 A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme; *Science* **282** 251–252

Jeffrey B-H, Tok, Junhyeong Cho and Robert R R 1999 Aminoglycoside antibiotics are able to specifically bind the 5′ Untranslated region of Thymidylate synthase messenger RNA; *Biochemistry* **38** 199–206

Michel F and Westhof E 1990 Modelling of the three-dimensional architecture of group-I catalytic introns based on comparative sequences analysis; *J. Mol. Biol.* **216** 585–610

Michel F, Ellington D A, Couture S and Szostak W J 1990 Phylogenetic and genetic evidence for base-triples in the catalytic domain of group-I introns; *Nature* (*London*) **47** 578–580

Murphy L F and Cech T R 1994 GAAA Tetraloop and conserved bulge stabilize Tertiary structure of a group-I intron domain; *J. Mol. Biol*. **236** 49–63

Pan B, Mitra N S and Sundaralingam M 1999 Crystal structure of an RNA 16-mer Duplex R(GCAGAGUUAAAUCUGC)2 with Nonadjacent G(Syn)* A+ (Anti Mispairs); *Biochemistry* **38** 2826–2831

Pieczenik G 1994 The theory of genetypic selection; predicting the direction of evolution as a consequence of GU base pairing and the existence of non-pathogenic strains of HIV-1; *Biochem. Mol. Biol. Int.* **5** 879–887

Pley H W, Flaherty K M and McKay D B 1994 Three dimensional structure of a hammerhead ribozymes; *Nature* (*London*) **372** 68–74

Ramos A and Varanu G 1997 Structure of the acceptor stem of *Escherichia coli* tRNA Ala: role of the G3.U70 base pair in synthetase recognition; *Nucleic Acid Res.* **25** 2083–2090

Walczak R, Carbon P and Krol A 1998 An essential non-Watson-Crick base pair motif in 3′ UTR to mediate selenoprotein translation; *RNA* **4** 74–84

MS received 17 May 2003; accepted 10 July 2003

Corresponding editor: DIPANKAR CHATTERJI