Constraining ribosomal RNA conformational space

Paola Favaretto, Arjun Bhutkar and Temple F. Smith*

BioMolecular Engineering Research Center, Boston University, 36 Cummington Street, Boston MA 02215, USA

ABSTRACT

Despite the potential for many possible secondary-structure conformations, the native sequence of ribosomal RNA (rRNA) is able to find the correct and universally conserved core fold. This study reports a computational analysis investigating two mechanisms that appear to constrain rRNA secondary-structure conformational space: ribosomal proteins and rRNA sequence composition. The analysis was carried out by using rRNA–ribosomal protein interaction data for the *Escherichia coli* 16S rRNA and free energy minimization software for secondary-structure prediction. The results indicate that selection pressures on rRNA sequence composition and ribosomal protein–rRNA interaction play a key role in constraining the rRNA secondary structure to a single stable form.

INTRODUCTION

The ribosome is a large molecular complex that catalyzes protein synthesis in all living organisms. Its basic structure, which consists of RNAs and proteins, is extremely conserved across the three phylogenetic domains and even more conserved within each phylodomain. Its universality and conservation, along with its fundamental role in biological processes, make the ribosome one of the most interesting and challenging complexes to study.

Comparative analyses (1,2) have established consensus predicted secondary structures for rRNA from a variety of organisms among the Bacteria, Archaea and Eukarya, and experimental data from crystallographic studies have established the tertiary structures of both large and small prokaryotic ribosomal subunits (3–6). Given the length of the rRNA and the potential for nucleotides to form alternative base pairs, there are numerous equally stable secondary structural folds associated with each rRNA sequence that contain nearly equal numbers of paired bases. Remarkably, even though from a kinetic point of view the rRNA could be trapped in one of many conformations, the rRNA finds the correct (and universally conserved) core fold among all the possible alternatives.

Various factors might contribute to this: the primary sequence, the binding of ribosomal proteins (both as 'guides' and 'stabilizers'), the action of RNA chaperone proteins (7–10) and possibly the ion concentration in the cell (11).

Previous work suggested a time-dependent hierarchical order to rRNA folding, although there is some evidence suggesting that this might not always be the case (12–15). The binding of ribosomal proteins adds stability to the RNA secondary structures and restricts the space of stable structural conformations (16). There is some evidence that initiation of RNA tertiary contacts might even precede the formation of the complete secondary structure (15). Moreover, the addition of magnesium ions might lead to the formation of stable intermediates progressing to a stable tertiary conformation with increasing magnesium ion concentration (11).

This study reports the computational analysis of potential constraints on the formation of secondary structures that are alternate or competing to the native 16S rRNA structure. The results indicate that selection pressures on rRNA sequence and rRNA–ribosomal proteins interactions play a key role in constraining the secondary structure to a single core fold.

METHODS

The free energy minimization software for secondary-structure prediction *mfold* (17,18), along with ribosomal protein–rRNA contact data from the crystal structure of the 30S subunit (3) were employed to investigate the role of ribosomal proteins and the rRNA sequence in constraining the rRNA secondary structure to a single fold. The minimum energy secondary structure of the 16S rRNA sequence of *Escherichia coli* was predicted under various bound-protein constraints and with a range of base-pair substitutions. To comply with *mfold* limitations, bound-protein constraints forcing non-canonical base pairs (U:U, A:G, G:G, A:C, A:A) or associated with pseudoknots were not considered. A window size of 20, temperature parameter corresponding to 37°C, free energy increment of 5%, and other *mfold* parameters set to default values were used in the simulations. The secondary structures obtained were compared with the native 16S rRNA secondary structure of *E.coli*, as established by comparative analyses (1,2). The predicted folds were scored by the percentage of base pairs predicted as in the native structure as well

---

as by the percentage of helices (totally or partially) predicted as in the native structure.

In order to assess the role and the contribution of the ribosomal proteins in the achievement of the correct rRNA folding, topological constraints representing the physical constraints imposed by the binding of ribosomal proteins were implemented. A detailed map of the protein–rRNA interactions in the 30S subunit of the bacterial ribosome is described in Brodersen *et al*. (3). Data for ribosomal protein S21 was obtained from Fink *et al*. (19). In each case where a protein was reported to make contact with the RNA backbone and/or a base, the protein was assumed to impose a constraint on that particular residue. More specifically, if the residue was part of a base pair, the protein was assumed to force that base pair in the fold, whereas if the residue was part of a loop or bulge, the protein was assumed to restrain that residue from pairing. Simulations were run independently for constraints imposed by each ribosomal protein and for constraints imposed by ribosomal proteins belonging to the same binding pathway, as determined by *in vitro* reconstitution studies (20).

Random protein constraints were generated as controls. Different numbers of bases with potential for canonical pairing were forced to pair simultaneously to emulate the effect of binding of random control proteins, without necessarily involving or avoiding base pairs found in the native rRNA secondary structure. In addition, hypothetical protein constraints were created by forcing pairing between some bases that are known to pair in the native rRNA structure but are not constrained by ribosomal proteins.

In order to investigate the role of nucleotide sequence composition, the 16S rRNA sequence of *E.coli* was modified by applying a range of random base-pair substitutions in which the native canonical base pairs were replaced by alternate canonical base pairs. In this way, the sequence composition was changed but the potential to form native secondary structures was maintained. The randomization in the base-pair substitutions was achieved on two levels: (i) the base-pair position was chosen randomly; (ii) the type of canonical substitution was chosen randomly with uniform distribution. However, the base-pair substitutions were made so as to guarantee changes in the targeted bases. Totally random sequences were also generated, both with and without preserving the background frequency of the nucleotides in the native sequence.

Given that G:C-tetraloop-closing base pairs are considered key to helix stability (21), the following substitutions were implemented to evaluate their importance: (i) all G:C base pairs that close tetraloops were modified into other canonical base pairs with equal probability; (ii) all G:C base pairs that close tetraloops were modified into non-canonical base pairs; and (iii) base pairs that were not G:C and were not closing tetraloops were modified to the same extent as the G:C closing base pairs present in the native structure.

RESULTS

Protein constraints

In the absence of any constraints, *mfold* reported a total of 27 secondary structures within 5% of the minimum energy. Applying the constraints imposed by the binding of the ribosomal proteins led to a significant reduction in the number of
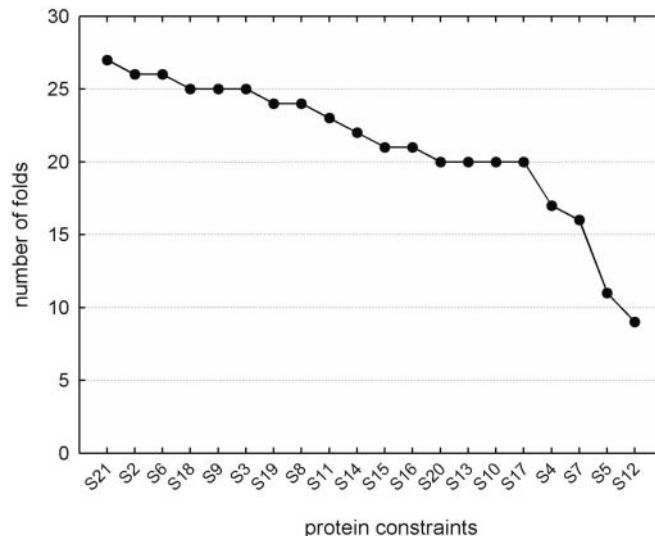

Figure 1. Number of folds within 5% of the minimum free energy fold predicted when individual protein constraints were applied. In each case where a protein was reported to make contact with an rRNA residue that is part of a base pair in the native structure, the protein was assumed to force that particular residue to pair. If a protein makes contact with an rRNA residue that is part of a loop or bulge in the native structure, the protein was assumed to prevent that residue from pairing.

possible folds (Figure 1), corroborating the hypothesis that ribosomal proteins play a critical role in enabling the attainment of the correct secondary structure by the rRNA (16). In particular, proteins S4, S7, S5 and S12 resulted in the most dramatic decrease in the number of alternative folds. Weak correlation ($r < 0.6$) was found between the number of constraints imposed by a given protein and the resulting number of folds associated with it. This rules out the possibility that the reduction in the variability of rRNA secondary-structure conformational space is solely due to the fact that some bases are constrained by the protein, either by being forced to pair or by being prevented from pairing. In other words, simply increasing the number of constraints does not necessarily result in a decrease in conformational variability.

When the constraints imposed by ribosomal proteins belonging to the same binding pathway (3,20,22) were applied (Figure 2), a trend towards decreased variability was observed (Table 1). By themselves, the early and intermediate binding proteins result in significant reduction in the number of folds and in increased accuracy. The late binding proteins by themselves do not have as significant an impact. When the early and intermediate proteins are combined, the accuracy is not much greater than when they are considered separately. When the late binding proteins are considered with the early and intermediate binding proteins, it is clear that their role is to incrementally increase the accuracy of the predicted structures, as well as to reduce the conformational variability, ultimately resulting in a single fold.

When the random and hypothetical protein constraints were applied to the 16S rRNA, no characteristic trend in the number of folds was observed. In several cases, although the constraints satisfied *mfold* requirements and involved segments with clear potential for helical structure, no stable fold could be found. In some cases, the number of obtained folds was
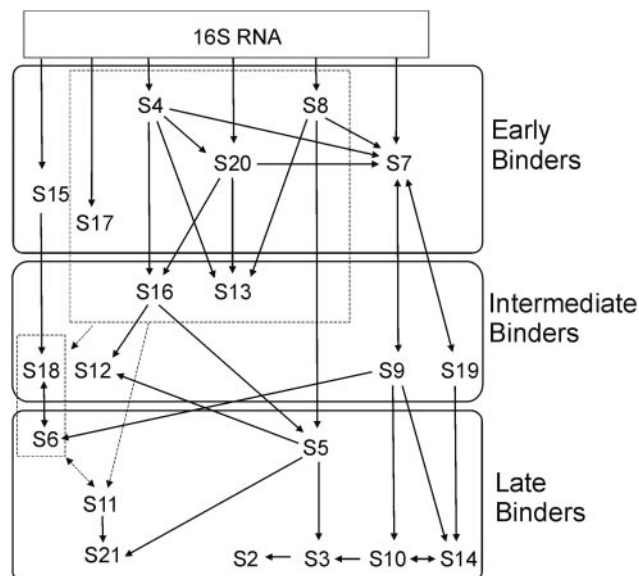
Figure 2. *E.coli* 16S ribosomal protein binding pathways determined from earlier *in vitro* studies (3,20). Arrows indicate ordered binding. Proteins are grouped together in terms of their temporal binding sequence as early, intermediate and late binders.

lower and in other cases higher than the number of folds obtained with comparable number of constraints from real ribosomal proteins. In general, we did not observe any specific trend in the data that would suggest that simply forcing bases to pair necessarily results in reduction of the secondary-structure conformational space of rRNA.

Sequences with base-pair substitutions

The analysis of the sequences with base-pair substitutions indicates that rRNA sequence composition has a strong influence in reducing the number of possible alternate secondary structures. In fact, even though the base-pair substitutions allowed the preservation of the native secondary structure, a statistically significant decrease ($p < 0.05$) in the number of base pairs and helices correctly predicted was observed, as the percentage of such canonical substitutions increased (Figure 3). This suggests that alternate structures can form when the primary sequence content is changed, even if the native secondary-structure elements can be preserved.

G:C-tetraloop-closing base pairs

Substitutions of G:C-tetraloop-closing base pairs by other canonical base pairs resulted in a decrease in the number of correctly predicted base pairs (Figure 4). Furthermore, as expected, substitutions of G:C-tetraloop-closing base pairs by non-canonical base pairs produced additional degradation—in some cases, no native base-pair matches were found. This decrease in the prediction accuracy cannot be attributed solely to changes in the native sequence, given that sequences with same number of substitutions to other canonical base pairs at random positions (rand 3%) or at positions corresponding to non-closing-G:C pairs (non closing GC) have a negligible effect on the number of correct predicted base pairs and helices (Figure 4).

DISCUSSION

Protein constraints

Among all the ribosomal proteins of the small subunit, S4, S7, S5 and S12 appear to have the greatest potential impact on reducing the number of possible folds. S4, S12 and S5 bind to the functional center and constrain critical base-pairing in the central region (3), thereby partition the 16S rRNA into its primary domains (central, 3′ major, 3′ minor, 5′). S5 and S7 bind to the head region (3′ major domain) of the rRNA, which presents significant conformational variability.

From the analysis of native, hypothetical and random protein constraints in terms of their ability to restrict the conformational space of the 16S rRNA, we can conclude that even though the ribosomal protein constraints might not represent the most efficient combination of constraints, they certainly are optimal and contribute to limiting the number of possible alternate secondary structures associated with the rRNA. The hypothetical and random constraints do not comply with three-dimensional limitations, fold-stability requirements, nor do they conform to other important aspects of protein functions (e.g. protein–protein interactions, binding specificity or extra-ribosomal activity). Consequently, it is entirely possible that some of these random and/or hypothetical constraints turned out, by chance, to be more effective than the ribosomal proteins in reducing the number of rRNA folds. On the other hand, ribosomal proteins are part of a very complex molecular machinery and are likely to have evolved to the present state by responding to several different selective pressures.

As part of their respective binding pathways, the bound ribosomal proteins dramatically reduce the rRNA secondary-structure conformational space (from 27 folds to 1 and from 48.44% correct base pairs to 81.13%). As seen in Table 1, early and intermediate binding proteins have the greatest impact in terms of reducing the number of folds (from 27 folds to 3 and 2), corresponding to a >89% decrease. On the other hand, the late binding proteins appear to improve the quality of the predicted secondary structure incrementally (from 77.25 to 81.13% correctly predicted base pairs).

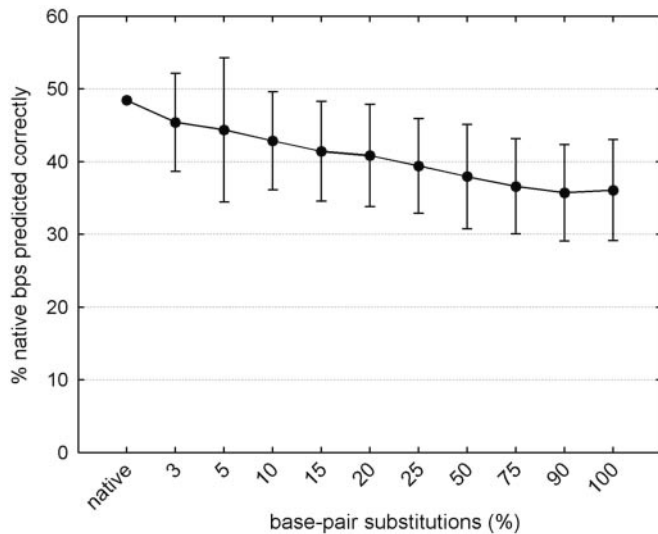### Sequences with base-pair substitutions

Regardless of the percentage of canonical base-pair substitutions, the number of base pairs predicted remains approximately the same (Figure 5). This is not very surprising given that the potential for forming all native secondary-structure elements remains, being that the substitutions are canonical. In the completely random sequences, the potential for helical structure is not preserved, but the total number of predicted base pairs is again not very different.

If the base pairs in the native sequence are substituted by other canonical base pairs, the number of stable alternate secondary structures remains approximately the same (Figure 6); therefore, the average number of stable folds does not depend on the percentage of base-pair substitutions. However, if bases not involved in native base-pairing are randomly substituted by other bases, the number of potential folds drastically increases (up to 54 folds) even though the average number of predicted total base pairs is approximately the same (Figure 5) and the nucleotide composition is the same.
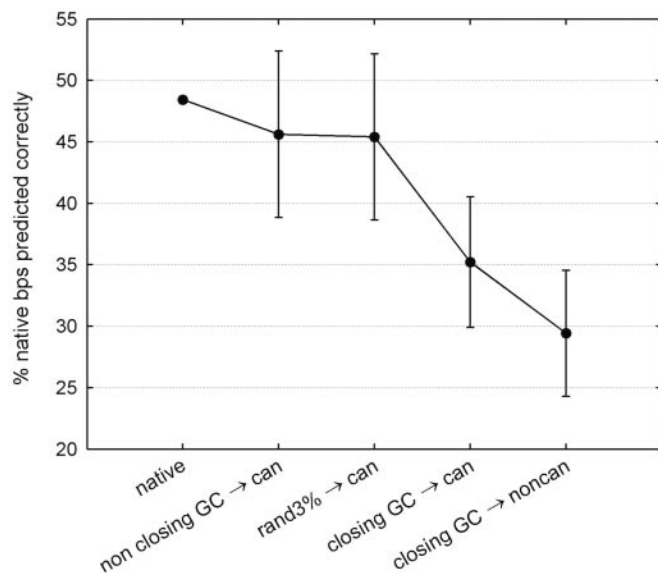
Table 1. Examples of protein binding pathways and their impact on the conformational variability and accuracy of predicted folds

Protein binding pathway	Num Folds	% Correct bps Avg	% Correct bps Max	% Correct bps Min	% Correct Full Helices Avg	% Correct All Helices Avg	% Correct All Helices Max	% Correct All Helices Min	Min Energy (kcal/mol)
Pathway S8_S5_S3_S2:									
S8	24	49.35	67.30	41.72	30.60	46.51	64.52	38.71	−604.8
S8_S5	12	64.66	70.23	60.17	37.46	62.63	67.74	58.06	−593.4
S8_S5_S3	11	63.26	68.76	58.70	36.56	61.78	66.67	56.99	−579.9
S8_S5_S3_S2	11	63.46	68.97	58.91	37.63	61.78	66.67	56.99	−579.1
Pathway S8_S4_S20_S7_S9_S19_S14:									
S8	24	49.35	67.30	41.72	30.60	46.51	64.52	38.71	−604.8
S8_S4	9	60.17	70.23	49.69	37.04	56.51	66.67	47.31	−598.7
S8_S4_S20	8	66.59	75.05	60.17	41.67	63.17	72.04	56.99	−590.8
S8_S4_S20_S7	4	70.02	76.52	62.05	41.94	66.40	73.12	58.06	−579.9
S8_S4_S20_S7_S9	3	71.42	75.26	67.51	40.86	67.03	70.97	62.37	−573.9
S8_S4_S20_S7_S9_S19	3	71.42	76.52	68.76	43.01	67.03	70.97	62.37	−573.2
S8_S4_S20_S7_S9_S19_S14	2	75.05	75.26	71.49	42.47	72.58	74.19	70.97	−564.9
Pathway S8_S4_S20_S17_S16_S5_S12:									
S8	24	49.35	67.30	41.72	30.60	46.51	64.52	38.71	−604.8
S8_S4	9	60.17	70.23	49.69	37.04	56.51	66.67	47.31	−598.7
S8_S4_S20	8	66.59	75.05	63.17	41.67	63.17	72.04	56.99	−590.8
S8_S4_S20_S17	7	67.09	75.05	60.17	41.78	63.75	72.04	56.99	−590.8
S8_S4_S20_S17_S16	7	66.76	75.05	58.28	41.94	63.59	72.04	55.91	−590.8
S8_S4_S20_S17_S16_S5	2	76.42	76.94	75.89	45.70	74.73	75.27	74.19	−580.7
S8_S4_S20_S17_S16_S5_S12	1	80.71	80.71	80.71	46.24	78.49	78.49	78.49	−567.7
Pathway S8_S4_S20_S17_S7_S15_S16_S13_S18_S9_S6:									
S8	24	49.35	67.30	41.72	30.60	46.51	64.52	38.71	−604.8
S8_S4	9	60.17	70.23	49.69	37.04	56.51	66.67	47.31	−598.7
S8_S4_S20	8	66.59	75.05	63.17	41.67	63.17	72.04	56.99	−590.8
S8_S4_S20_S17	7	67.09	75.05	60.17	41.78	63.75	72.04	56.99	−590.8
S8_S4_S20_S17_S7_S15	3	72.26	77.15	66.46	42.29	68.46	74.19	62.37	−571.6
S8_S4_S20_S17_S7_S15_S16	3	72.26	77.15	66.46	42.29	68.46	74.19	62.37	−571.6
S8_S4_S20_S17_S7_S15_S16_S13	2	75.58	77.57	73.58	43.55	71.51	74.19	68.82	−567.8
S8_S4_S20_S17_S7_S15_S16_S13_S18	2	76.42	78.41	74.42	43.55	70.43	73.12	67.74	−563.4
S8_S4_S20_S17_S7_S15_S16_S13_S18_S9	2	75.16	77.15	73.17	41.40	68.28	70.97	65.59	−557.4
S8_S4_S20_S17_S7_S15_S16_S13_S18_S9_S6	2	75.16	77.15	73.17	41.40	68.28	70.97	65.59	−557.4
Pathway S15_S18_S6_S11_S21:									
S15	21	49.95	71.49	40.04	30.97	48.01	68.82	38.71	−605
S15_S18	20	52.03	72.33	40.88	31.40	48.12	67.74	37.63	−600.6
S15_S18_S6_S11_S21	20	52.45	72.75	41.30	32.47	49.19	68.82	38.71	−599.8
Pathway S17_S12_S5_S3_S2_S21:									
S17	20	51.91	68.97	44.23	32.90	49.46	65.59	43.01	−611.9
S17_S12_S5	1	74.63	74.63	74.63	41.94	72.04	72.04	72.04	−583.8
S17_S12_S5_S3_S2_S21	1	75.05	75.05	75.05	41.94	72.04	72.04	72.04	−565.8
Pathway S7_S19_S9_S14:									
S7	16	59.64	70.44	50.73	35.28	55.65	66.67	49.46	−599.8
S7_S19	15	59.99	70.44	50.73	35.27	56.13	66.67	49.46	−599.1
S7_S19_S9	15	58.73	69.18	49.48	33.12	53.98	64.52	47.31	−593.1
S7_S19_S9_S14	14	60.32	70.86	51.15	33.95	57.30	67.74	50.54	−584.8
Early, intermediate, and late binders:									
No proteins	27	48.44	70.86	37.74	30.67	45.32	67.74	35.48	−615.2
Early binders only	3	72.26	77.15	66.46	42.29	68.46	74.19	62.37	−571.6
Intermediate binders only	2	77.88	79.87	75.89	42.47	69.35	72.04	66.67	−575.5
Late binders only	7	65.80	71.91	62.05	38.10	63.75	69.89	60.22	−582.6
Early + intermediate binders	2	77.25	79.25	75.26	42.47	69.35	72.04	66.67	−547.1
Early + intermediate + late binders	1	81.13	81.13	81.13	44.09	76.34	76.34	76.34	−536.2

The bottom part of the table shows the impact of early-, intermediate- and late-binding proteins considered as individual categories and also considered together. 'Num_folds' reports the number of folds produced by *mfold* after the corresponding pathway protein constraints have been applied. '% Correct bps Avg' reports the average base-pair prediction accuracy across all folds compared to base pairs in the native structure. '% Correct Full Helices Avg.' reports the average number of native helices predicted fully (all helix base pairs) across all the corresponding folds. '% Correct All Helices Avg.' reports the average number of native helices predicted fully or partially (sum of the two) across all the corresponding folds. A helix is considered to have been predicted partially if some, but not all, of the base pairs in the native helix structure have been predicted. Maximum and minimum values for base pair and helix prediction accuracy are also shown in the corresponding columns. The minimum folding energy across all folds for each set of protein constraints is reported in the last column. Samples of protein constraints corresponding to binding pathways from Figure 2 are shown. Some combinations of constraints appear to reduce the rRNA secondary-structure conformational variability more than others. In some cases, the final conformation is restricted to a single fold, as in the case of all proteins (early-, intermediate-, late-binders combined). See text for further discussion.
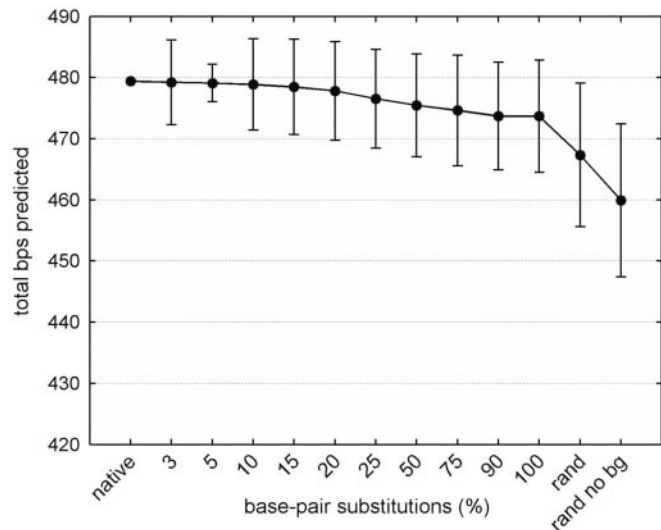
Figure 3. Average percentage of native base pairs predicted correctly for sequences with canonical base-pair substitutions. For each class of base-pair substitutions, the predicted base pairs that are present in the native structure are considered correct. The average is taken across the 100 samples of each class. A one-factor ANOVA test showed that the difference in the average percentage of base pairs correctly predicted is statistically significant ($p < 0.05$).

Figure 5. Average percentage of native base pairs predicted correctly for sequences with canonical base-pair substitutions. The native sequence is the 16S rRNA sequence of *E.coli*; 'rand' indicates random sequences of the same length and base distribution as the native sequence; 'rand no bg' indicates random sequences of the same length as the native sequence with a uniform base distribution.
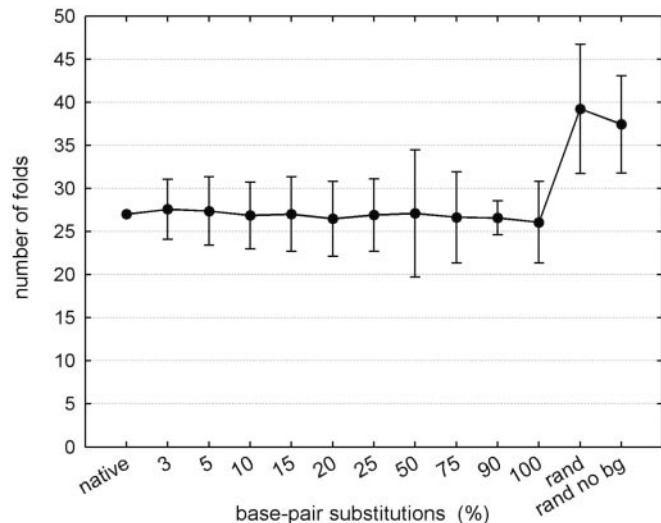

Figure 4. Average percentages of native base pairs predicted correctly for the following classes of sequences: 'native' sequence—16S rRNA sequence of *E.coli*; 'non closing GC → can'—*E.coli* 16S rRNA sequences in which 3% of base pairs that are not G:C-tetraloop-closing base pairs are substituted by other canonical base pairs; 'rand 3% → can'—*E.coli* 16S rRNA sequences in which ∼3% of the native base pairs are substituted by other canonical base pairs; 'closing GC → can'—*E.coli* 16S rRNA sequences in which all the G:C-tetraloop-closing base pairs are substituted by other canonical base pairs (the total number of substitutions is equivalent to ∼3% of the total number of native base pairs); 'closing GC → noncan'—*E.coli* 16S rRNA sequences in which all the G:C-tetraloop-closing base pairs are substituted by non-canonical base pairs (the total number of substitutions is equivalent to ∼3% of the native base pairs).

This clearly suggests that the positional composition of the sequence is under selection, not just for biochemical function and protein binding, but also for not competing in alternative base-pairing.

Figure 6. Average number of secondary-structure folds predicted within 5% of the minimum free energy for sequences with canonical base-pair substitutions. Sequence categories are the same as in Figure 5.

G:C-tetraloop-closing base pairs

Some positions of the 16S rRNA sequence seem more significant than others in limiting the number of possible folds, e.g. the G:C base pairs that close tetraloops. The degradation in prediction quality observed when these were replaced by other canonical and non-canonical base pairs is energetically expected. This is in line with experimental observations that one of the earliest structural motifs to form *in vitro* are tetraloops with highly stable G:C closing pairs (21).

Domain variability

When folded independently, without protein constraints, the 3′ major domain exhibits the highest number of alternate secondary structures, as compared with the other domains. This variability, and the fact that a considerable number of ribosomal proteins bind into the 3′ major domain, is further evidence of the structural role played by the ribosomal proteins. The rRNA regions that present the most variability necessitate extensive ribosomal protein–rRNA interactions to attain the correct secondary structure and consequently the correct fold.

In conclusion, while we are aware of the potential limitations of considering only secondary-structure energetics, our results suggest that both rRNA primary sequence and constraints imposed by ribosomal proteins are critical to the attainment of the correct fold of the rRNA.

REFERENCES

1. Cannone,J.J., Subramanian,S., Schnare,M.N., Collett,J.R., D'Souza,L.M., Du,Y., Feng,B., Lin,N., Madabusi,L.V., Müller,K.M., Pande,N., Shang,Z., Yu,N. and Gutell,R.R. (2002) The Comparative RNA Web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *Bioinformatics*, **3**, 2.
2. Gutell,R.R., Weiser,B., Woese,C.R. and Noller,H.F. (1985) Comparative anatomy of the 16-S-like ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **32**, 155–216.
3. Brodersen,D.E., Clemons,W.E.,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.
4. Klein,D., Moore,P. and Steitz,T. (2004) The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J. Mol. Biol*, **340**, 141–177.
5. 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.
6. 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 A resolution. *Science*, **289**, 905–920.
7. Cristofari,G. and Darlix,J. (2002) The ubiquitous nature of RNA chaperone proteins. *Prog. Nucleic Acid Res. Mol. Biol.*, **72**, 223–268.
8. Herschlag,D. (1995) RNA chaperones and the RNA folding problem. *J. Biol. Chem.*, **270**, 20871–20874.
9. Lorsch,J. (2002) RNA chaperones exist and DEAD box proteins get a life. *Cell*, **109**, 797–800.
10. Maki,J., Schnobirch,D. and Culver,G. (2002) The DnaK chaperone system facilitates 30S ribosomal subunit assembly. *Mol. Cell*, **10**, 129–138.
11. Rangan,P., Masquida,B., Westhof,E. and Woodson,S. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA*, **100**, 1574–1579.
12. Brion,P. and Westhof,E. (1997) Hierarchy and dynamics of RNA folding. *Ann. Rev. Biophys. Biomol. Struct.*, **26**, 113–137.
13. Tinoco,I. and Bustamante,C. (1999) How RNA folds. *J. Mol. Biol.*, **293**, 271–281.
14. Westhof,E. and Massire,C. (2004) Evolution of RNA architecture. *Science*, **306**, 62–63..
15. Wu,M. and Tinoco,I. (1998) RNA folding causes secondary structure rearrangement. *Proc. Natl Acad. Sci. USA*, **95**, 11555–11560.
16. Noller,H. (2004) The driving force for molecular evolution of translation. *RNA*, **10**, 1833–1837..
17. Zuker,M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.
18. Mathews,D., Sabina,J., Zuker,M. and Turner,D. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structures. *J. Mol. Biol.*, **288**, 911–940.
19. Fink,D., Chen,R., Noller,H. and Altman,R. (1996) Computational methods for defining the allowed conformational space of 16S rRNA based on chemical footprinting data. *RNA*, **2**, 851–866.
20. Held,W.A., Mizushima,S. and Nomura,M. (1973) Reconstitution of *Escherichia coli* 30S ribosomal subunit from purified molecular components. *J. Biol. Chem.*, **248**, 5720–5730.
21. Woese,C., Winker,S. and Gutell,R. (1990) Architecture of ribosomal RNA: constraints on the sequence of tetra-loops. *Proc. Natl Acad. Sci. USA*, **87**, 8467–8471.
22. Nomura,M. (1973) Assembly of bacterial ribosomes. *Science*, **179**, 864–873.