SURVEY AND SUMMARY
The interaction networks of structured RNAs

A. Lescoute and E. Westhof*

Architecture et Réactivité de l'ARN, Université Louis Pasteur, IBMC, CNRS, 15 rue R.Descartes,
F-67084 Strasbourg, France

ABSTRACT

All pairwise interactions occurring between bases which could be detected in three-dimensional structures of crystallized RNA molecules are annotated on new planar diagrams. The diagrams attempt to map the underlying complex networks of base–base interactions and, especially, they aim at conveying key relationships between helical domains: co-axial stacking, bending and all Watson–Crick as well as non-Watson–Crick base pairs. Although such wiring diagrams cannot replace full stereographic images for correct spatial understanding and representation, they reveal structural similarities as well as the conserved patterns and distances between motifs which are present within the interaction networks of folded RNAs of similar or unrelated functions. Finally, the diagrams could help devising methods for meaningfully transforming RNA structures into graphs amenable to network analysis.
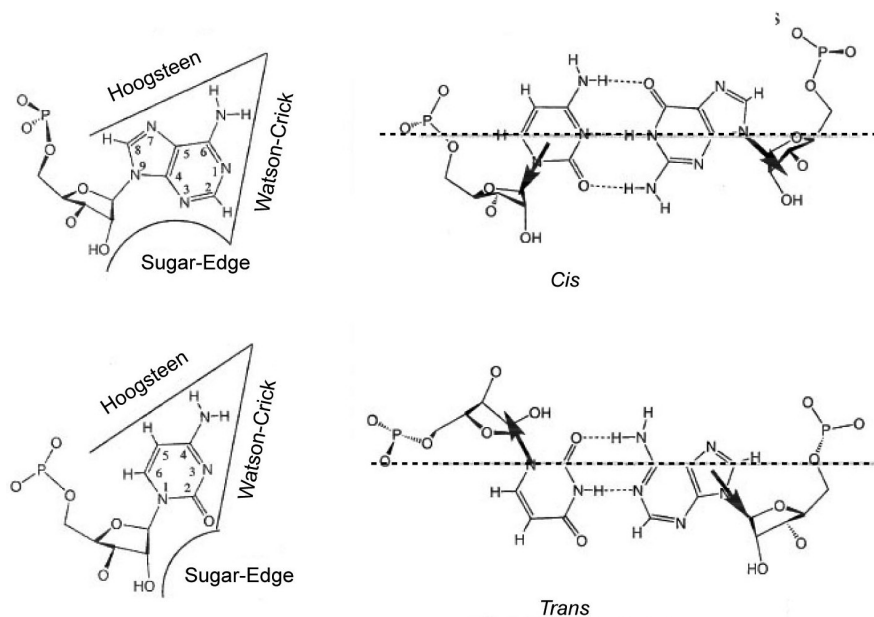
INTRODUCTION

RNA architecture is now visualized as the hierarchical assembly of preorganized double-stranded helices, formed by Watson–Crick base pairs, and which are connected and interlinked by RNA modules maintained by non-Watson–Crick base pairs. The folded and native architecture of a structured RNA molecule is so complex and intimate that each nucleotide forms several types of non-bonded interactions with the neighbouring nucleotides which are brought into contact by the folding process. Among the interactions, one can find: (i) phosphate–phosphate contacts mediated by water molecules or positively charged cations; (ii) phosphate–sugar H-bonding usually mediated by the $2'$-hydroxyl group; (iii) sugar–sugar H-bonding interactions; (iv) base to phosphate or base to sugar H-bonding contacts; (v) base to phosphate or base to sugar stacking interactions; (vi) base to base H-bonding and (vii) base to base stacking interactions. Interactions (ii) to (vii) can occur either directly or indirectly via water molecules or ions (sometimes both modes occur simultaneously). The number of interacting strands generally varies between two and four. Since the polynucleotide sugar–phosphate backbone is monotonous, the specificity of the architectural fold and of the molecular recognition properties resides essentially in the base sequence, although as stated above several interactions either do not involve the bases or are sequence independent. In several instances, such contacts can be considered opportunistic or contingent to the presence of correct neighbouring base–base interactions. These molecular interactions do, however, contribute to the stability of the overall fold. Furthermore, the recent crystal structures clearly reveal the dominant role continuous base stacking plays in the maintenance of RNA architecture. It is, therefore, of central importance to represent and annotate base–base interactions as systematically and concisely as possible. Here, we present maps of the complex networks of interactions between nucleotide bases in known three-dimensional structures of RNAs. One major aim of the present analysis and survey is to facilitate the extraction and the recognition of the sequence constraints underlying a fold or three-dimensional motif. The hope is that the knowledge of such sequence constraints will help predictions of tertiary structure of RNAs on the sole basis of a genomic sequence.

The mapping is based on a nomenclature for base–base contacts that has been previously proposed (1). In such a scheme, any interacting assembly involving base–base interactions can be decomposed as a set of one-to-one base–base contacts. Using the proposed annotation symbols, one can therefore represent on a diagram the ensemble of all base–base contacts in a folded RNA, the canonical Watson–Crick base pairs of the secondary structure and the non-Watson–Crick pairs of the tertiary interactions. Here, we present planar diagrams for several types of structured RNAs. When generating such diagrams, attention is paid to the coaxial stacking or parallel orientation of helices. Thus, such diagrams represent the network of base–base interactions occurring within a RNA architecture. Such diagrams cannot include all the overwhelming richness of any RNA architecture and, thus, cannot compensate for three-dimensional views of the structured RNAs. However, recurring motifs and the networks they form are easily recognizable and local comparisons between structures facilitated (2). Such diagrams contain more information than that conveyed, with

*To whom correspondence should be addressed. Tel/Fax: +33 388 41 70 46; Email : E.Westhof@ibmc.u-strasbg.fr

Figure 1. The three edges of the purine and pyrimidine nucleotides (Watson–Crick, Hoogsteen and Sugar) with the two possible orientations of the attached glycoside bongs. They can be *cis* or *trans* depending on whether they are on the same side (*cis*) or on opposite sides (*trans*) of a line median to the base-base H-bonds. Redrawn after (1).The symbols used for annotating the diagrams are shown below. A nucleotide with a base in the *syn* orientation with respect to the sugar is represented either as a hollowed or as a bold uppercase letter. The annotation for stacking (empty rectangle used as a placeholder) was introduced by (9). In some instances where a single base–base H-bond is present, the connected atoms are given and separated by a dashed line or the symbols are separated by a dashed line, in order to help visualize the facing edges. Within the text, the non-Watson–Crick base pairs is indicated by an 'o' between the interacting nucleotides instead of the symbols and when necessary the full name is given. There is no special symbol for the backbone connectivity; the line or the arrow is chosen so as to prevent confusion with other base-base symbols.

deceptive simplicity, by secondary structure drawings. The improvement in information content is achieved at a cost: (i) the need to know the symbolic nomenclature for the interactions; (ii) the transformation of the conventional secondary structure drawings into new ones displaying co-axial stacking of helices. Finally, whatever the extra information, the diagrams cannot replace three-dimensional views of the folded RNAs. On the contrary, they should be used as a complement to stereo views helping to organize and order the striking and amazing complexity of three-dimensional structures.

MATERIALS AND METHODS

RNA purine and pyrimidine bases present three edges for hydrogen bonding interactions (Figure 1): the Watson-Crick edge, the Hoogsteen edge (for purine) or the C–H edge (for pyrimidine), and the Sugar edge (which includes the sugar 2′-hydroxyl group) (1). A given edge of one base can potentially interact in a plane with any one of the three edges of a second base. The base–base interactions can occur in either the *cis* or *trans* orientation of the glycosyl bonds, i.e. with the attached sugars on the same side or on opposite sides of a line linking the interacting edges. This scheme leads to twelve possible, distinct edge-to-edge base-pairing geometries (or families). Each pairing geometry is designated by stating the interacting edges of the two bases (Watson–Crick represented by a circle, Hoogsteen represented by a square, or Sugar edge represented by a triangle) and the relative orientation of the glycosidic bonds, *cis* (filled symbols) or *trans* (empty symbols) (Figure 1). A historically based priority

Table 1. List of structures with the Protein Data Bank (93) identification code and the crystallographic resolution

Structure		Organism	PDB id	Resolution (Å)	References
Group I intron	with exons	*Azoarcus*	1U6B	3.10	(9)
	ribozyme	*Tetrahymena*	1X8W	3.80	(11)
	ribozyme/product	*Twort*	1Y0Q	3.60	(10)
RNaseP Type A RNA	Spec. Domain	*T.thermophilus*	1US9	2.90	(33)
	Full length	*T.maritima*	2A2E	3.85	(32)
RNaseP Type B RNA	Spec. Domain	*B. subtilis*	1NBS	3.15	(29)
	Full length	*B.stearothermophilus*	2A64	3.30	(31)
16S rRNA		*T.thermophilus*	1J5E	3.05	(35)
		E.coli	2AVY	3.46	(36)
23S rRNA	Transpeptidation site	*H.marismortui*	1S72	2.40	(91)
	TPC + CCA	*H.marismortui*	1Q86	3.00	(92)
	TPC	*E.coli*	2AW4	3.46	(36)
Hepatitis delta virus			1DRZ	2.30	(40)
Diels-Alder ribozyme		*Aptamer*	1YLS	3.00	(37)
Hairpin ribozyme		Viroïd RNA	1HP6	2.40	(44)
Hammerhead ribozyme	All RNA		1MME	3.10	(48)
	Whole	*S.mansoni*	2G0Z	2.20	(49)
Guanine riboswitch	With hypoxanthine	*B.subtilis*	1U8D	1.95	(51)
	With guanine		1Y27	2.40	(50)
	With adenine		1Y26	2.10	(50)
TPP riboswitch		*E.coli*	2GDI	2.05	(54)
SAM riboswitch		*T.tengcongensis*	2GIS	2.90	(52)

rule is invoked for listing the bases in a pair: Watson–Crick edge > Hoogsteen edge > Sugar edge.

This nomenclature allows for computer-based automatic search of base pairings (3–5) with planar drawings of the connected pairs (6). Here all contacts have been visually verified and all drawings done anew in Adobe Illustrator in order to emphasize their biological relevance as described above. Table 1 gives the list of the structures for which we present the diagram of the network of interactions. The selection spans the sizes of the small ribozymes to the large 16S rRNA of the 30S ribosomal subunit. An editable version of all drawings can be retrieved from the laboratory website (http://www-ibmc.u-strasbg.fr/arn/Westhof/).

The evolving view of a structured RNA

The secondary structures of most structured RNAs with a defined and precise function occur in many flavours. For example, group I introns are classified in at least four subgroups characterized by the number and positions of helices attached to an invariant secondary structure core (7). Despite this variability, the overall architectures of the ribozymes promote the stabilization of the helical stems building up the core and the correct positioning of the helical substrates [e.g. in group I introns, see Refs. (8–13)]. This is mainly achieved by the properties of the RNA anchoring motifs which allow for the formation of different and often mutually exclusive long–range contacts between non-homologous peripheral elements. Recurrent and systematic use of essentially two main types of long-range RNA–RNA anchors are observed in domain assembly: GNRA tetraloops with their receptors (14,15) and loop–loop Watson–Crick or non-Watson–Crick base pairings (8,15,16). In order to promote such long-range contacts, sub-domains, which are usually subtended by complex and diverse sets of molecular interactions, have to be assembled. Among them, three-way

junctions constitute frequent and critical structured sub-domains necessary to promote further long-range RNA–RNA contacts. For example, the three-way junction that forms the catalytic core of the hammerhead ribozyme is constrained by tertiary interactions between peripheral elements. These constraints accelerate the folding of the ribozyme which is hundred-times more efficient than the minimally reduced ribozyme (17–19).

Group I introns

Figure 2 illustrates the evolving process of our structural understanding of a large RNA molecule, the prototypical group I introns. First, the secondary structure elements were established using sequence alignments (20) and a planar structure diagram agreed upon (21). After modelling of the three-dimensional structure (7), a new secondary structure diagram was proposed (22). The latter representation (22) included numerous crossings of strands but allowed to visualize the co-axial stacking of helices and the relative position of the co-axial stacks with respect to each other. On the basis of such an architecture of the core, long-range and unsuspected contacts between peripheral elements could be established and modelled (8). Finally, upon completion of crystal structures at sufficient resolution, full interaction diagrams could be produced (9–11).

Supplementary Figure S1 shows the interaction network diagrams for the three crystal structures of group I introns published to date. A comparison between those diagrams shows quickly several points. The numerous A-minor contacts [e.g. J3/4-P6 (23)] and especially those made by the GNRA tetraloops with the 11nt-receptor or two stacked helical base pairs are conspicuous. The 11nt-receptor displays variability in the conformation of the bulging U (compare J8/8a and J5/5a in the *Azoarcus* intron). Differences between the various subgroups appear also. L9 always contact the end
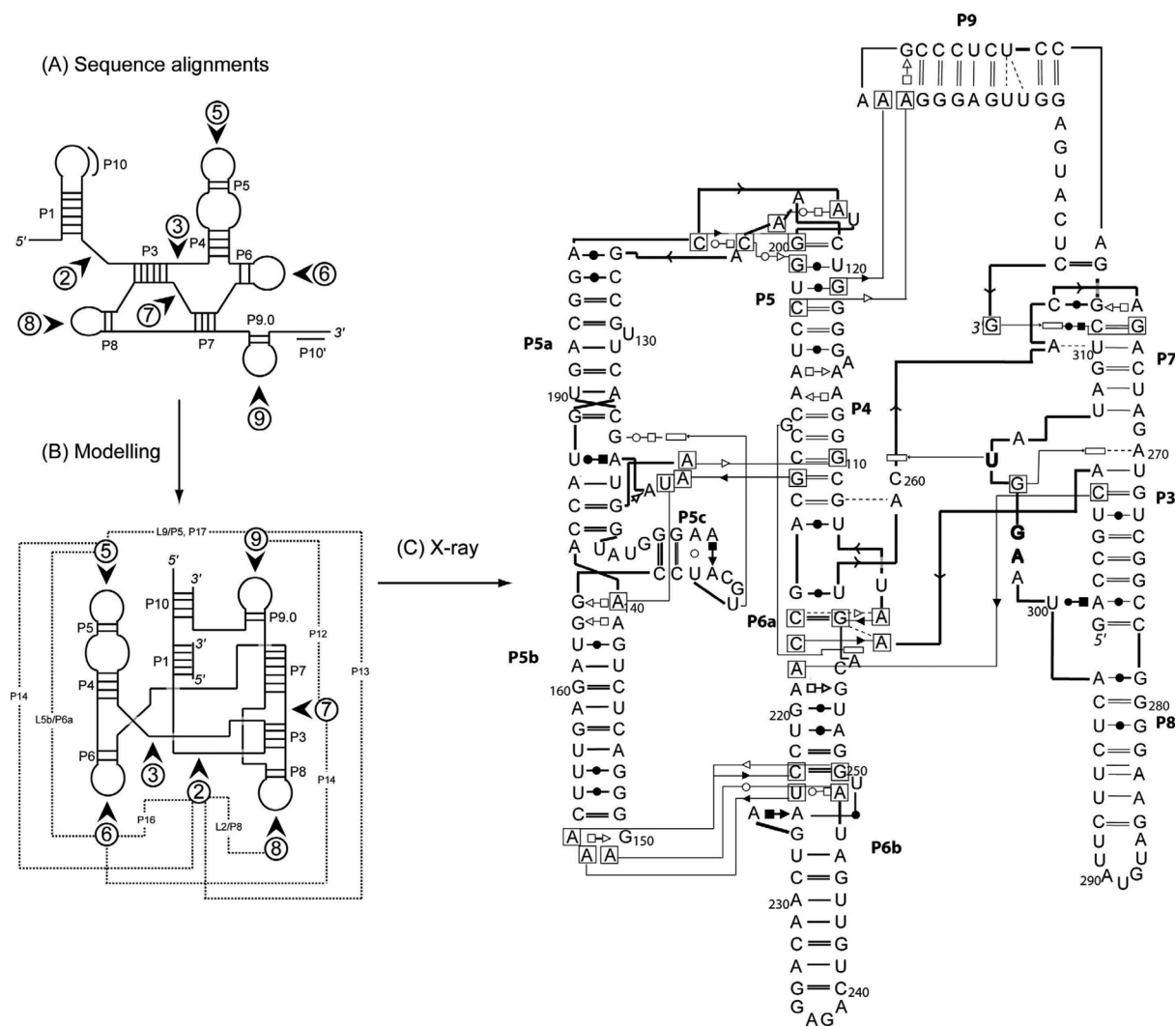
Figure 2. Evolution of the structural representations of group I introns. (**A**) Drawing based on the analysis of sequences (21). (**B**) Drawing with the inclusion of the co-axial stacking based on three-dimensional modelling (7,22). In (A) and (B), the numbers in circles correspond to the domain numbers. (**C**) Drawing with annotated tertiary contacts based on the crystal structure of the *Tetrahymena thermophyla* group I intron (11).

of P5, but P9 folds back on P5 through distinct local topology (a three-way junction in Twort; a 'reverse Kink-turn' in *Azoarcus* (24); an undefined bend in *Tetrahymena*). Depending on the P1/P10 substrate, J8/7 displays different conformations; some nucleotides are in the *syn* conformation in the absence of substrate and new A-minor contacts occur in presence of substrate.

The RNase P structures

A similar evolving process is apparent in the RNase P structures (Figure 3). First, phylogenetic analysis led to the establishment of the secondary structures (25) for the two families (types A and B) (26). Later, modelling studies (27,28) led to the predictions of co-axial stacking between definite helices in coherence with long-range contacts. This formed the basis for a second type of drawings. Recently, crystal structures of the specificity domains (29,30) and of the catalytic domains (31,32) were published.

A striking result of the two structures of the specificity domain was the fold of the large L11/12 internal loop which displayed two T-like loops (33) characterized by a 5-membered loop closed by a *trans* Watson–Crick/Hoogsteen base pair. Despite the variability between the two structures, in part due to the different junctions, some unexpected contacts are maintained, e.g. the last residue of L11/12, C175 in type A, stacks with a purine residue following the first T-like loop while it is G168, part of a bacterial loop E motif in P10.1 of type B (28), which forms the same stack. This is not the only example (see Figure 4): P11 contains a stack of two base triples made of a C=G and A–U Watson–Crick pair to which, respectively, a A forms a Sugar/Hoogsteen *cis* and an amino base a Watson–Crick/Hoogsteen *trans*. But, in type A, the amino base is a A which comes from an internal loop 5′ of P11 and, in type B, it is a C which is a single-stranded residue linking the 3′end of P10 to P10.1. However, the A forming the triple with the C=G base pair is structurally identical in both types. Further, in
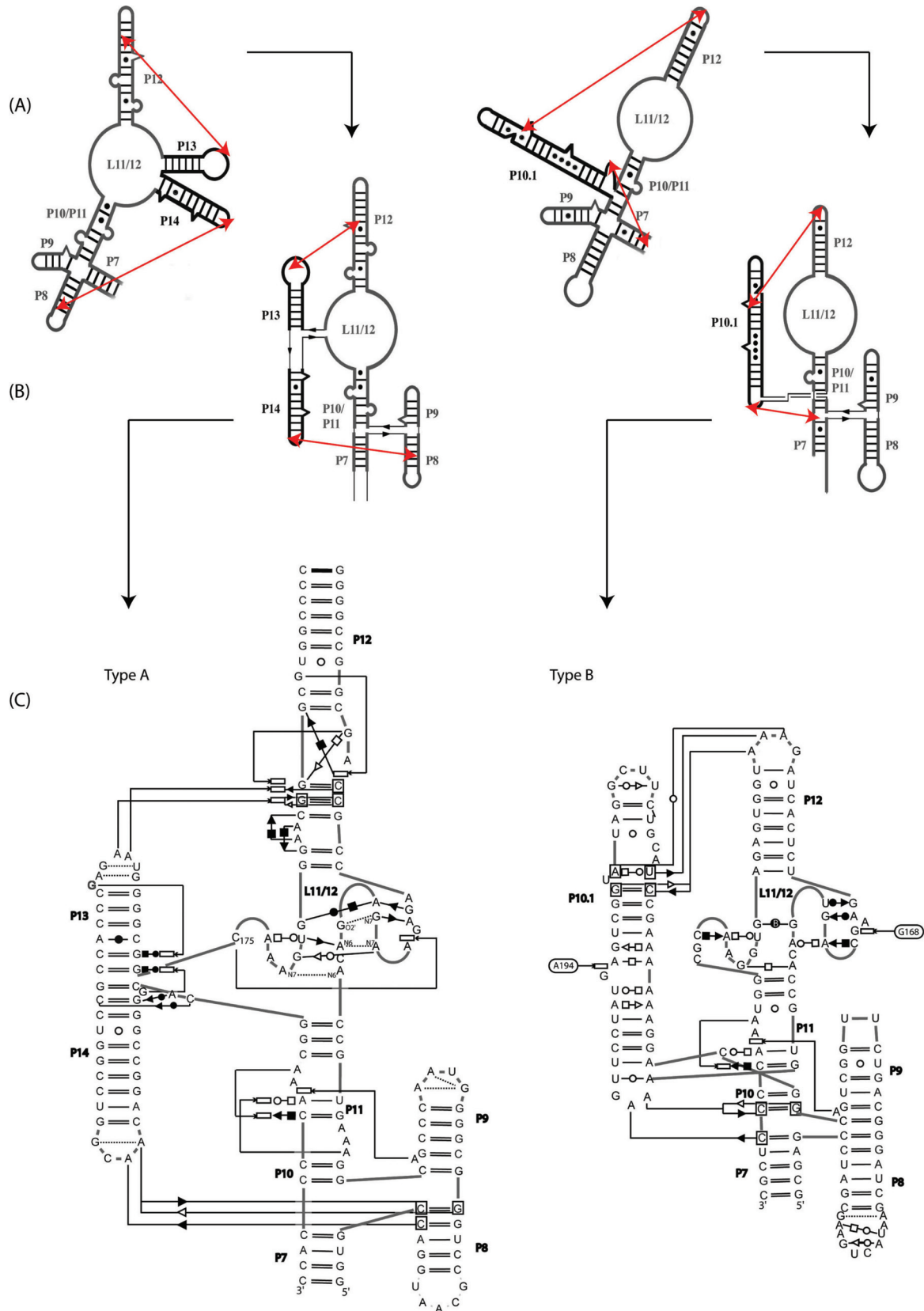
Figure 3. Evolution of the structural representations of the RNA of the RNase P ribozyme. On the left for the type A subgroup and on the right for the type B subgroup. At the top, the representation based on the phylogeny with some non-Watson–Crick base pairs shown as filled circles. (26). In the middle, the representation following the 3D modelling (28). The double ended red arrows correspond to the long-range contacts between secondary structure elements. Below, the annotated tertiary contacts based on the crystal structures (29,30).
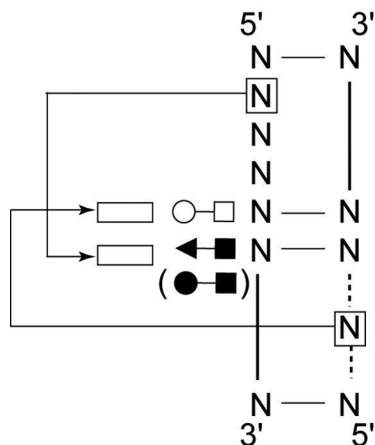
Figure 4. Consensus diagram of a motif observed in P RNA in stem P11 and in the transpeptidylation center of the 23S rRNA.

both types a bulging A from stem P9 stacks 5′ of the A of the A–U pair. These structures stress the role of bulging residues for linking two helical domains (e.g. bulged A in P9 linking P9 and P11 or the last residue of L13 which stacks 5 base pairs below on a bulged G coming from the co-axially stacked P14) (34). An almost identical situation occurs in helix h11 of both 16S rRNA structures (35,36) (see below). For completeness, the catalytic domain (32) and the full ribozyme (31) have been also drawn but a somewhat lower resolution precludes comparisons between these two structures (see Supplementary Figure S2). The multibranch junction around P1–P4 is particularly intriguing and present some analogies with the active site of the 23S rRNA (see below).

The Diels-Adler and nucleolytic ribozymes

In Figure 5 are shown the interaction diagrams for the published crystal structures of small ribozymes. The Diels-Adler ribozyme (Figure 5a) (37) displays three non-Watson–Crick pairs and a two base pair pseudoknot at the 5′ end (38). The hepatitis delta virus ribozyme presents a pseudoknotted helix at its 3′ end (39) and an additional two base pair pseudoknot below the cleavage site (Figure 5b) (40,41). This latter pseudoknot is stabilized by two base triples. The single-strand leading to the 3′ end pseudoknot forms A-minor contacts with a co-axially stacked helix (40,42). It is striking to see how four nucleotides of the 6-membered hairpin loop L3 form the stable UNCG tetraloop characterized by a G in the *syn* conformation (43). The hairpin ribozyme (Figure 5c) has a four-way junction with interactions mediated by structured internal loops between two of the helices (44). The ribose zipper (between nucleotides 10/11 and 24/25) critical for the folding (45,46) is clearly apparent. The loop E like structure of loop B (46,47) displays the characteristic set of non-Watson–Crick pairs. It is interesting to notice, however, that the bulging U42 forms a *cis* Watson–Crick pair with A23.

The two main crystal structures of the hammerhead ribozymes are shown on Figure 5d and 5e (48,49). Recently, it has been clear that the structure shown on Figure 5d does not correspond to the active structure of the hammerhead

ribozyme. The new crystal structure of the hammerhead ribozyme from *Schistosome mansoni* (49) is stunning in several respects. Interestingly, the loop–loop interactions are dominated by *trans* Watson–Crick/Hoogsteen pairs. A comparison between the structures of the hairpin and hammerhead ribozymes reveals an intriguing pattern (Figure 5f). In both cases, cleavage occurs between two *trans* Hoogsteen/Sugar–edge pairs (tandem sheared base pairs) with an intermediate residue forming a canonical Watson–Crick G=C pair.

The riboswitches

Five riboswitch structures have now been determined, all complexed with activating ligand (Figure 6): two structures of the guanine riboswitch (50,51), the SAM riboswitch (52) and two structures of the thiamine pyrophosphate (TPP) riboswitch (53,54). Each of those structures contains at least one motif found recurrently in other RNAs: a common three-way junction in the guanine riboswitch (Figure 6a) (55,56); a kink-turn motif (57,58) in the SAM riboswitch (Figure 6b); a T-like loop (33) in the TPP riboswitch (Figure 6c). As in the *Schistosome* hammerhead, the *trans* Watson–Crick/Hoogsteen pairs are frequent in the loop–loop interactions of the guanine riboswitch. Interestingly, the latter type of base pairs leads to a local antiparallel orientation of the strands. Although the SAM riboswitch displays a standard A-minor type of contact with alternating type I and type II pairs (59), the TPP riboswitch presents an example of the unfrequent double type I/type I contact (Figure 7) (60).

### The transpeptidylation center of the 23S rRNA structure

The transpeptidylation center is a multibranch loop in domain V of the 23S rRNA (61). Two interaction diagrams are shown on Figures 8 and 9. Figure 8 corresponds to the structure in the *Haloarcula marismortui* 23S rRNA and shows also the contacts with the -CCA end of a tRNA in the A site and P site (62). Figure 9 corresponds to the *E.coli* structure (36). To help the understanding, the usual secondary structure and a 3D representation are shown in Figure 10. The compactness of the PTC multibranch loop leads to a T-like shape with the parallel stacks of H89 and H90/H91 forming the vertical branch of the 'T', helices H73 and H93 forming the horizontal branch of the 'T' and the active nucleotides at the center of the 'T'. Helices H90/H91/H92 adopt the frequent family C fold of three-way junction (56), which brings the apical loop of H92 close to the important residues U2506 and U2585 thereby forming part of the binding site for the A site tRNA on the 23S rRNA. The center contains three occurrences (helices H73, H74 and H91) of a motif described above (Figure 4) which is present in P10/P11 of RNase P RNAs. One of those motifs includes A2451 of the active site (H74) and forms part of the binding site of P site tRNA on the 23S rRNA.

The 16S rRNA structures

The interaction network diagrams for the 16S rRNA of *T.thermophilus* and *E.coli* are shown in Figures 11 and 12. The characteristic feature of the 16S rRNA is the clear decomposition into large domains (35): the 5′, the central,

the 3′major and the 3′minor domains. The 5′domain forms the 'body', the 3′major domain forms the 'head' and the central domain forms the 'platform' (63,64). In the interaction network diagram, the 5′pseudoknot (formed between the loop of h1 and the junction between helices h27 and h28) is set at the center because it interconnects the 5′domain, the 3′major and the central domain. The 3′major domain has pronounced intra-domain connections and only two inter-domain
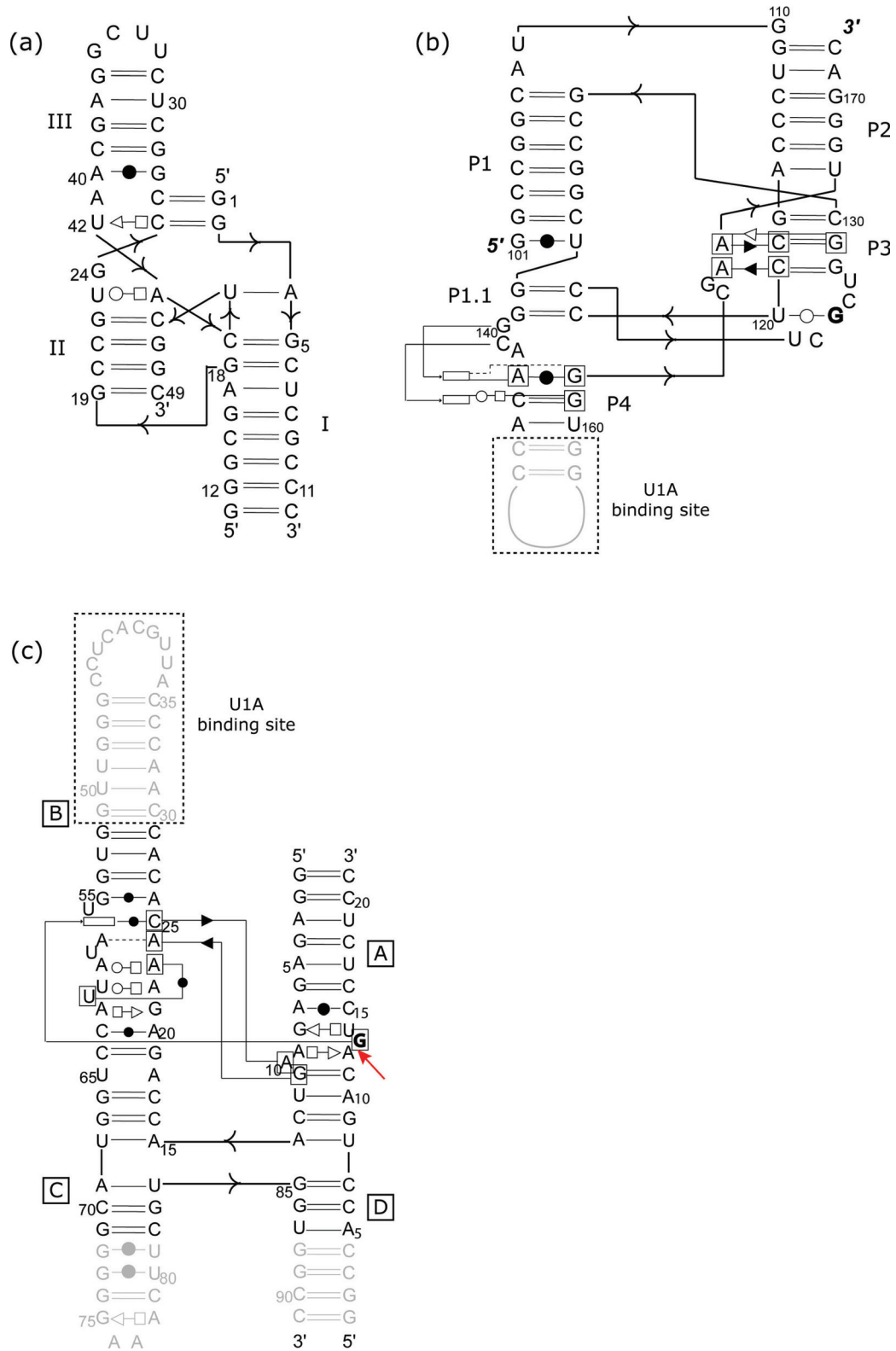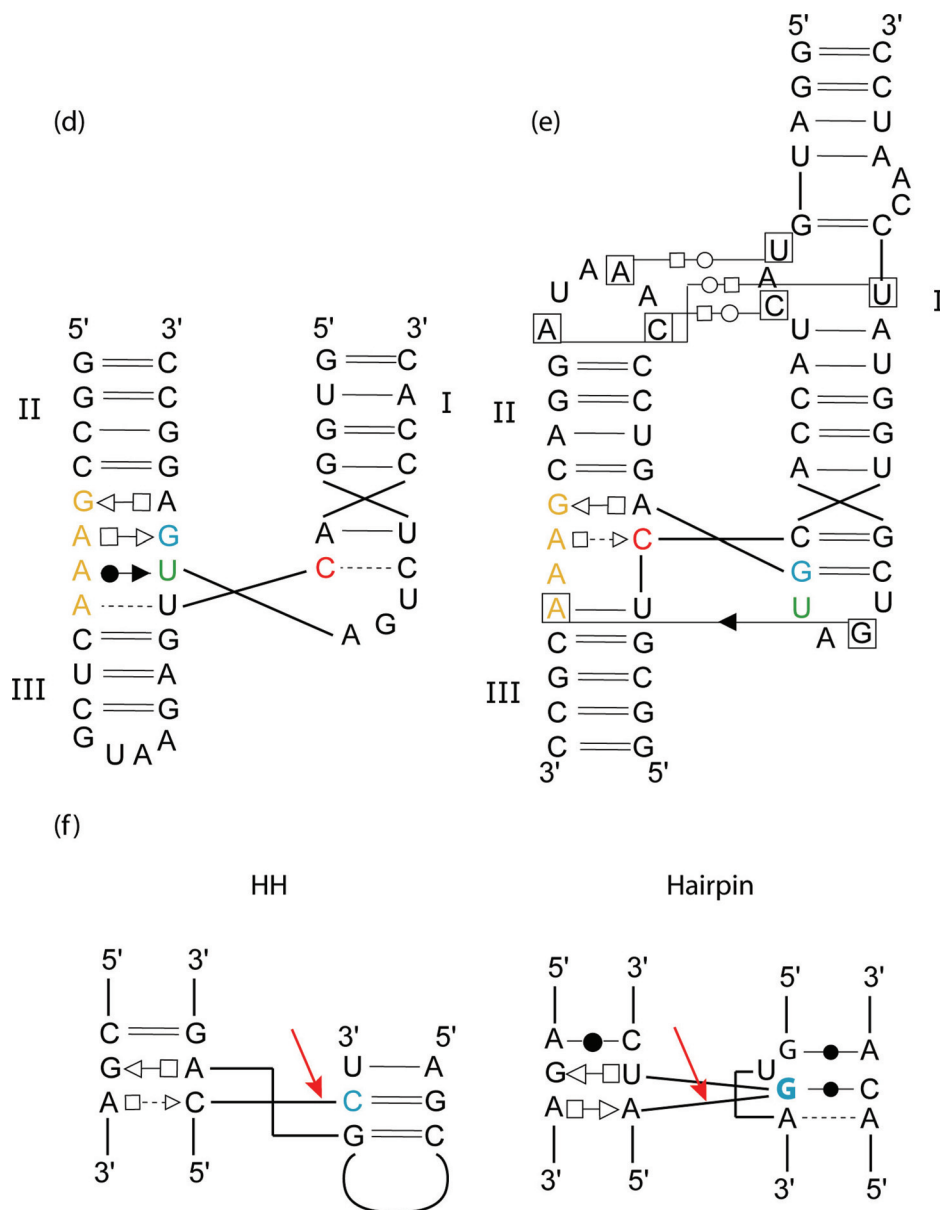
Figure 5. Interaction network diagrams for (**A**) the Diels-Adler ribozyme (37); (**B**) the hepatitis delta virus ribozyme (40); (**C**) the hairpin ribozyme (44); (**D**) the minimal hammerhead ribozyme (48); (**e**) the *Schistosome* hammerhead ribozyme (49); (**F**) comparison of the active sites of the hairpin and hammerhead ribozymes. The single ended red arrows indicate the cleavage site. The colored nucleotides indicate those that change state between the minimal and complete hammerhead ribozymes.

connections (h36 with the 5′ pseudoknot and h28, which is stacked with the 5′pseudoknot, interacts with the base of helix h44 in the 3′minor domain). A single primary binding protein (65–67), S7, interacts in the multibranch junctions which follow helix h28. On the contrary, the central domain has less intra-domain contacts but several inter-domain contacts: one with the 3′minor (h24 with h45); two with the 5′pseudoknot (h26a and h27), and four with the 5′domain (h19 with h12, h27 with h11 and the long helix h21 with h4 and h12). Two primary binding proteins bind to the central domain: proteins S8 and S15. The 5′domain presents rather extensive intra-domain contacts, has two primary binding proteins S4 and S17 binding to it, and besides the four

inter-domain contacts with the central domain presents additional links with the long decoding helix h44 (via h13 in both structures and also via h8 in the *E.coli* structure). In short, the head is rather free to move with respect to the other domains, while the platform and body could move in concert with the movements of the body transmitted to the penultimate helix h44. These long-range contacts therefore reflect the known conformational changes occurring in the ribosome (9,36,68,69).

The *E.coli* structure has revealed the molecular details of three important RNA–RNA inter-subunit contacts (36,69–71). The inter-subunit bridge, B7a, occurs between h23 and H68 (see Figure 13). The inter-subunit bridge, B2a, occurs
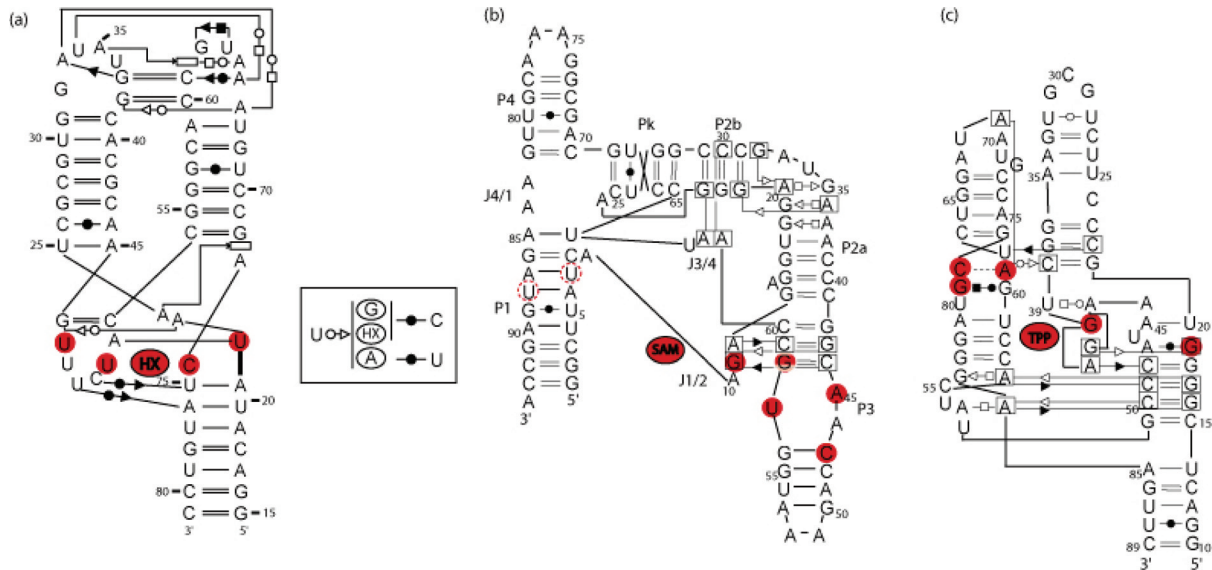
Figure 6. Interaction network diagrams for (**A**) the guanine riboswitch in presence of hypoxanthine (HX) (51) and on the insert at the right the variations in presnce of guanine and adenine (50); (**B**) the SAM riboswitch (52); (**C**) the thiamine pyrophosphate riboswitch (54). The filled red circles indicate those bases contacted by the ligand in each case and the dashed red circles indicate van der Waals contacts (52).
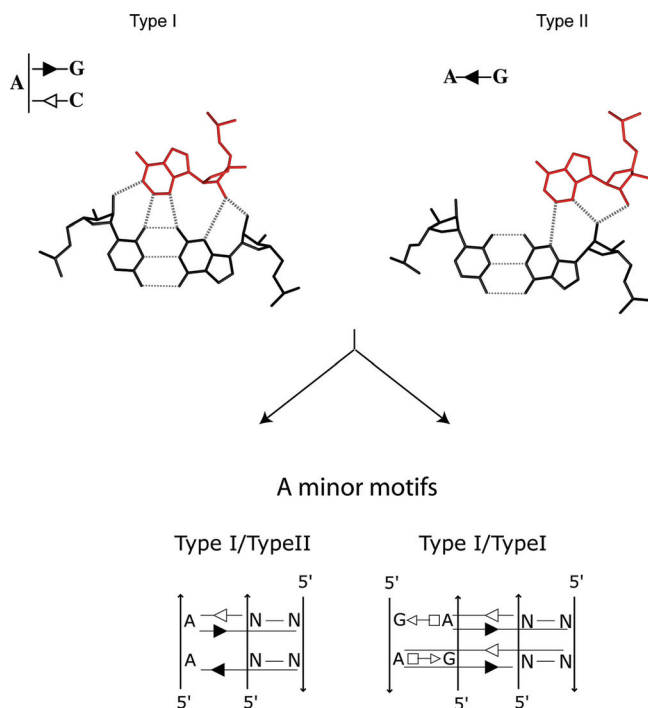

Figure 7. Example of the type I and type II A-minor contacts (59) and their observed associations into motifs. The most frequently observed one is typeII/typeI from the 5′ end to the 3′ end. The double combination typeI/typeI does occur also but it is less frequent.

between the A site of h44 and the loop of H69, while the inter-subunit bridge, B3, occurs below the A site of h44 and with helix H71 (see Figure 14). The B7a bridge is essentially a *cis* sugar-edge/sugar-edge base pair which interestingly involves an unusual K-turn in the 16S [KT-23, see (58)]. The B2a bridge involves the A site of the 16S and especially an invariant UoU pair. Although the conformation of that UoU pair in the free state is not precisely known, in presence of antibiotics of the aminoglycoside family, the conformation of the UoU pair is always bifurcated with the O4 of U1406 pointing to the N3-H of U1495 (72,73). Thus, the effects of antibiotic binding could be transmitted to the large subunit via the bridge B2a. The B3 bridge exploits the less frequent A-minor motif with two type I contacts (see Figure 7) between a tandem *trans* Hoogsteen/sugar-edge GoA pairs (sheared) and two stacked G=C pairs of H71.

Structural consequences of modularity

The modular architecture of folded RNAs implies that distances between interacting parts are conserved in functionally homologous molecules. Thus, elongator tRNAs have generally four base pairs in the dihydrouridine helix and five in the thymine helix and the maintenance of those distances allows for specific base pairing contacts between a conserved set of two Gs in the dihydrouridine loop and the conserved TΨCG segment of the thymine loop. Also, in group I introns (see Figures 2 and Supplementary Figure S1), the distance between the GoU pair at the cleavage site of helix P1 and the GAAA capping loop of the co-axially stacked P2 helix is always of 12 bp. This invariant distance guarantees formation of the contact between GAAA and its receptor in P8 (3bp below the interface with helix P3) (7).

The 16S rRNA presents similar examples of conserved numbered of base pairs between interacting modules. The distance between the two sets of interacting As in helix h21 is 7 bp (with two unpaired nucleotides); the two sets of As contact identical positions of base pairs in h4 and h12. In helix h44, the distance between the motif involved in bridge B3 (tandem *trans* Hoogsteen/sugar-edge GoA pairs followed
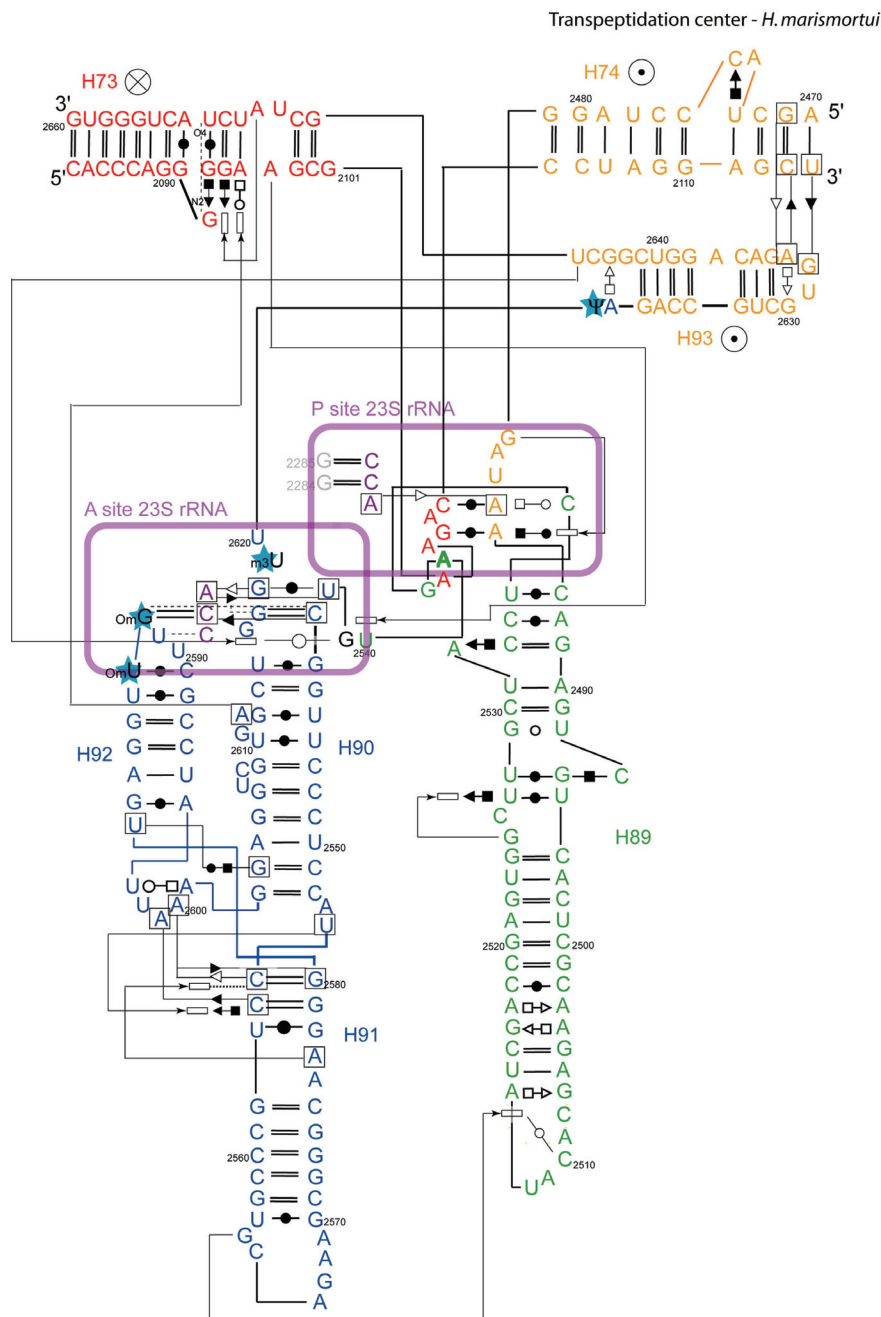
Figure 8. Interaction network diagrams for the transpeptidylation center of the 23S rRNA of *H. marismortui* (62) (69,94) showing the contacts with the –CCA ends of A- and P-site tRNAs. The circles with a central dot indicates that the helix points towards the reader and the circle with a cross the reverse. Known modifications are indicated by blue stars (95).

by a GoU pair) and the motif forming an inter-domain contact with h13 is of 12 bp (despite the fact that the motif of h44 interacting with h13 is not conserved). In helix h17, the distance between the interface with h16 and the A contacting a G=C pair in h15 is 13 bp (not counting bulged residues); at the same time, the distance between the same G=C pair in h15 and the end of the co-axially stacked h4 is 11 bp. Both helices h22 and h26 start with a *trans* Hoogsteen/sugar-edge base pair and the sixth base pair

above it in h22 forms a sugar-edge/sugar-edge base pair with the eighth base pair in h26 of both 16S rRNA structures (the overall length of h26 is variable).

The capping GNRA loop of h8 forms an identical contact with the last base pair of h14 and 6 bp from the capping GNRA a striking motif in h8 forms identical contacts with an identical motif at the base of h6. However, beyond that conserved motif, the length of h6 is variable. In *Thermus thermophilus* 16S rRNA, helix h9 is longer than in
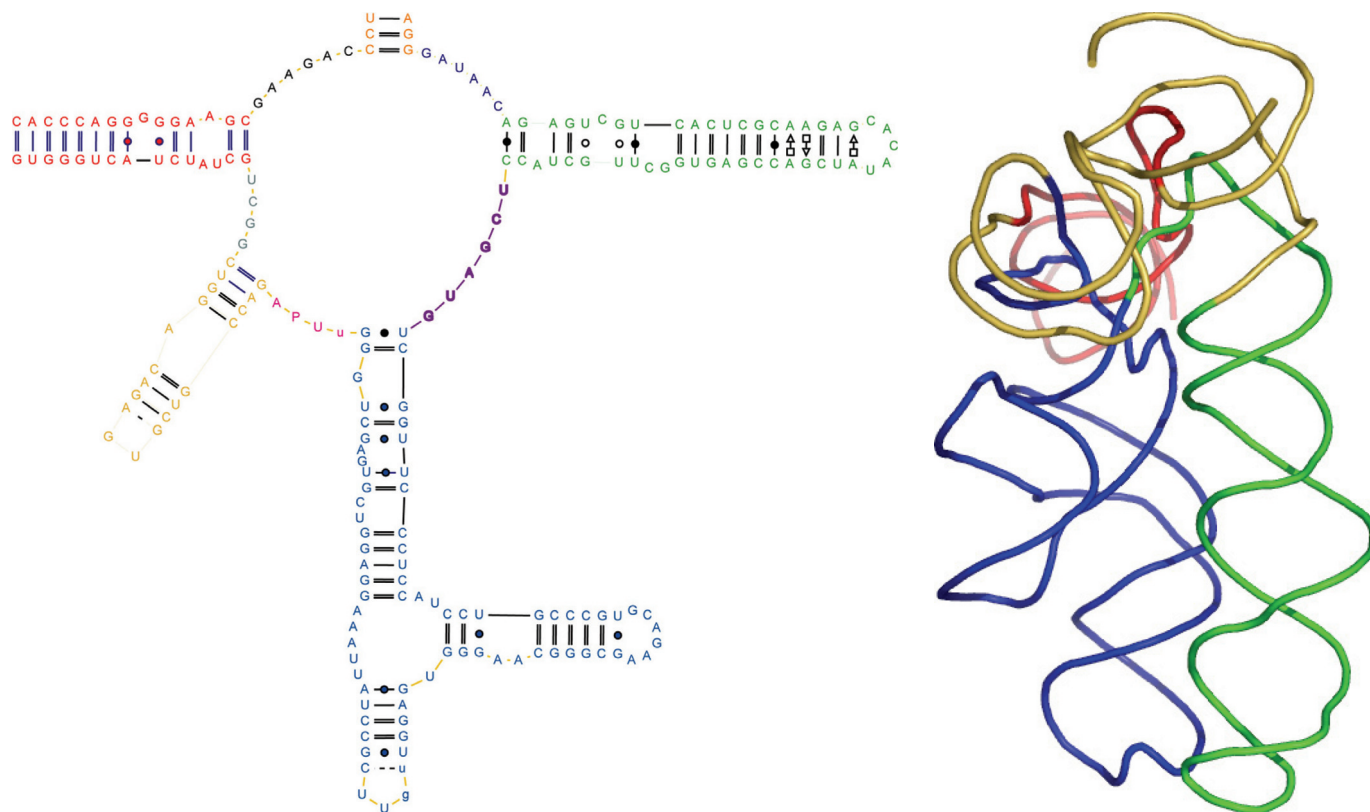
Figure 9. Interaction network diagrams for the transpeptidylation center of the 23S rRNA of *E.coli* (36). Known modifications are indicated by blue stars (96).

E.coli 16S rRNA, but there is an additional residue in h7 which forms a *trans* Watson–Crick/sugar-edge within the capping loop of h9. At the same time, helix h10 is shorter in *T.thermophilus*, which allows an unpaired U of h8 to form a *trans* Hoogsteen/sugar-edge with the first base pair of h10. There is an interdomain contact between the tip of h44 and the beginning of h8 in *E.coli* which is absent in *T.thermophilus* where h44 is too short to form such contacts. Such conservations in distances and contacts are significant considering the distance between the two bacterial species to which these 16S rRNAs belong to.

Similarities in motifs and contacts made are apparent from such diagrams. For example, the striking motif in h8, forming a tight contact with h6 (74), is found also in h41 where it forms also a tight contact with h43. Helices h20 and h24 both contain an identical motif of three consecutive non-Watson–Crick pairs in *trans* (Hoogsteen/sugar-edge, Watson–Crick/sugar-edge). Surprisingly, they form no further contacts but, two base pairs above, two Watson–Crick pairs are engaged in A-minor interactions (in which the As originate from different types of motifs in h23a and h23b, respectively).

Figure 10. The conventional secondary structure and the tertiary fold of the transpeptidylation center of *H.marismortui* (62). The lowercase letters indicate nucleotides which are modified.

CONCLUSIONS

The wiring diagrams representing the networks of all pairwise base–base interactions in architectures of structured RNAs illustrate several principles of RNA folding. A striking observation is the ubiquitous presence of RNA motifs with a defined order of non-Watson–Crick base pairs. They are found recurrently in structured RNAs with different functions and stemming from the three kingdoms of life. Some motifs are more frequent and more pervasive than others (e.g. loop E or K-turn motifs). For those motifs, we have progressed in the establishment of their sequence constraints (58,75). Although a complete survey of RNA motifs is not yet achieved, their number is clearly limited (76). The task of enumerating and cataloguing RNA motifs is problematic because of the lack of accepted definition on the nature and size of RNA motifs (2,77). This is compounded by the observation that most RNA motifs can be assembled like Russian dolls such that only portions of RNA motifs can occur (78,79). For example, the variants of the T-loop motifs (33,80,81) or those of the loop E family (79,82).

In statistical physics, complex systems are analysed as networks of interactions between the components of the system. Recently, network theory has permeated all realms of science (83,84). Some major real-life networks have striking properties. They are said 'scale-free', meaning that their

properties are controlled by a small number of highly connected nodes. In such scale-free networks, any two nodes can be connected by a very small number of intermediate nodes, the small-world effect. The main attribute of such networks is the presence of clusters of local interactions with long-range interactions between the clusters (85,86). Visually, in the wiring diagrams of the 16S rRNAs, the clusters are apparent as well as the connections between them. Such diagrams appear like hierarchical networks which have the properties of scale-free networks with embedded modularity (84,87). However, as shown for protein structures (88–90), in the process of conceptualising macromolecules as networks several simplifications and hypotheses have to be formulated and evaluated. How should the number and types of contacts be considered? Should we consider all atomic contacts or only hydrogen bonds? Such analysis has not been performed yet for RNA systems. The hypothesis that RNA three-dimensional structures form scale-free networks could lead to valuable knowledge on the evolution pathways of RNA molecules. The origin of scale-free topology is rooted in the self-organization of networks because networks grow by preferential linking and attachment of links to nodes with a high number of connections (86). The reverse process could therefore lead to the set of nucleotides constituting the common ancestor of 16S rRNA.
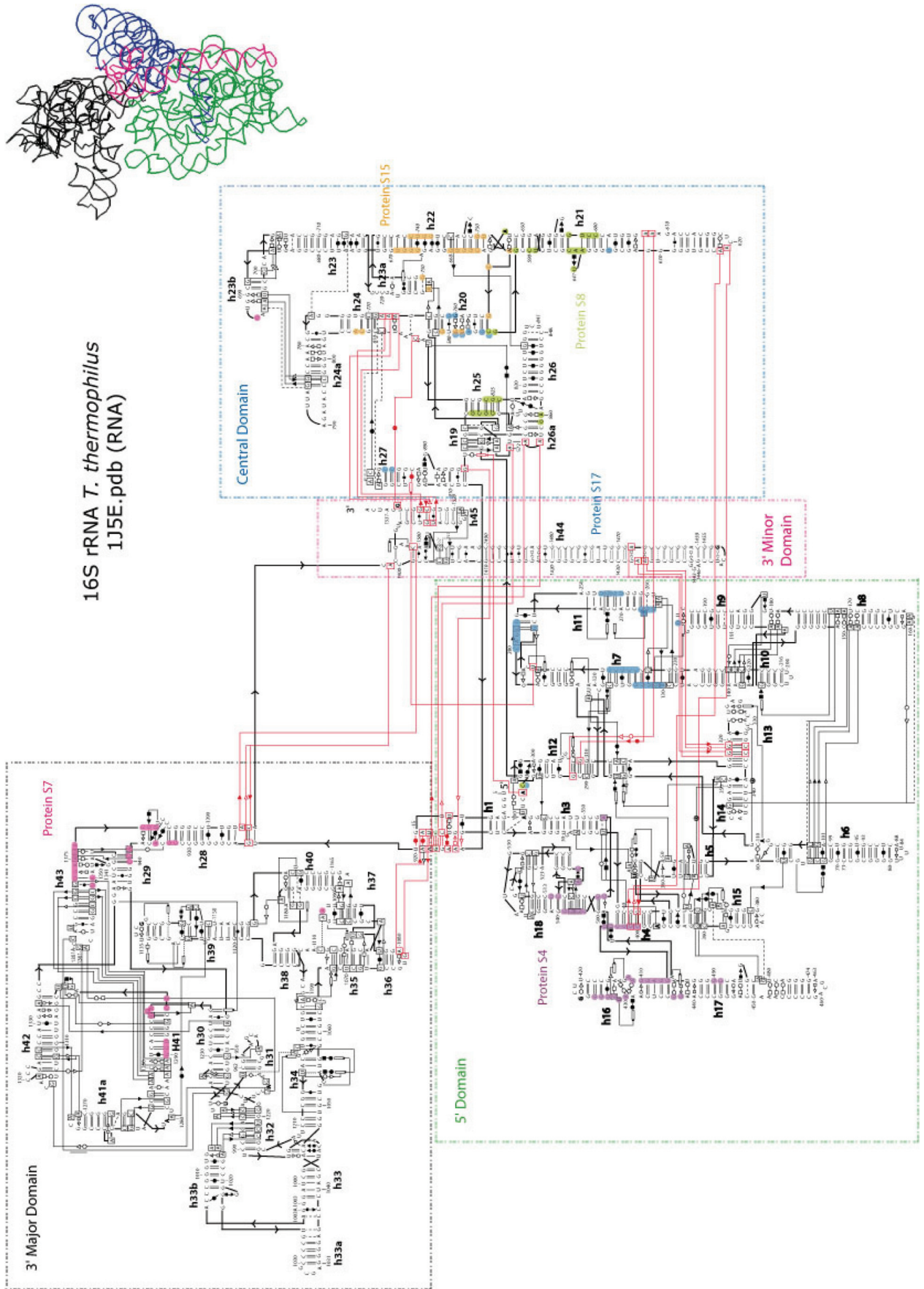
Figure 11. Interaction network diagrams for the 16S rRNA of *T.thermophilus* in the 30S subunit (35). The contacts of the primary proteins (S4, S7, S8, S15, S17) are colored on the contacted nucleotides following the indicated color code (97).
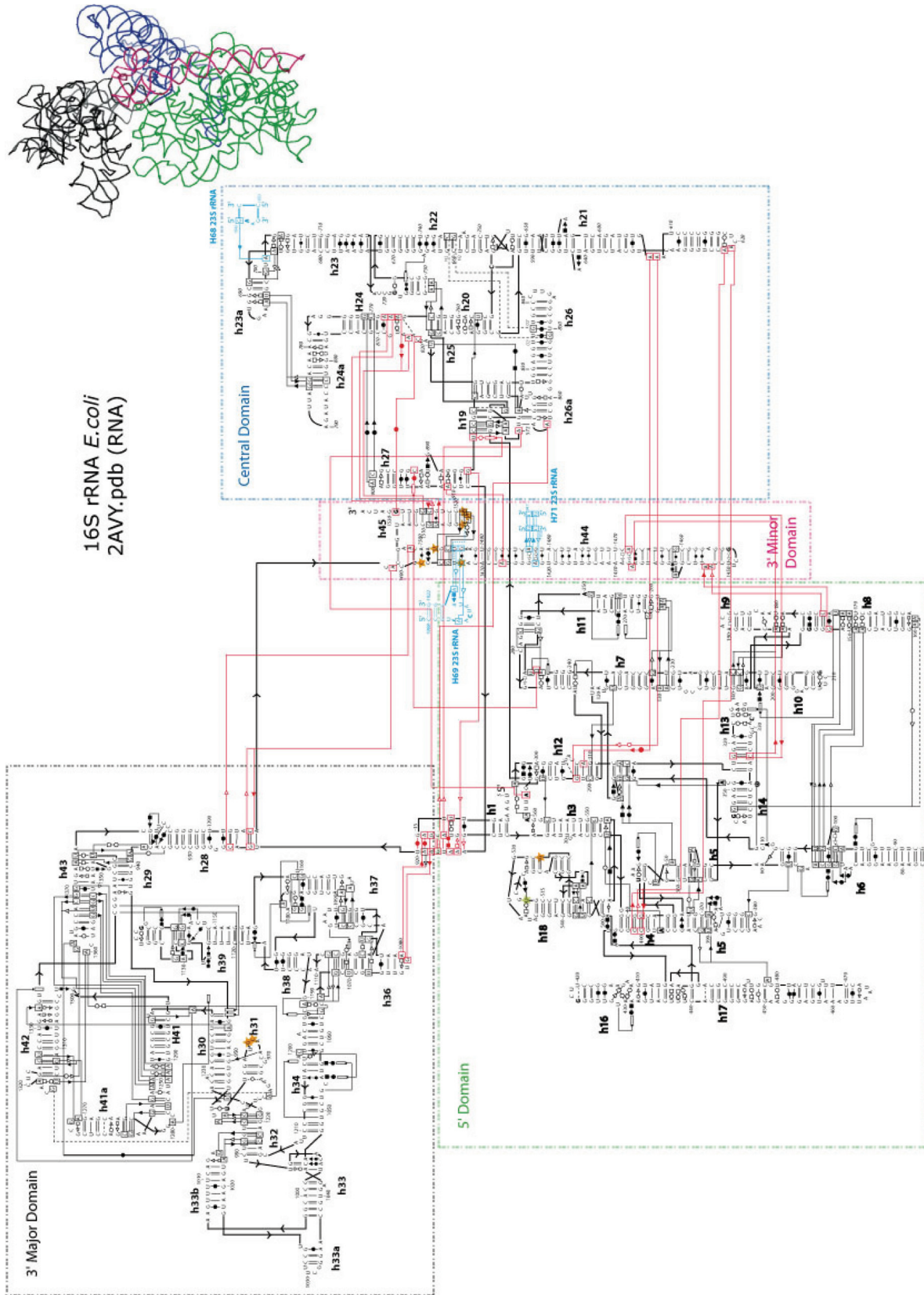
Figure 12. Interaction network diagrams for the for the 16S rRNA of *E.coli* in the 70S ribosome (36). The three inter-subunit bridges RNA mediated are shown in blue (36). The modifications are shown by stars (in orange: 2′-O-methylations; green: pseudouridines) (98).
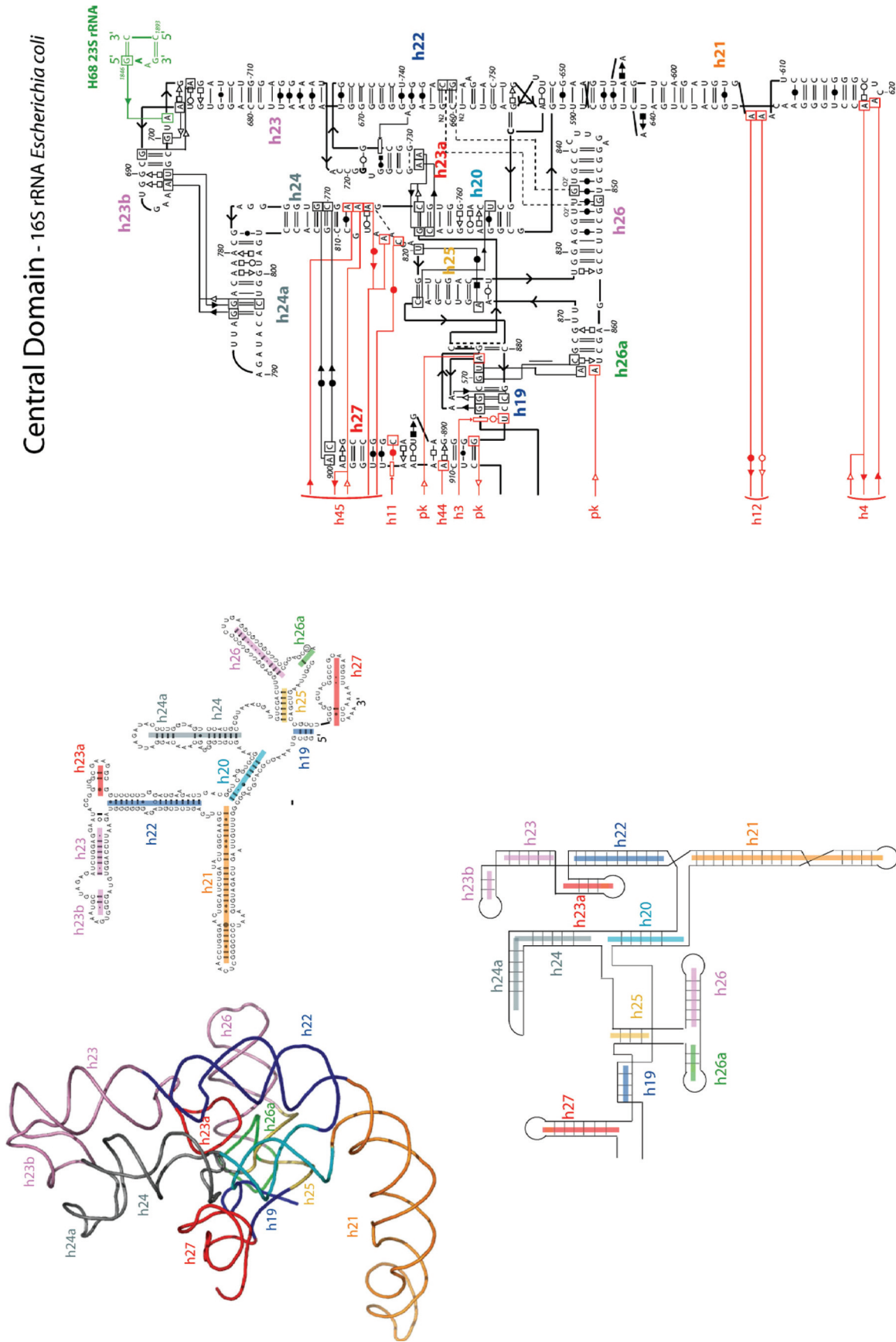
Figure 13. The central domain of the 16S rRNA in the *E.coli* ribosome (36) showing the tertiary fold, the conventional secondary structure, the view emphasizing co-axial stacking and the network diagram.
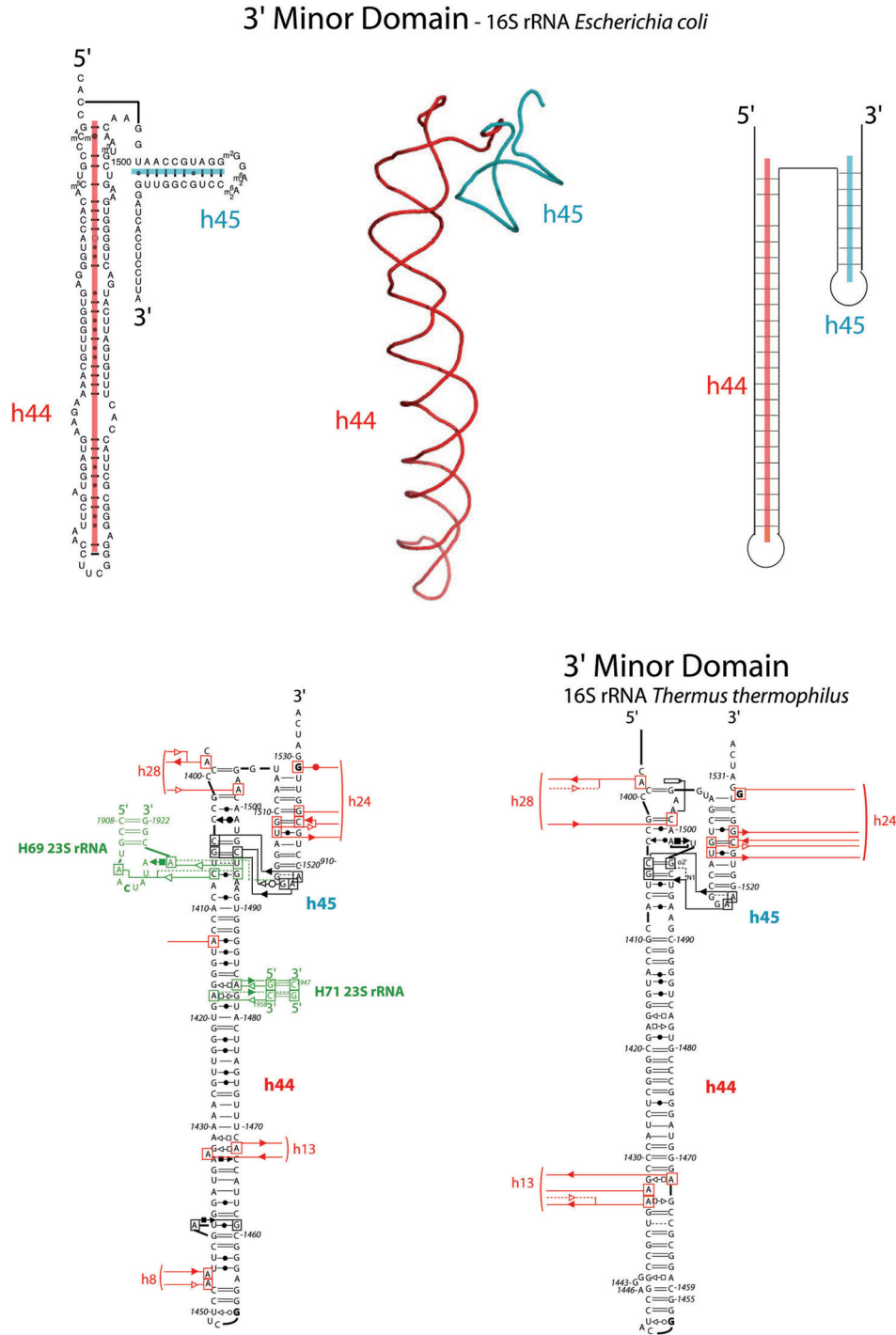
Figure 14. The 3′minor domain of the 16S rRNA in the *E.coli* ribosome (36) showing the tertiary fold, the conventional secondary structure, the view emphasizing co-axial stacking and the network diagram. For comparison, the 3′minor domain of the 16S rRNA in the *T.thermophilus* 30S particle (35) is also shown.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.

2. Leontis,N.B., Lescoute,A. and Westhof,E. (2006) The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.*, **16**, 279–287.

3. Yang,H., Jossinet,F., Leontis,N., Chen,L., Westbrook,J., Berman,H. and Westhof,E. (2003) Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res.*, **31**, 3450–3460.

4. Gendron,P., Lemieux,S. and Major,F. (2001) Quantitative analysis of nucleic acid three-dimensional structures. *J. Mol. Biol.*, **308**, 919–936.

5. Lemieux,S. and Major,F. (2002) RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res.*, **30**, 4250–4263.

6. Jossinet,F. and Westhof,E. (2005) Sequence to Structure (S2S): display, manipulate and interconnect RNA data from sequence to structure. *Bioinformatics*, **21**, 3320–3321.

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

8. Lehnert,V., Jaeger,L., Michel,F. and Westhof,E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila ribozyme*. *Chem. Biol.*, **3**, 993–1009.

9. Adams,P.L., Stahley,M.R., Gill,M.L., Kosek,A.B., Wang,J. and Strobel,S.A. (2004) Crystal structure of a group I intron splicing intermediate. *RNA*, **10**, 1867–1887.

10. Golden,B.L., Kim,H. and Chase,E. (2005) Crystal structure of a phage Twort group I ribozyme-product complex. *Nature Struct. Mol. Biol.*, **12**, 82–89.

11. Guo,F., Gooding,A.R. and Cech,T.R. (2004) Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol. Cell*, **16**, 351–362.

12. Woodson,S.A. (2005) Structure and assembly of group I introns. *Curr. Opin. Struct. Biol.*, **15**, 324–330.

13. Vicens,Q. and Cech,T.R. (2006) Atomic level architecture of group I introns revealed. *Trends Biochem. Sci.*, **31**, 41–51.

14. Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276–1285.

15. Costa,M. and Michel,F. (1997) Rules for RNA recognition of GNRA tetraloops deduced by *in vitro* selection: comparison with *in vivo* evolution. *EMBO J.*, **16**, 3289–3302.

16. Costa,M., Michel,F. and Westhof,E. (2000) A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J.*, **19**, 5007–5018.

17. Khvorova,A., Lescoute,A., Westhof,E. and Jayasena,S.D. (2003) Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature Struct. Biol.*, **10**, 708–712.

18. Penedo,J.C., Wilson,T.J., Jayasena,S.D., Khvorova,A. and Lilley,D.M. (2004) Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements. *RNA*, **10**, 880–888.

19. Canny,M.D., Jucker,F.M., Kellogg,E., Khvorova,A., Jayasena,S.D. and Pardi,A. (2004) Fast cleavage kinetics of a natural hammerhead ribozyme. *J. Am. Chem. Soc.*, **126**, 10848–10849.

20. Michel,F., Jacquier,A. and Dujon,B. (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie*, **64**, 867–881.

21. Burke,J.M., Belfort,M., Cech,T.R., Davies,R.W., Schweyen,R.J., Shub,D.A., Szostak,J.W. and Tabak,H.F. (1987) Structural conventions for group I introns. *Nucleic Acids Res.*, **15**, 7217–7221.

22. Cech,T.R., Damberger,S.H. and Gutell,R.R. (1994) Representation of the secondary and tertiary structure of group I introns. *Nature Struct. Biol.*, **1**, 273–280.

23. Michel,F., Ellington,A.D., Couture,S. and Szostak,J.W. (1990) Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. *Nature*, **347**, 578–580.

24. Strobel,S.A., Adams,P.L., Stahley,M.R. and Wang,J. (2004) RNA kink turns to the left and to the right. *RNA*, **10**, 1852–1854.

25. James,B.D., Olsen,G.J., Liu,J. and Pace,N.R. (1988) The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell*, **52**, 19–26.

26. Brown,J.W. (1999) The Ribonuclease P Database. *Nucleic Acids Res.*, **27**, 314.

27. Harris,M.E., Kazantchev,A.V., Chen,J.L. and Pace,N.R. (1997) Analysis of the tertiary structure of the ribonuclease P ribozyme-substrate complex by site-specific photoaffinity crosslinking. *RNA*, **3**, 561–574.

28. Massire,C., Jaeger,L. and Westhof,E. (1998) Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J. Mol. Biol.*, **279**, 773–793.

29. Krasilnikov,A.S., Xiao,Y., Pan,T. and Mondragon,A. (2004) Basis for structural diversity in homologous RNAs. *Science*, **306**, 104–107.

30. Krasilnikov,A.S., Yang,X., Pan,T. and Mondragon,A. (2003) Crystal structure of the specificity domain of ribonuclease P. *Nature*, **421**, 760–764.

31. Kazantsev,A.V., Krivenko,A.A., Harrington,D.J., Holbrook,S.R., Adams,P.D. and Pace,N.R. (2005) Crystal structure of a bacterial ribonuclease P RNA. *Proc. Natl Acad. Sci. USA*, **102**, 13392–13397.

32. Torres-Larios,A., Swinger,K.K., Krasilnikov,A.S., Pan,T. and Mondragon,A. (2005) Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*, **437**, 584–587.

33. Krasilnikov,A.S. and Mondragon,A. (2003) On the occurrence of the T-loop RNA folding motif in large RNA molecules. *RNA*, **9**, 640–643.

34. Hermann,T. and Patel,D.J. (2000) RNA bulges as architectural and recognition motifs. *Structure*, **8**, R47–R54.

35. Wimberly,B.T., Brodersen,D.E., Clemons,W.M., Jr, Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.

36. Schuwirth,B.S., Borovinskaya,M.A., Hau,C.W., Zhang,W., Vila-Sanjurjo,A., Holton,J.M. and Cate,J.H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution. *Science*, **310**, 827–834.

37. Serganov,A., Keiper,S., Malinina,L., Tereshko,V., Skripkin,E., Hobartner,C., Polonskaia,A., Phan,A.T., Wombacher,R., Micura,R. *et al.* (2005) Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation. *Nature Struct. Mol. Biol.*, **12**, 218–224.

38. Keiper,S., Bebenroth,D., Seelig,B., Westhof,E. and Jaschke,A. (2004) Architecture of a Diels-Alderase ribozyme with a preformed catalytic pocket. *Chem. Biol.*, **11**, 1217–1227.

39. Perrotta,A.T. and Been,M.D. (1991) A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature*, **350**, 434–436.

40. Ferre-D'Amare,A.R., Zhou,K. and Doudna,J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature*, **395**, 567–574.

41. Wadkins,T.S., Perrotta,A.T., Ferre-D'Amare,A.R., Doudna,J.A. and Been,M.D. (1999) A nested double pseudoknot is required for self-cleavage activity of both the genomic and antigenomic hepatitis delta virus ribozymes. *RNA*, **5**, 720–727.

42. Tanner,N.K., Thill,G., Petit-Kostas,E., Crain-Denoyelle,A.M. and Westhof,E. (1994) A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr. Biol.*, **4**, 488–498.

43. Ennifar,E., Nikulin,A., Tishchenko,S., Serganov,A., Nevskaya,N., Garber,M., Ehresmann,B., Ehresmann,C., Nikonov,S. and Dumas,P. (2000) The crystal structure of UUCG tetraloop. *J. Mol. Biol.*, **304**, 35–42.

44. Rupert,P.B. and Ferre-D'Amare,A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780–786.

45. Chowrira,B.M., Berzal-Herranz,A., Keller,C.F. and Burke,J.M. (1993) Four ribose 2′-hydroxyl groups essential for catalytic function of the hairpin ribozyme. *J. Biol. Chem.*, **268**, 19458–19462.

46. Earnshaw,D.J., Masquida,B., Müller,S., Sigurdsson,S.T., Eckstein,F., Westhof,E. and Gait,M.J. (1997) Inter-domain cross-linking and molecular modelling of the hairpin ribozyme. *J. Mol. Biol.*, **274**, 197–212.

47. Butcher,S.E., Allain,F.H.-T. and Feigon,J. (1999) Solution structure of the loop B from the hairpin ribozyme. *Nature Struct. Biol.*, **6**, 212–216.

48. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an all-RNA hammerhead ribozyme. *Nucleic Acids Symp. Ser.*, **34**, 214–216.

49. Martick,M. and Scott,W.G. (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell*, **126**, 309–320.

50. Serganov,A., Yuan,Y.R., Pikovskaya,O., Polonskaia,A., Malinina,L., Phan,A.T., Hobartner,C., Micura,R., Breaker,R.R. and Patel,D.J. (2004) Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.*, **11**, 1729–1741.

51. Batey,R.T., Gilbert,S.D. and Montange,R.K. (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature*, **432**, 411–415.

52. Montange,R.K. and Batey,R.T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*, **441**, 1172–1175.

53. Thore,S., Leibundgut,M. and Ban,N. (2006) Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science*, **312**, 1208–1211.

54. Serganov,A., Polonskaia,A., Phan,A.T., Breaker,R.R. and Patel,D.J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*, **441**, 1167–1171.

55. Lescoute,A. and Westhof,E. (2005) Riboswitch structures: purine ligands replace tertiary contacts. *Chem. Biol.*, **12**, 10–13.

56. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA*, **12**, 83–93.

57. Klein,D.J., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.*, **20**, 4214–4221.

58. Lescoute,A., Leontis,N.B., Massire,C. and Westhof,E. (2005) Recurrent structural RNA motifs, Isosteriality Matrices and sequence alignments. *Nucleic Acids Res.*, **33**, 2395–2409.

59. Nissen,P., Ippolito,J.A., Ban,N., Moore,P.B. and Steitz,T.A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl Acad. Sci. USA*, **98**, 4899–4903.

60. Lescoute,A. and Westhof,E. (2006) The A-minor motifs in the decoding recognition process. *Biochimie*, **88**, 993–999.

61. Barta,A., Steiner,G., Brosius,J., Noller,H.F. and Kuechler,E. (1984) Identification of a site on 23S ribosomal RNA located at the peptidyl transferase center. *Proc. Natl Acad. Sci. USA*, **81**, 3607–3611.

62. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905–920.

63. Lata,K.R., Agrawal,R.K., Penczek,P., Grassucci,R., Zhu,J. and Frank,J. (1996) Three-dimensional reconstruction of the *Escherichia coli* 30S ribosomal subunit in ice. *J. Mol. Biol.*, **262**, 43–52.

64. Clemons,W.M., Jr, May,J.L., Wimberly,B.T., McCutcheon,J.P., Capel,M.S. and Ramakrishnan,V. (1999) Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature*, **400**, 833–840.

65. Mizushima,S. and Nomura,M. (1970) Assembly mapping of 30S ribosomal proteins from E.coli. *Nature*, **226**, 1214.

66. Nomura,M. and Erdmann,V.A. (1970) Reconstitution of 50S ribosomal subunits from dissociated molecular components. *Nature*, **228**, 744–748.

67. Grondek,J.F. and Culver,G.M. (2004) Assembly of the 30S ribosomal subunit: positioning ribosomal protein S13 in the S7 assembly branch. *RNA*, **10**, 1861–1866.

68. Serdyuk,I., Baranov,V., Tsalkova,T., Gulyamova,D., Pavlov,M., Spirin,A. and May,R. (1992) Structural dynamics of translating ribosomes. *Biochimie*, **74**, 299–306.

69. Yusupov,M.M., Yusupova,G.Z., Baucom,A., Lieberman,K., Earnest,T.N., Cate,J.H. and Noller,H.F. (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science*, **292**, 883–896.

70. Frank,J., Verschoor,A., Li,Y., Zhu,J., Lata,R.K., Radermacher,M., Penczek,P., Grassucci,R., Agrawal,R.K. and Srivastava,S. (1995) A model of the translational apparatus based on a three-dimensional reconstruction of the *Escherichia coli* ribosome. *Biochem. Cell Biol.*, **73**, 757–765.

71. Cate,J.H., Yusupov,M.M., Yusupova,G.Z., Earnest,T.N. and Noller,H.F. (1999) X-ray crystal structures of 70S ribosome functional complexes. *Science*, **285**, 2095–2104.

72. Vicens,Q. and Westhof,E. (2001) Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site. *Structure*, **9**, 647–658.

73. Francois,B., Russell,R.J., Murray,J.B., Aboul-ela,F., Masquida,B., Vicens,Q. and Westhof,E. (2005) Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. *Nucleic Acids Res.*, **33**, 5677–5690.

74. Noller,H.F. (2005) RNA structure: reading the ribosome. *Science*, **309**, 1508–1514.

75. Leontis,N.B. and Westhof,E. (1998) A common motif organizes the structure of multi-helix loops in 16S and 23S ribosomal RNAs. *J. Mol. Biol.*, **283**, 571–583.

76. Moore,P.B. (1999) Structural motifs in RNA. *Annu. Rev. Biochem.*, **68**, 287–300.

77. Leontis,N.B., Altman,R.B., Berman,H.M., Brenner,S.E., Brown,J.W., Engelke,D.R., Harvey,S.C., Holbrook,S.R., Jossinet,F., Lewis,S.E. *et al.* (2006) The RNA Ontology Consortium: an open invitation to the RNA community. *RNA*, **12**, 533–541.

78. Westhof,E. and Fritsch,V. (2000) RNA folding: beyond Watson–Crick pairs. *Structure*, **8**, R55–R65.

79. Leontis,N.B., Stombaugh,J. and Westhof,E. (2002) Motif prediction in ribosomal RNAs lessons and prospects for automated motif prediction in homologous RNA molecules. *Biochimie*, **84**, 961–973.

80. Leontis,N.B. and Westhof,E. (2003) Analysis of RNA motifs. *Curr. Opin. Struct. Biol.*, **13**, 300–308.

81. Lee,J.C., Cannone,J.J. and Gutell,R.R. (2003) The lonepair triloop: a new motif in RNA structure. *J. Mol. Biol.*, **325**, 65–83.

82. Lee,J.C., Gutell,R.R. and Russell,R. (2006) The UAA/GAN internal loop motif: a new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions. *J. Mol. Biol.*, **360**, 978–988.

83. Strogatz,S.H. (2001) Exploring complex networks. *Nature*, **410**, 268–276.

84. Barabasi,A.L. and Oltvai,Z.N. (2004) Network biology: understanding the cell's functional organization. *Nature Rev. Genet.*, **5**, 101–113.

85. Watts,D.J. and Strogatz,S.H. (1998) Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–442.

86. Barabasi,A.L. and Albert,R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.

87. Ravasz,E., Somera,A.L., Mongru,D.A., Oltvai,Z.N. and Barabasi,A.L. (2002) Hierarchical organization of modularity in metabolic networks. *Science*, **297**, 1551–1555.

88. Greene,L.H. and Higman,V.A. (2003) Uncovering network systems within protein structures. *J. Mol. Biol.*, **334**, 781–791.

89. Amitai,G., Shemesh,A., Sitbon,E., Shklar,M., Netanely,D., Venger,I. and Pietrokovski,S. (2004) Network analysis of protein structures identifies functional residues. *J. Mol. Biol.*, **344**, 1135–1146.

90. del Sol,A., Fujihashi,H., Amoros,D. and Nussinov,R. (2006) Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.*, **2**, 2006. 0019.

91. Klein,D.J., Moore,P.B. and Steitz,T.A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, **10**, 1366–1379.

92. Hansen,J.L., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2002) Structural insights into peptide bond formation. *Proc. Natl Acad. Sci. USA*, **99**, 11670–11675.

93. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.

94. Nissen,P., Hansen,J., Ban,N., Moore,P.B. and Steitz,T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920–930.

95. Mengel-Jorgensen,J., Jensen,S.S., Rasmussen,A., Poehlsgaard,J., Iversen,J.J. and Kirpekar,F. (2006) Modifications in *Thermus thermophilus* 23S ribosomal RNA are centered in regions of RNA-RNA contact. *J. Biol. Chem.*, **281**, 22108–22117.

96. Kowalak,J.A., Bruenger,E. and McCloskey,J.A. (1995) Posttranscriptional modification of the central loop of domain V in *Escherichia coli* 23S ribosomal RNA. *J. Biol. Chem.*, **270**, 17758–17764.

97. Brodersen,D.E., Clemons,W.M., Jr, Carter,A.P., Wimberly,B.T. and Ramakrishnan,V. (2002) Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16S RNA. *J. Mol. Biol.*, **316**, 725–768.

98. McCloskey,J.A. and Rozenski,J. (2005) The small subunit rRNA modification database. *Nucleic Acids Res.*, **33**, D135–D138.