

# Nanometer scale pores similar in size to the entrance of the ribosomal exit cavity are a common feature of large RNAs

MARIO RIVAS,<sup>1,3</sup> QUYEN TRAN,<sup>2,3</sup> and GEORGE E. FOX<sup>2,4</sup>

<sup>1</sup>Laboratorio de Origen de la Vida, Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México, Coyoacán, C.P. 0451, Mexico

<sup>2</sup>Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, USA

## ABSTRACT

The highly conserved peptidyl transferase center (PTC) of the ribosome contains an RNA pore that serves as the entrance to the exit tunnel. Analysis of available ribosome crystal structures has revealed the presence of multiple additional well-defined pores of comparable size in the ribosomal (rRNA) RNAs. These typically have dimensions of 1–2 nm, with a total area of  $\sim 100 \text{ \AA}^2$  or more, and most are associated with one or more ribosomal proteins. The PTC example and the other rRNA pores result from the packing of helices. However, in the non-PTC cases the nitrogenous bases do not protrude into the pore, thereby limiting the potential for hydrogen bonding within the pore. Instead, it is the RNA backbone that largely defines the pore likely resulting in a negatively charged environment. In many but not all cases, ribosomal proteins are associated with the pores to a greater or lesser extent. With the exception of the PTC case, the large subunit pores are not found in what are thought to be the evolutionarily oldest regions of the 23S rRNA. The unusual nature of the PTC pore may reflect a history of being created by hybridization between two or more RNAs early in evolution rather than simple folding of a single RNA. An initial survey of nonribosomal RNA crystal structures revealed additional pores, thereby showing that they are likely a general feature of RNA tertiary structure.

**Keywords:** RNA tertiary structure; ribosome evolution; RNA nanopores; exit tunnel; peptidyl transferase center

## INTRODUCTION

The ribosome is responsible for coded protein synthesis (Palade et al. 1955; Watson 1963; Noller et al. 1992; Schmeing and Ramakrishnan 2009). The core transpeptidation reaction occurs in a region of the large ribosomal RNA, (rRNA), known as the peptidyl transferase center (PTC), (Nierhaus et al. 1980; Nissen et al. 2000; Sievers et al. 2004; Schmeing et al. 2005; Trobro and Aqvist 2005; Rodnina et al. 2006; Voorhees et al. 2009; Leung et al. 2011; Zaher et al. 2011). From an origins perspective it has been proposed that the PTC originated as a result of noncovalent interactions between two L-shaped RNA molecules that were ultimately ligated together (Agmon et al. 2005; Davidovich et al. 2009). One of these hypothetical RNAs would represent the core of the A-site and the other the P-site in the modern 50S subunit. When these two regions of the modern RNA interact, they create a well-defined cavity that accommodates the termini

of the A- and P-site tRNAs (Voorhees et al. 2009) and therefore forms at least part of the catalytic site.

This cavity (Fig. 1A,B) is properly regarded as a 1.5–2 nm pore. It serves as the entrance to the exit tunnel (Voss et al. 2006), which in turn is effectively a nanotube. In modern protein synthesis, the tunnel allows the nascent protein to leave the ribosome and prevents premature (mis)folding (Nakatogawa and Ito 2002). The tunnel is targeted by macrolide antibiotics that bind RNA functional groups lining its walls, with the result that peptide elongation is inhibited (Starosta et al. 2010). Two key issues that have not been explored previously are the frequency of occurrence of RNA pores and the general nature of the RNA structures that form them.

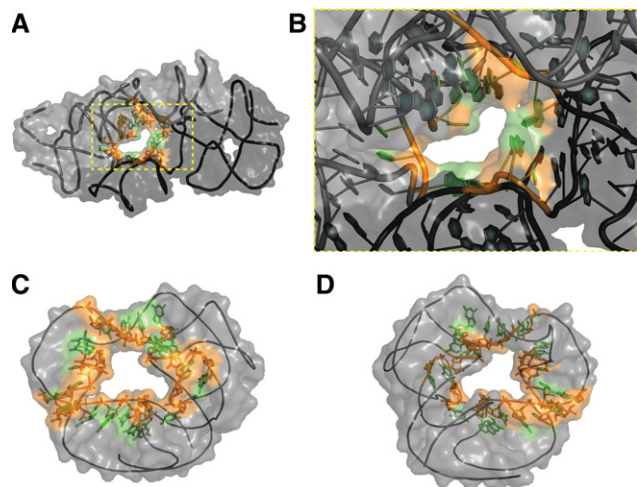
Voss et al. (2006) previously looked for cavities in the ribosome as part of their effort to characterize the exit tunnel. To accomplish this, they used a rolling ball algorithm to detect cavities in the complete ribosome structure. Because they included both proteins and RNA, their approach would necessarily overlook a pore that was formed by the RNA alone, but occluded by one or more proteins. In their effort, the largest number of cavities was found when the probe radius was in the 3–5 Å size range. A second peak occurred when the rolling ball radius was in the 8–9.5 Å size range, but the specific responsible structural features were not described.

<sup>3</sup>These authors contributed equally to this work.

<sup>4</sup>Corresponding author

E-mail fox@uh.edu

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.038828.113>.



**FIGURE 1.** Although it is of similar size, the PTC Pore is distinctly different from other pores. (A) The pocket-like symmetrical region of the ribosome (Agmon et al. 2005) includes the PTC pore. Atoms that line the pore are colored in green (nitrogenous bases) and orange (phosphate backbone and sugars). (B) Close-up view of the PTC pore reveals a unique ordering of the RNA residues whereby the phosphate backbone is twisted away from the lumen of the pore, while nitrogenous bases are exposed or likely to be in contact with any molecule that occupied the space. (C) The region covering residues 2103–2186 in *Thermus thermophilus* 23S rRNA (crystal 2WDL) is shown. This region encompasses pore 10 (Table 1). In contrast with the PTC pore, the nitrogenous bases (in green) are twisted away, leaving the phosphate backbone (in orange) to line the lumen of the pore. This geometry is the same for all of the non-PTC pores found in the large and small ribosomal subunit. (D) View of pore 10 from the opposite side.

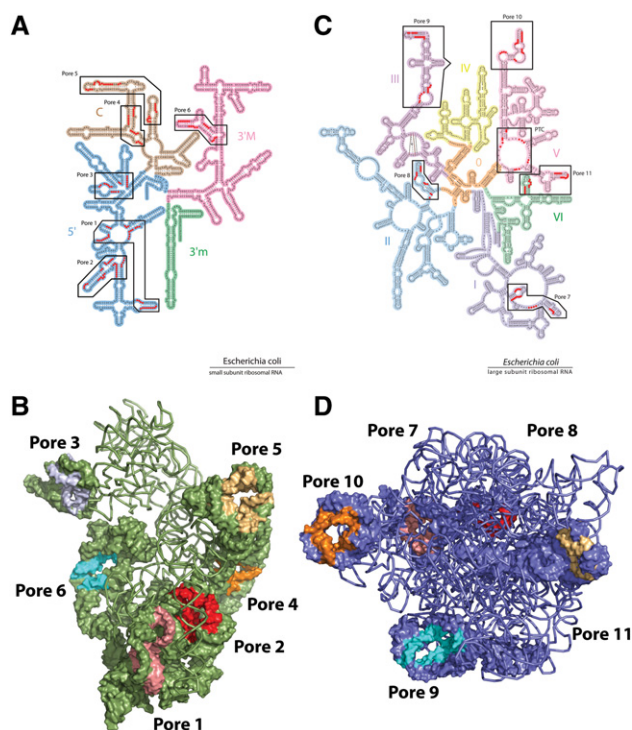
## RESULTS AND DISCUSSION

An initial examination of the *Thermus thermophilus* 16S and 23S rRNAs revealed the presence of 11 nanometer-size pores in addition to the PTC pore as listed in Table 1. Their locations in the context of the rRNA secondary structure are shown in Figure 2, A and C. With regard to the tertiary structure of the ribosome, they are mostly on the surface and avoid the interface between the two subunits. The location of the pores relative to the ribosome tertiary structure is shown in Figure 2, B and D.

It might be argued that these pores were unique features of the rRNAs and therefore would not occur in other RNAs. To address this, the crystal structures of several other RNAs were examined. No nanometer-size pores were found in a partial spliceosome structure or 4.5S RNA. However, additional pores were found in several riboswitch and RNase P structures (Table 2), bringing the total found to date to 16. For each pore, a continuous subsequence of ~54–293 residues encompasses the pore. The pores were never exactly round, so we characterized their size in terms of the smallest and largest distance across the opening from which one can estimate an “area.” Typically, these distances range from 10 to 20 Å (1–2 nm), though several of the pores seen in various riboswitch structures are somewhat smaller (7–12 Å). False pores are some-

times seen when one looks down a helical stem. The true pores are actually built, in part, on such stems, typically utilizing at least four or more residues on at least one strand of a helical element. Although it is somewhat subjective, we also sought to identify specific residues that define the boundaries of each pore. This varies between 15 and 26 residues in regions containing only one pore. A detailed summary for each pore is provided in Table 1.

Comparison of the various RNA pores suggests that there are at least two architectures. The far more common type is a phosphate-lined pore defined by the phosphodiester backbone of the RNA, as illustrated by pore 10 (Fig. 1C). Such pores will present a negative charged surface to the environment. The second type is exemplified by the single case of the PTC pore (Fig. 1B), where the pore is lined in part by the nucleic acid bases rather than just the RNA backbone. Pores of this type will present a slightly hydrophobic or positively charged



**FIGURE 2.** Location of pores relative to ribosome structure. (A) The individual RNA residues defining the pores as seen in Table 1 are highlighted in red on a secondary structure diagram obtained from the Ribovision site (<http://apollo.chemistry.gatech.edu/RiboVision>). The approximate region of RNA that bounds the pore is demarcated by a black box. (B) The distribution of pores in the 30S subunit is viewed from the 50S side. Each pore is labeled. The portion of the rRNA that bounds each pore is shown as a green surface and the residues that define the pore are in various colors. All of the remaining portions of the rRNA are shown as wire colored in green. The central medial axis of the 30S subunit is devoid of nanometer scale pores. (C) Pores in the large subunit are shown as in A. The underlying secondary structure diagram (Petrov et al. 2013) was obtained from the Ribovision website. (D) Crown view of the distribution of pores found in the 50S subunit. The pores are distributed along two planes perpendicular to the vertical axis passing through the central protuberance.

**TABLE 1.** Pores found in the *T. thermophilus* ribosome crystals PBD ID 2WDK and 2WDL

Pore number	RNA	Residues defining the pore	Residues count	Size (Å)	Protein interaction
1	16S	60–62, 101–109, 150–153, 161–169, 331–338, 348–350	36	15 to 21	S20
2	16S	119–123, 233–241, 248–254, 264–267, 276–277, 282–287	33	13 to 21	S17
3	16S	405–408, 427–428, 430–434, 498–499, 541–546	19	15 to 19	S4
4	16S	578–581, 655–659, 742–746, 754–758	19	10 to 15	S15
5	16S	689–692, 699–700, 702–710, 713–714, 775–780, 796–800	28	9 to 13	S11
6	16S	990–993, 1004–1008, 1016–1021, 1038–1041	19	11 to 18	NONE
7	23S	203–207, 217–220, 235–242, 253–257	21	16 to 18	L15, L28, L34, L35*
8	23S	1197–1199, 1213–1217, 1228–1231, 1239–1241	15	10 to 16	L20, L4, L21*
9	23S	1424–1429, 1480–1487, 1498–1503, 1559–1564	26	7 to 14	L2
10	23S	2115–2118, 2134–2138, 2148–2154, 2159–2162, 2173–2174	22	10 to 15	NONE
11	23S	2522–2526, 2531–2537, 2647–2651, 2664–2669	23	9 to 15	L6, L36*
PTC	23S	2061–2064, 2439, 2441–2442, 2450–2452, 2505–2506, 2585–2587	15	10 to 17	NONE

Instances where multiple proteins interact with the pore space are marked with an asterisk.

surface, depending on the hydration and protonation state of the bases in the solvent. In two cases, pores 1 and 2 (Table 1; Supplemental Fig. S1) and pores 4 and 5 (Table 1), a somewhat larger region of the rRNA was found to encompass two pores. As shown for the pair consisting of pores 4 and 5 (Fig. 3A), both pores are again clearly lined with backbone atoms rather than nucleotides. Single pores 3 and 7 are shown in Supplemental Figure S2.

The presence of the non-PTC rRNA pores appears to be a conserved feature of many rRNAs. An examination of the crystal structure of the *E. coli*, *Haloarcula marismortui*, and *Deinococcus radiodurans* 50S subunit structures revealed that all five non-PTC pores found in the 23S rRNA (Table 1, pores 7–11) are seen in all three structures. Although the available structures for the 30S subunit do not include a member of the Archaea, equivalent pores were again found at similar locations in *E. coli* and *Thermus thermophilus*. However, many of the pores will clearly be missing in the most minimal mitochondrial rRNAs, where large regions of the rRNAs are deleted.

Given that the RNA secondary structures themselves are very conserved, the occurrence of equivalent pores is not unexpected. To explore the matter further, we examined the conservation of the individual residues lining each pore using preprocessed data available on the Comparative RNA web site and project (Cannone et al. 2002). The results in Supplemental Table S1 revealed that residues lining pores 1, 2, 7, and especially 5 and 11, were very conserved. The conservation, however, did not match the level of the PTC pore, where only two residues showed significant variation in all of the Bacteria Domain. In contrast, the residues that line pores 6, 8, and 9 were far less conserved, even when only the Bacterial Domain was considered.

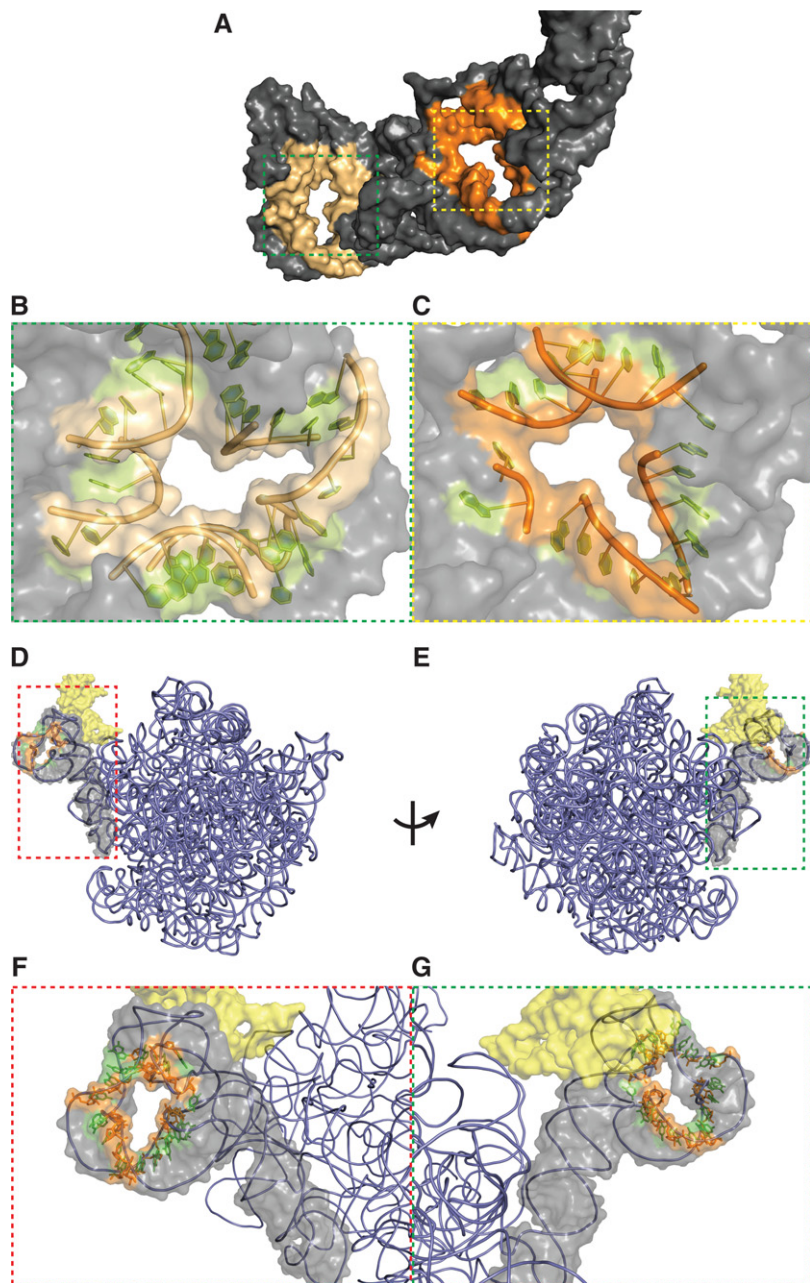
The ribosome crystal structures are very porous, and water residues reach the interior in many places (Voss et al. 2006). It is thus unlikely that the nanometer scale pores described here

serve as channels for water flow. A more likely utility may be to stabilize the binding of various ribosomal proteins. With the possible exception of pores 6 and 10, all of the rRNA pores are associated to some extent with ribosomal proteins. An example of extensive association occurs in pore 3, where ribosomal protein S4 actually protrudes into and completely fills the pore, Supplemental Figure S3.

In order to facilitate future studies of the properties of RNA pores, it will be useful to have an effective model system. All of the rRNA pores seen to date are in the larger context of the structure of an entire rRNA. Thus, it is possible that they are stabilized directly or indirectly in part by tertiary interactions with other parts of the rRNA or ribosomal proteins and, hence, might cease to exist in the absence of such interactions. This is unlikely for at least pores 1 and 2, as it has been shown that the 5' domain of *E. coli* 16S rRNA already forms all of the expected tertiary interactions in the absence of the ribosomal proteins (Adilakshmi et al. 2005). Nevertheless, pore 10 is an interesting exception. It is located in a region of the larger ribosome structure near the ribosomal protein L1 stalk that is devoid of interactions with other parts of the RNA (Fig. 3D, E). Ribosomal protein L1 circumscribes the perimeter of the pore on the side furthest from the 30S subunit, but does not protrude into the pore as seen in the *T. thermophilus* 50S crystal structure (Fig. 3F,G). Its proximity to the pore and the E-site tRNA make it possible that the combination may play a role in the dissociation of the tRNA from the ribosome. Absent this, pore 10 may serve as a model system. It is an apparently self-contained pore that is encompassed by positions 2103–2186, and thus can be synthesized as an 85-residue sequence.

Nanometer-scale RNA pores can be regarded as a novel form of RNA tertiary structure. From the perspective of RNA design, future work may focus on efforts to predict their occurrence from primary sequence alone and as a possible construct for use in synthetic biology applications. For





**FIGURE 3.** Region of *T. thermophilus* 16S rRNA covering residues 575–810 that contains two adjacent pores and a region from the 23S rRNA showing a self-contained pore. (A) Overall surface rendering of the region with pore 4 in orange and pore 5 in yellow. (B) Close-up view of pore 4 with backbone colored in orange and nitrogenous bases in green. (C) Close-up view of pore 5 with backbone now colored in yellow. As in Figure 1, C and D, the nitrogenous bases in both cases avoid the lumen of the pore. (D) The ribosomal region that contains pore 10 that was previously highlighted in Figure 1, C and D, is shown. The 23S rRNA is displayed as a blue ribbon. Ribosomal protein L1 is rendered as a light-yellow surface, and the region containing pore 10 is highlighted by the red dotted lines. (E) Same as D but rotated 180 degrees. (F,G) Close-up view for both orientations, respectively.

example, it may be possible to create pores that facilitate catalysis. The PTC pore clearly binds RNA ligands, e.g., the termini of the A and P site tRNAs (Kim and Green 1999; Voorhees et al. 2009; Zaher et al. 2011). Most of the other

pores likely interact with peptide ligands as well, as is illustrated for ribosomal protein S4, which interacts with pore 3 (Supplemental Fig. S3).

Given that pore 10 is much more self-contained than the other pores, it could be used as a scaffold to design and test the feasibility of practical applications that utilize the aforementioned interactions. Perhaps a naïve but intuitive starting point would be to duplicate the architecture of pore 10 multiple times in a larger RNA. The primary sequence would have to be carefully designed so that, in solution, such an RNA would form many pores arrayed in a net-like macrostructure. This net-like structure would have a regular distribution of both large and small pores through which filtrates/small biomolecules could flow. The stacking and tessellation of such a structure will conceivably produce a molecular sieve that could have superior properties to those of ceramic and membrane variants. Furthermore, RNA is less toxic and its stability could be enhanced by modifications such as the addition of a 2' O-methylation on one or more residues.

Although RNA pores appear to be a common feature in many RNAs, the PTC example remains unusual in that it is not a phosphate-lined pore. It is, in fact, the only pore that is found in what are regarded as the most ancient areas of the 23S rRNA. The other 23S rRNA pores are all in regions of the RNA that are relatively recent additions to the structure (Hury et al. 2006; Bokov and Steinberg 2009; Hsiao et al. 2009; Fox 2010). The unique nature of the PTC pore may reflect a history of being created by hybridization between two or more RNAs in an RNA world rather than simple folding of a single RNA (Davidovich et al. 2009; Fox et al. 2012).

## MATERIALS AND METHODS

A heuristic search for RNA pores in the crystal structures of various large RNAs was undertaken. Crystal structures were obtained from the Protein Data Bank (PDB, <http://www.pdb.org>). Initially, the crystal structures of ribosomes from *T. thermophilus* (PDB 2WDK and 2WDL) were used to identify pores. As previously shown by Voss et al. (2006), there are numerous small cavities in the large

**TABLE 2.** Pores found in nonribosomal RNA crystals

RNA's type	PDB ID	Components	Residues defining the pore	# Residues	Size (Å)
Riboswitch	2YDH	Chain A	13, 22–23, 37–40, 60–62	10	7 to 17
	2YIF	Chain X	10–13	13	16
		Chain Z	95–103		
RNase-P	3DHS	Chain A	60–64, 77–80, 258–260, 288–290, 297–299	18	5 to 17
		Chain B	223–226, 240–243, 255–263	21	10 to 18
		Chain C	83–86		
	3Q1R	Chain B	60–62, 223–226, 240–243, 257–263	25	11 to 20
		Chain C	80–86		

RNAs, which are not considered here. All of the pores found in the 16S rRNA and 23S rRNA measure at least 0.7 nm in the shorter direction with an overall area that typically exceeds 100 Å<sup>2</sup>. The location of the pore-lining residues relative to the rRNA secondary structures is shown in Figure 2, A and B. After pores were found in these structures, we verified the presence and prevalence of corresponding pores in *Escherichia coli* (PDB 3R8N, 3R8O, 3R8S, and 3R8T), *H. marismortui* (PDB 1JJ2), and *D. radiodurans* (PDB 2D3O). Currently, there are no small subunit crystals for either *H. marismortui* or *D. radiodurans*, so we couldn't look for equivalent pores in these 16S rRNAs. Subsequently, we searched for pores in crystals of riboswitches (PDB 2YDH, 3SUH, and 2YIF) the 4.5S RNA (PDB 1DUL and 1HQ1), RNase-P (PDB 3DHS, 3Q1Q, and 3Q1R), and in the partial structure of a RNA spliceosome (PDB 3SIV).

In order to find pores, the RNA portion of the crystals was visualized using Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC). The whole RNA chain was first represented as a Connolly surface through PyMOL's internal surface-finding algorithm. This representation allowed easy recognition of structural pores in the RNA folding. The proteins were not removed, but rather were initially hidden from view. Once the pores had been localized, we used both the ribbon and lines representations to define their properties and nature. We also made distance measurements using PyMOL. In the future, it will be useful to develop a more formal procedure for identifying pores so that none are overlooked.

## SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

## ACKNOWLEDGMENTS

This work was supported in part by a subcontract to G.E.F. from the NASA Center for Ribosome Evolution and Adaptation at the Georgia Institute of Technology (NNA09DA78A). Q.T. was supported by a NASA Earth and Space Science Fellowship (10-Planet10R-0025). M.R. was supported by the Marine Biological Laboratory's NASA Planetary Biology Internship Program. Computational and graphics facilities were provided by the Texas Learning and Computation Center at the University of Houston. We are indebted to Dr. Steve Harvey for helpful discussion and Dr. Antonio Lazcano for his encouragement and support of M.R.

Received February 28, 2013; accepted June 19, 2013.

## REFERENCES

- Adilakshmi T, Ramaswamy P, Woodson SA. 2005. Protein-independent folding pathway of the 16S rRNA 5' domain. *J Mol Biol* **351**: 508–519.
- Agmon I, Bashan A, Zarivach R, Yonath A. 2005. Symmetry at the active site of the ribosome: Structural and functional implications. *Biol Chem* **386**: 833–844.
- Bokov K, Steinberg SV. 2009. A hierarchical model for evolution of 23S ribosomal RNA. *Nature* **457**: 977–980.
- Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Müller KM, et al. 2002. The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BioMed Central Bioinformatics* **3**: 2. [Correction: *BioMed Central Bioinformatics*. **3**: 15.]
- Davidovich C, Belousoff M, Bashan A, Yonath A. 2009. The evolving ribosome: From non-coded peptide bond formation to sophisticated translation machinery. *Res Microbiol* **160**: 487–492.
- Fox GE. 2010. Origin and evolution of the ribosome. *Cold Spring Harb Perspect Biol* **2**: a003483.
- Fox GE, Tran Q, Yonath A. 2012. An exit cavity was crucial to the polymerase activity of the early ribosome. *Astrobiology* **12**: 57–60.
- Hsiao C, Mohan S, Kalahar BK, Williams LD. 2009. Peeling the onion: Ribosomes are ancient molecular fossils. *Mol Biol Evol* **26**: 2415–2425.
- Hury J, Nagaswamy U, Larios-Sanz M, Fox GE. 2006. Ribosome origins: The relative age of 23S rRNA domains. *Orig Life Evol Biosph* **36**: 421–429.
- Kim DF, Green R. 1999. Base-pairing between 23S rRNA and tRNA in the ribosomal A-site. *Mol Cell* **4**: 859–864.
- Leung EK, Suslov N, Tuttle N, Sengupta R, Piccirilli JA. 2011. The mechanism of peptidyl transfer catalysis by the ribosome. *Annu Rev Biochem* **80**: 527–555.
- Nakatogawa H, Ito K. 2002. The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**: 629–636.
- Nierhaus KH, Schulze H, Cooperman BS. 1980. Molecular mechanisms of the ribosomal peptidyl transferase center. *Biochem Int* **1**: 185–192.
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**: 920–930.
- Noller HF, Hoffarth V, Zimniak L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**: 1416–1419.
- Palade GE. 1955. A small particulate component of the cytoplasm. *J Biophys Biochem Cytol* **1**: 59–68.
- Petrov AS, Bernier CR, Hershkovits E, Xue Y, Waterbury CC, Hsiao C, Stepanov VG, Gaucher EA, Grover MA, Harvey SC, et al. 2013. Secondary structure and domain architecture of the 23S rRNA. *Nucleic Acids Res* (in press). doi: 10.1093/nar/gkt513.

- Rodnina MV, Beringer M, Wintermeyer W. 2006. Mechanism of peptide bond formation on the ribosome. *Q Rev Biophys* **39**: 203–225.
- Schmeing TM, Ramakrishnan V. 2009. What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**: 1234–1242.
- Schmeing TM, Huang KS, Kitchen DE, Strobel SA, Steitz TA. 2005. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol Cell* **20**: 437–448.
- Sievers A, Beringer M, Rodnina MV, Wolfenden R. 2004. The ribosome as an entropy trap. *Proc Natl Acad Sci* **101**: 7897–7901.
- Starosta AL, Karpenko VV, Shishkina AV, Mikolajka A, Sumbatyan NV, Schluenzen F, Korshunova GA, Bogdanov AA, Wilson DN. 2010. Interplay between the ribosomal tunnel, nascent chain, and macrolides influences drug inhibition. *Chem Biol* **16**: 1180–1189.
- Trobro S, Aqvist J. 2005. Mechanism of peptide bond synthesis on the ribosome. *Proc Natl Acad Sci* **102**: 12395–12400.
- Voorhees RM, Weixlbaumer A, Loakes D, Kelley AC, Ramakrishnan V. 2009. Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* **16**: 528–533.
- Voss NR, Gerstein M, Steitz TA, Moore PB. 2006. The geometry of the ribosomal polypeptide exit tunnel. *J Mol Biol* **360**: 893–906.
- Watson JD. 1963. Involvement of RNA in the synthesis of proteins. *Science* **140**: 17–26.
- Zaher HS, Shaw JJ, Strobel SA, Green R. 2011. The 2'-OH group of the peptidyl-tRNA stabilizes an active conformation of the ribosomal PTC. *EMBO J* **30**: 2445–2453.