HYPOTHESIS

The driving force for molecular evolution of translation

HARRY F. NOLLER

Center for Molecular Biology of RNA and Department of Molecular, Cell and Developmental Biology, University of California at Santa Cruz, Santa Cruz, California 95064, USA

ABSTRACT

It is widely argued that protein synthesis evolved out of an RNA world, in which catalytic and other biological functions now carried out by proteins were performed by RNAs. However, it is not clear what selective advantage would have provided the driving force for evolution of a primitive translation apparatus, because of the unlikelihood that rudimentary polypeptides would have contributed sufficiently useful biological functions. Here, I suggest that the availability of even simple peptides could have significantly enlarged the otherwise limited structure space of RNA. In other words, translation initially evolved not to create a protein world, but to extend the structural, and therefore the functional, capabilities of the RNA world. Observed examples of substantial structural rearrangements in RNA that are induced by binding of peptides and other small molecules support this possibility.

Keywords: ribosomes; RNA world; peptides; protein synthesis

One of the most baffling problems in molecular evolution is to explain how the process of translation emerged. Early thinking about the origins of the ribosome was confounded by the chicken-or-the-egg question: How could the first ribosome have evolved if it required proteins for its own function? Crick's answer was to ask whether the first ribosomes might have been made solely of RNA (Crick 1968). The realization that the mechanism of action of the ribosome is indeed based primarily on its RNA, and not its proteins, has now largely resolved this question (Noller et al. 1992; Nissen et al. 2000; Ogle et al. 2001; Yusupov et al. 2001; Hoang et al. 2004). Moreover, in vitro selection experiments have directly demonstrated the ability of simple RNA molecules to carry out the four basic reactions of protein synthesis (Zhang and Cech 1997; Illangasekare and Yarus 1999; Lee et al. 2000; Kumar and Yarus 2001).

But recent crystallographic studies have revealed the ribosome to be a structure of unprecedented complexity (Ban et al. 2000; Nissen et al. 2000; Ogle et al. 2001; Yusupov et al. 2001). This introduces a new problem. If, as is widely argued, translation evolved out of an RNA world (Gesteland et al. 1999), populated by molecules of only modest structural complexity, the ribosome and its associated tRNAs

and mRNAs must have evolved in a step-wise fashion from small, preexisting RNAs, which therefore had different (RNA-world) functions. Most importantly, they cannot have evolved initially to make functional proteins, in the modern sense, because of the vanishingly small probability that the first attempts at polypeptide synthesis by a primitive translational apparatus could yield a protein with any useful enzymatic activity (Woese 1967). That is to say, the RNA world could not have anticipated that the evolution of a macromolecular machine with the complexity of the ribosome would in turn eventually lead to the evolution of long polypeptide chains of specific sequence that fold into stable, three-dimensional structures with desirable biological functions. Clearly, until the first active proteins emerged, no selective advantage would exist for evolution of a translational machinery, if its only purpose was to synthesize functional proteins. Yet, we know it evolved, because here it is (Fig. 1).

More likely, protein synthesis initially arose not to create fully functional enzyme-like proteins, but for some other purpose. An attractive alternative is that early primitive peptides could have enhanced the functions of existing RNAs by providing additional functional groups such as positively charged or hydrophobic side chains, acting essentially as cofactors for ribozymes (e.g., see Szathmáry and Maynard Smith 1997). A more subtle possibility, suggested by Poole et al. (1998) is that primitive peptides synthesized in the RNA world could have played a "chaperone-like" role in binding to and stabilizing RNA structures. Here, I would like to make the specific proposal that an even more potent
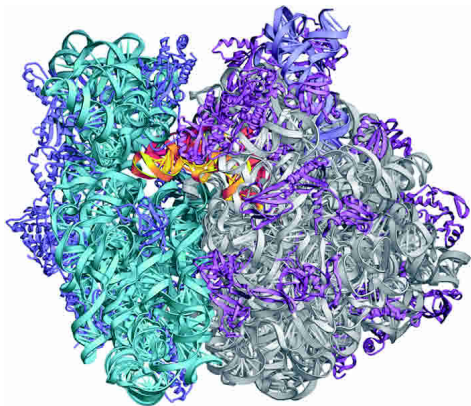
FIGURE 1. The 5.5 Å crystal structure of the *Thermus thermophilus* 70S ribosome (Yusupov et al. 2001). The 30S subunit (*left*) contains 16S rRNA (cyan) and 19 different ribosomal proteins (dark blue). The 50S subunit (*right*) contains 23S rRNA (gray), 5S rRNA (gray-blue), and more than 30 different proteins (magenta).

driving force for selection of a primitive protein synthesis capability was to enlarge the structural repertoire of RNA. In other words, polypeptide synthesis initially evolved not to create a protein world, but to extend the structural (and therefore the functional) capabilities of the RNA world. The elucidation of numerous new RNA structures, for both naturally occuring RNAs and ones obtained through in vitro selection, has revealed a wide variety of RNA folding motifs that extend well beyond the basic double helix and the tertiary folds originally seen in tRNA. Nevertheless, it is becoming clear that the structural versatility of RNA pales by comparison with that of proteins, a condition most likely imposed by the stereochemical constraints inherent in the nucleotide monomer itself (Saenger 1984), as well as the dominant influence of base stacking on RNA folding. This, in addition to the availability of a greater array of chemical groups on the amino acid side chains, is invoked by RNA world advocates to explain why nearly all biological functions are now carried out by proteins. Thus, evolution of function in the RNA world must have been severely restricted by the inherent limitations of RNA structure.

In recent years, structural studies of RNA and RNA–ligand complexes have provided extensive evidence that binding not only of proteins, but even peptides and small-molecular-weight ligands can dramatically influence the three-dimensional structure of RNA, increasing the number and types of stable RNA folds. In the majority of cases, the ligand-binding fold is only observed in the presence of the bound ligand. Most relevant are examples of conformational changes induced by peptide ligands, the most likely products of a primitive translational apparatus.

INTERACTIONS OF RNA WITH SMALL MOLECULES

RNA structure can be profoundly influenced by the binding of many different types of small molecules, as exemplified by structural studies on RNA aptamer–ligand complexes as well as by naturally occurring ligand-binding RNA motifs. A striking example is the structuring of an in vitro selected RNA aptamer by AMP (Dieckmann et al. 1996). Here, a largely unstructured 11-base bulge loop and an opposed G fold into an elaborate three-dimensional cage that surrounds the ligand upon complex formation, incorporating the bound AMP into a GNRA tetraloop-like structure. Small-molecule ligand-induced RNA folding has also been observed in structural studies of several other aptamers, including those that bind vitamin B12 (Sussman et al. 2000), theophylline (Zimmermann et al. 1997), FMN (Fan et al. 1996), and neomycin B (Jiang et al. 1999b).

Thus, the many small molecules that must have inhabited the RNA world could also have played a role in the evolution of RNA structure, by binding to RNAs and extending their structural possibilities. "Riboswitches," naturally occuring RNA modules that are imbedded in mRNA (Mandal and Breaker 2004), may represent relics of such early RNA–ligand interactions. Several kinds of riboswitches have been discovered, with binding specificities for a wide variety of small-molecule cellular ligands, including vitamin B12, thiamine pyrophosphate, flavin mononucleotide, S-adenosylmethionine, lysine, guanine, and adenine (Mandal and Breaker 2004). Riboswitch domains typically consist of ~70–170 nt, and bind their ligands with K_ds from micromolar to nanomolar values. Ligand binding induces substantial conformational changes in the RNA structures of riboswitches that effect transcriptional or translational regulation of the mRNA in which they are imbedded.

PEPTIDE-INDUCED REARRANGEMENT OF RNA STRUCTURE

Although such small-molecule interactions could well extend the range of stable RNA structures, the advent of oligopeptides would present an even richer spectrum of RNA-binding ligands. The combinatorics of peptide sequence possibilities, even given a smaller, more primitive set of amino acids than our present one, could rapidly lead to a vast number of novel molecular shapes and chemical surfaces, offering a large number of potential RNA ligand types. It is not clear why amino acids, rather than some other class of bifunctional molecule, would have been the preferred monomers, but their ready availability from abiotic chemistries (Miller 1953) is a possible explanation.

The classical example of the HIV Tat–TAR interaction provides a vivid illustration of a peptide-induced conformational change in an RNA (Puglisi et al. 1992). The hairpin-like TAR element of the HIV RNA is recognized specifically by the Tat protein, an event that is essential for viral function. The region of the Tat protein essential for binding to TAR RNA was narrowed down to an arginine-rich se-
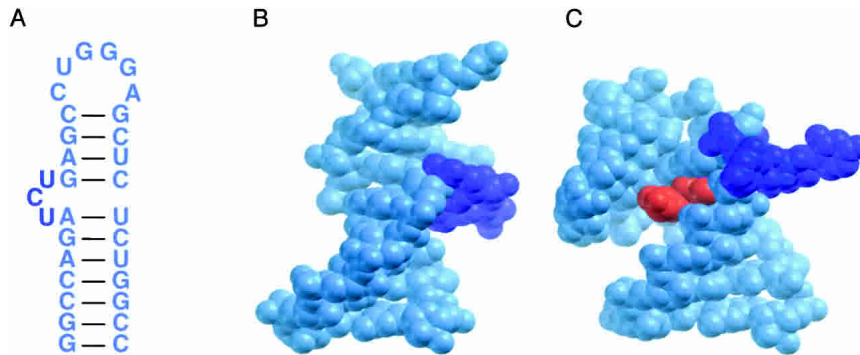
FIGURE 2. The Tat–TAR interaction: An example of ligand-induced structural rearrangement of RNA. (*A*) Secondary structure and NMR solution structures of (*B*) free TAR RNA and (*C*) a complex of argininamide bound to TAR RNA (Puglisi et al. 1992), rendered as space-filling van der Waals surfaces. The UCU bulge loop is shown in dark blue, and the bound argininamide ligand in orange. The structure of TAR RNA in the argininamide complex is the same as in the TAR–peptide complex (Puglisi et al. 1992), but the K_d for the latter is five to six orders of magnitude lower (Tao and Frankel 1992).

quence of nine amino acids (Calnan et al. 1991). Model peptides from this region of Tat are able to recognize and bind specifically to the TAR RNA with nanomolar K_ds. Remarkably, even a single argininamide was shown to bind specifically to TAR (Tao and Frankel 1992), although with greatly decreased affinity. The NMR structures of the free and argininamide-complexed TAR RNA (Puglisi et al. 1992) revealed an extensive ligand-induced structural rearrangement (Fig. 2), which includes unstacking of the bases in the bulge loop, coaxial stacking of the two helical stems, and formation of an A-U-U base triple. NMR studies indicate that the structure of the TAR RNA in the peptide–RNA complex is essentially identical to that observed for the argininamide complex, although the binding affinity of the peptide is five to six orders of magnitude greater than that of the single amino acid derivative (Tao and Frankel 1992).

The Tat–TAR interaction provides one reason for the evolution of polypeptide polymerization. Although the simple monomer argininamide clearly makes all of the molecular interactions with TAR RNA that are needed to stabilize its folded structure, it does so only at very high concentrations that are unlikely to be found in an RNA world setting. By polymerization of arginine with additional amino acids, which apparently contribute mainly nonspecific electrostatic interactions, the same rearrangement is achieved at nanomolar concentrations of peptide. Thus, one rationale for the emergence of a peptidyl transferase activity is to amplify the affinities of small-molecule effectors. A second, related advantage would be to increase the specificity of the interaction, by building in additional amino acids whose presence in the folded complex would be compatible only with a subset of possible structures, for steric or other reasons. One can then imagine a third advantage, as in a hypothetical case where the crucial nucleating interactions themselves (e.g., those provided by the argininamide example) require a dipeptide, tripeptide, and so on.

A second example comes from the binding of an HIV Rev peptide to RRE, another hairpin-like RNA element from the *env* gene of HIV. Binding of the 17-mer peptide induces conformational changes in the RRE RNA that include formation of two noncanonical purine–purine base pairs and stabilization of two Watson–Crick pairs that are not observed in the free RNA (Battiste et al. 1996). This complex is of additional interest, in that the α-helical structure of the REV peptide is formed only upon binding to the RNA (Tan and Frankel 1994). Thus, the structures of both the RNA and the peptide are mutually induced by complex formation. In a further example, the BIV Tat–TAR interaction, the Tat peptide was observed to change from a completely unfolded state to a β-hairpin upon complex formation, while the TAR RNA rearranged to form a base triple (Puglisi et al. 1995). The HIV Rev–RRE and BIV Tat–TAR findings raise the intriguing possibility that the evolution of protein folding could have been bootstrapped by protein–RNA interactions. For example, fixing of marginally stable β-sheet structures by interaction with RNA could have preceded the evolution of more stable domains.

A fourth example is provided by the structure of a peptide derived from the HTLV-1 Rex protein bound to an in vitro selected RNA aptamer (Jiang et al. 1999a). Binding of the 15-mer peptide to the RNA results not only in formation of three base triples but also stabilizes the orientations of three double-helical stems. Thus, in addition to promoting formation of unusual tertiary structural features or noncanonical base pairs, a peptide can even influence or stabilize the relative geometry of separate RNA structural elements, and so direct the overall three-dimensional shape of an RNA.

INFLUENCE OF PROTEINS ON RNA FUNCTION

The idea that simple peptides could expand the functional capabilities of RNA is supported by the fact that most of the functional RNAs found in present-day cells depend on proteins for their activity. Ribosomes themselves, although fundamentally ribozymes in nature, still require proteins to fold their RNAs into biologically active conformations and to optimize the speed and accuracy of their functions (Garrett et al. 2000; CSHSQB 2001). Like the peptides discussed above, ribosomal proteins have been observed to stabilize unusual tertiary structural elements and noncanonical base pairs (Stern et al. 1989), and are often found at multihelix junctions, where they have in some cases been shown to fix

the relative orientation of different helical elements (Orr et al. 1998). The activity of ribonuclease P has been shown to be determined by its RNA component, but is inactive under physiological conditions in the absence of its protein subunit (Guerrier-Takada et al. 1983). The activity of spliceosomes, responsible for splicing introns out of pre-mRNA, is believed to be based on spliceosomal RNAs. Spliceosomes employ a mechanism that closely resembles the one used by the group II self-splicing introns, which bear striking structural similarity to the spliceosomal RNAs (Sharp 1991). Yet, the spliceosome contains on the order of a hundred proteins, without which the RNA components are inactive. The influence of proteins on the RNA components of these nucleoprotein complexes may provide clues to the ways in which small peptides could have triggered the transition into the protein world.

Generation of functionally successful peptide ligands would provide a selective advantage for the primitive protein synthesis apparatus to its RNA hosts. At some point, the existence of a useful suite of RNA-binding polypeptides would in turn influence the evolution of the RNAs to develop improved peptide-binding properties, as well as a greater range of peptide-induced structures and functions. Thus, the polypeptides themselves would begin to play an important role in further improvements in RNA function (including protein synthesis) and in the overall evolution of an increasingly interdependent protein–RNA system.

It is easiest to think of synthesis of the first simple peptides as taking place independently of coding, giving rise to a comparatively limited number of short peptides of fixed or even random sequence. Specification of amino acids by RNA sequences most likely emerged later, requiring coevolution of the ribosome and its tRNAs (Noller 1993; Schimmel and Henderson 1994) ultimately leading to our present, protein-dominated form of biology. Most difficult to explain is this final phase of evolution of the ribosome, the emergence of coding, which gave rise to "one of the great evolutionary saltations…that would generate a truly enormous new, totally unique evolutionary phase space" (Woese 2004).

ACKNOWLEDGMENTS

REFERENCES

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.

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

Calnan, B.J., Tidor, B., Biancalana, S., Hudson, D., and Frankel, A.D. 1991. Arginine-mediated RNA recognition: The arginine fork. *Science* **252:** 1167–1171.

Crick, F.H.C. 1968. The origin of the genetic code. *J. Mol. Biol.* **38:** 367–379.

CSHSQB. 2001. *The ribosome.* In *Cold Spring Harbor Symp. Quant. Bio. 64.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Dieckmann, T., Suzuki, E., Nakamura, G.K., and Feigon, J. 1996. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* **2:** 628–640.

Fan, P., Suri, A.K., Fiala, R., Live, D., and Patel, D.J. 1996. Molecular recognition in the FMN–RNA aptamer complex. *J. Mol. Biol.* **258:** 480–500.

Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B., and Noller, H.F. 2000. *The ribosome: Structure, function, antibiotics and cellular interactions.* ASM Press, Washington, DC.

Gesteland, R.F., Cech, T., and Atkins, J.F., eds. 1999. *The RNA world*, 2d ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849–857.

Hoang, L., Fredrick, K., and Noller, H.F. 2004. Creating ribosomes with an all-RNA 30S subunit P site. *Proc. Natl. Acad. Sci.* **101:** 12439–12443.

Illangasekare, M. and Yarus, M. 1999. Specific, rapid synthesis of Phe-RNA by RNA. *Proc. Natl. Acad. Sci.* **96:** 5470–5475.

Jiang, F., Gorin, A., Hu, W., Majumdar, A., Baskerville, S., Xu, W., Ellington, A., and Patel, D.J. 1999a. Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples. *Struct. Fold. Des.* **7:** 1461–1472.

Jiang, L., Majumdar, A., Hu, W., Jaishree, T.J., Xu, W., and Patel, D.J. 1999b. Saccharide–RNA recognition in a complex formed between neomycin B and an RNA aptamer. *Struct. Fold. Des.* **7:** 817–827.

Kumar, R.K. and Yarus, M. 2001. RNA-catalyzed amino acid activation. *Biochemistry* **40:** 6998–7004.

Lee, N., Bessho, Y., Wei, K., Szostak, J.W., and Suga, H. 2000. Ribozyme-catalyzed tRNA aminoacylation. *Nat. Struct. Biol.* **7:** 28–33.

Mandal, M. and Breaker, R.R. 2004. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell. Biol.* **5:** 451–463.

Miller, S.L. 1953. A production of amino acids under possible primitive earth conditions. *Science* **117:** 528–529.

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.

Noller, H.F. 1993. On the origin of the ribosome: Coevolution of subdomains of tRNA and rRNA. In *The RNA world* (eds. R.F. Gesteland et al.), pp. 137–156. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Noller, H.F., Hoffarth, V., and Zimniak, L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256:** 1416–1419.

Ogle, J.M., Brodersen, D.E., Clemons, W.M., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292:** 897–902.

Orr, J.W., Hagerman, P.J., and Williamson, J.R. 1998. Protein and Mg(2+)-induced conformational changes in the S15 binding site of 16 S ribosomal RNA. *J. Mol. Biol.* **275:** 453–464.

Poole, A.M., Jeffares, D.C., and Penny, D. 1998. The path from the RNA world. *J. Mol. Evol.* **46:** 1–17.

Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D., and Williamson, J.R. 1992. Conformation of the TAR RNA–arginine complex by NMR spectroscopy. *Science* **257:** 76–80.

Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. 1995. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide–RNA complex. *Science* **270:** 1200–1203.

Saenger, W. 1984. *Principles of nucleic acid structure.* Springer-Verlag, New York.

Schimmel, P. and Henderson, B. 1994. Possible role of aminoacyl–RNA complexes in noncoded peptide synthesis and origin of coded synthesis. *Proc. Natl. Acad. Sci.* **91:** 11283–11286.

Sharp, P.A. 1991. "Five easy pieces." *Science* **254:** 663.

Stern, S., Powers, T., Changchien, L.M., and Noller, H.F. 1989. RNA–protein interactions in 30S ribosomal subunits: Folding and function of 16S rRNA. *Science* **244:** 783–790.

Sussman, D., Nix, J.C., and Wilson, C. 2000. The structural basis for molecular recognition by the vitamin B 12 RNA aptamer. *Nat. Struct. Biol.* **7:** 53–57.

Szathmáry, E. and Maynard Smith, J. 1997. From replicators to re-producers: The first major transitions leading to life. *J. Theor. Biol.* **187:** 555–571.

Tan, R. and Frankel, A.D. 1994. Costabilization of peptide and RNA structure in an HIV Rev peptide–RRE complex. *Biochemistry* **33:** 14579–14585.

Tao, J. and Frankel, A.D. 1992. Specific binding of arginine to TAR RNA. *Proc. Natl. Acad. Sci.* **89:** 2723–2726.

Woese, C.R. 1967. *The genetic code. The molecular basis for genetic expression.* Harper & Row, New York.

———. 2004. A new biology for a new century. *Microbiol. Mol. Biol. Rev.* **68:** 173–186.

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

Zhang, B. and Cech, T.R. 1997. Peptide bond formation by in vitro selected ribozymes. *Nature* **390:** 96–100.

Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P., and Pardi, A. 1997. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat. Struct. Biol.* **4:** 644–649.

RNA

A PUBLICATION OF THE RNA SOCIETY

The driving force for molecular evolution of translation

HARRY F. NOLLER

RNA 2004 10: 1833-1837

References	This article cites 33 articles, 18 of which can be accessed free at: http://rnajournal.cshlp.org/content/10/12/1833.full.html#ref-list-1
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**