

RNA and the Small Molecule World

Yitzhak Tor*

Our understanding of the flow of biochemical information from deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) and then to proteins has evolved enormously in the last few decades.^[1, 2] DNA sequences coding for specific genes are transcribed in the cell nucleus to yield heterogeneous nuclear RNAs. These primary transcripts are processed to yield mature messenger RNAs that are then transported to the cytoplasm where translation, namely, protein synthesis, takes place in the ribosomes. The transport, stability, cellular localization, and translation efficiency of individual messenger RNA molecules are all well controlled. This multi-level regulation, mediated by numerous RNA-binding proteins and ribonucleoprotein complexes, ultimately governs gene expression and its functional manifestation.^[3] RNA is therefore no longer viewed as a passive carrier of genetic information. Rather, it is appreciated as a structurally and functionally sophisticated biomolecule intimately involved in key cellular processes. As such, it is quickly becoming a central target for drug design.^[4, 5] It is illustrative to follow how RNA has assumed this pivotal role in modern bioorganic chemistry.^[6]

Experimental evidence correlating RNA with protein synthesis started emerging in the late 1930s from the laboratories of Caspersson and Brachet. In their 1939 *Nature* paper, Caspersson and Schultz stated: "The presence of pentose nucleotides in high concentrations in rapidly dividing tissues is probably thus a general phenomenon".^[7] At about the same time, small, RNA-rich particles, initially named microsomes and later ribosomes, had been noticed by Claude.^[8] In 1946 (at the first major biological symposium focusing on nucleic acids), Brachet put forward the hypothesis that these granules are responsible for protein biosynthesis.^[9] Classic electron microscopy work by Palade,^[10] and pulse-chase experiments by Zamecnik,^[11] provided the necessary experimental support for this hypothesis almost a decade later.^[12]

In summarizing a 1947 *Experientia* paper by Boivin and Vendrely, an anonymous editor concluded: "Through catalytic actions the macromolecular desoxyribonucleic acids govern the building of macromolecular ribonucleic acids, and, in turn, these control the production of cytoplasmic enzymes".^[13] This statement, although formulating the es-

sence of biological information flow, did not specify the exact components and mechanisms involved. After the 1953 elucidation of the double helical structure of DNA by Watson and Crick,^[14] attention turned to the roles RNA plays in gene expression. The view held in the late 1950s was that each gene controls the synthesis of a single specialized ribosome, which in turn governs the synthesis of a specific protein. In this scheme, epitomized as the one gene–one ribosome–one protein hypothesis, ribosomal RNA was postulated to be the template for protein synthesis. The difficulties encountered with this model led Jacob and Monod to hypothesize that the information is carried by a short-lived RNA molecule and that the ribosomes were nonspecialized structures.^[15] Their 1961 prediction was soon confirmed by the discovery of messenger RNA (mRNA), a labile template that turns over rapidly.^[16] The fundamentals of translation began to unfold when a soluble RNA, named transfer RNA (tRNA), was shown to be involved in protein synthesis by carrying an activated amino acid,^[17, 18] an observation that supported Crick's adaptor hypothesis.^[19] In 1966, the genetic code—an algorithm relating triplets of nucleotides in genes to specific amino acids in proteins—was established.^[20]

The central machinery for protein biosynthesis, where both decoding and peptidyl transferase take place, is the ribosome: a complex of ribosomal RNA (rRNA) and numerous proteins.^[21] This large and as yet unsolved structure comprises 85 % of all cellular RNA. Intriguing experiments reported by Noller and co-workers in 1992 suggested that protein-depleted rRNA has, by itself, a peptidyl transferase activity.^[22] These fascinating observations, indicating catalytic activity of rRNA, have been very recently substantiated by Watanabe et al. They showed that the six domains of the 23S rRNA of *E. coli* stimulated peptide-bond formation when assembled together in the absence of any ribosomal proteins.^[23] Taken together with the well-characterized RNA enzymes (ribozymes),^[24] RNA is once more taking center stage as a key functional biomolecule.^[25]

The functional diversity of RNA can be attributed to the intricate three-dimensional folds it can assume through a multitude of secondary structures and tertiary interactions.^[26] This structural sophistication, together with the spatial projection of functional groups and the electrostatic field generated by the RNA fold, create potential binding pockets for ions, small molecules, and proteins.^[27] Ideally, one should be able to take advantage of this promising recognition capacity to identify common binding motifs and utilize them

[*] Prof. Dr. Y. Tor
Department of Chemistry and Biochemistry
University of California, San Diego
La Jolla, CA 92093–0358 (USA)
Fax: (+1)619-534-5383
E-mail: ytor@ucsd.edu

for the design of new RNA binders as modulators of cellular functions.^[5]

The search for small molecules as RNA binders has been inspired by early observations that indicated the ability of antibiotics to interfere with protein biosynthesis.^[28] In a seminal paper published in 1987, Noller and Moazed showed that several classes of antibiotics, among them aminoglycosides, interact with functional sites on 16S rRNA.^[29] This discovery, suggesting the direct binding of low molecular weight drugs to rRNA, triggered a renewed interest in RNA–ligand recognition. Since then, aminoglycoside antibiotics have been shown to interact with a number of functional RNA molecules including group I introns,^[30] hammerhead ribozymes,^[31] and the human hepatitis delta virus ribozyme.^[32] Their binding to the A-site analogue of 16S rRNA has been established,^[33–35] and their interactions with tRNA have been studied in detail.^[36] Aminoglycoside-specific RNA aptamers have been selected in vitro,^[37] and NMR studies have revealed the structures of the complexes with their cognate aminoglycosides.^[38] Certain aminoglycoside antibiotics and synthetic ligands have been shown to interfere with RNA–protein binding.^[39–43] Issues of specificity in RNA–ligand interactions,^[44] as well as binding mechanism and dynamics have been recently addressed.^[45–47] Although the challenges are still enormous, RNA has become a promising and fertile target for ligand design.^[5]

Werstuck and Green have recently added an exciting new dimension to this rapidly growing area of RNA recognition. They have demonstrated that gene expression can be controlled in living cells through specific interactions between RNA and small molecules.^[48] Specifically, Werstuck and Green engineered unique RNA sequences into bacterial and mammalian mRNA and mediated essential cellular functions by adding small and cell-permeable molecules. Although artificially constructed, the systems reported, provide a proof of principle and the first demonstration that small molecules can act as cellular regulators at the mRNA level.

Werstuck and Green have chosen RNA aptamers, short RNA sequences that are selected in vitro to bind small molecules with high affinity and selectivity, as the targeted sites. Initially, aptamers recognizing tobramycin and kanamycin, two aminoglycoside antibiotics, have been employed (Figure 1). Plasmids expressing these RNA aptamers have been constructed by cloning the corresponding DNA sequences into a T7 RNA polymerase-driven expression vector. An *Escherichia coli* strain containing an inducible T7 RNA polymerase was then transformed with these plasmids. To test for the in vivo function of these RNA sequences, Werstuck and Green have looked for a drug-resistant phenotype. Thus, bacteria expressing the kanamycin-binding aptamer were able to grow in the presence of kanamycin A, the cognate bactericidal antibiotic. In contrast, the growth of untransformed bacteria and a strain producing a tobramycin-binding aptamer was negligible. Similarly, bacteria containing the tobramycin-recognizing aptamer were able to grow in the presence of tobramycin, the

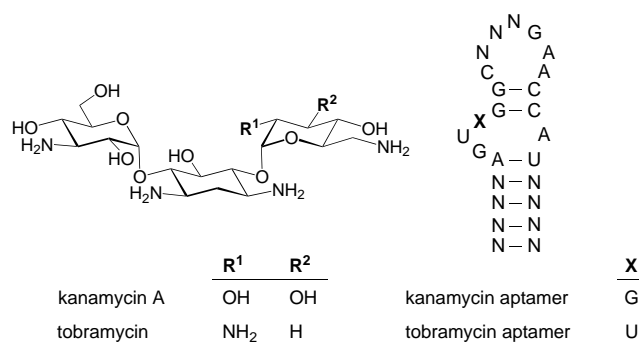


Figure 1. Aminoglycoside antibiotics and their corresponding RNA aptamers.^[48]

cognate antibiotic, while their growth was significantly repressed by kanamycin A, the noncognate small molecule.^[48]

This series of experiments established that a drug-resistant bacterial strain can be realized by the expression of a drug-binding RNA molecule. These RNA aptamers generated in vivo act as a scavenger of a substance that otherwise would be toxic to the cell. In certain respects, the approach taken by Werstuck and Green is reminiscent of actual resistance mechanisms that operate in bacteria. The most prominent mechanism for the acquired microbial resistance to aminoglycosides is the inactivation of these antibiotics by plasmid-encoded enzymes.^[49] Numerous enzymes that phosphorylate, adenylylate, or acetylate the aminoglycosides have been identified. It is worth noting that these aminoglycoside-modifying enzymes have severely hampered the clinical usefulness of aminoglycoside antibiotics.

To demonstrate the utilization of small molecules as modulators of eukaryotic gene expression, Werstuck and Green have focused on regulating the translation initiation.^[48] The template for this process, the eukaryotic mRNA, is commonly posttranscriptionally modified at the 5'-end by a N7-methylated guanosine nucleotide, which acts as a cap. Translation is typically preceded by a scanning process through the 5'-untranslated region (UTR)—the mRNA sequence spanning between the 5'-cap and a start codon (Figure 2). In this key process a pre-initiation complex composed of a ribosomal subunit, initiation factors, and a charged initiator tRNA^{Met}, migrates along the 5'-UTR until it encounters a start codon, normally the first AUG triplet. Translational regulation via the 5'-UTR can therefore be mediated by several mechanisms, including modification at

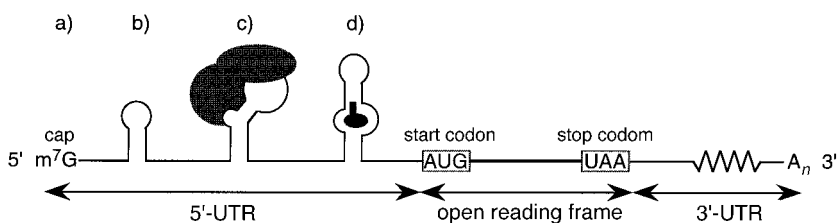


Figure 2. Schematic illustration of functional elements and their organization within eukaryotic mRNA. Translational regulation at the 5'-UTR can be mediated by a) modification of the cap structure, b) a stable secondary structure, c) RNA–protein interactions, and, as hypothesized by Werstuck and Green, d) interactions between RNA and small molecules. Note that additional regulation elements that exist in the 3'-UTR are not shown.^[3b]

the cap structure (Figure 2a), the introduction of stable secondary structures into the UTR (Figure 2b), and through protein–RNA binding within the UTR (Figure 2c).^[3b, 50] Werstuck and Green hypothesized that the incorporation of a stable small molecule–aptamer complex into the 5'-UTR may impede the initiation apparatus by blocking the scanning processes or the ribosome–mRNA binding (Figure 2d).

Aminoglycoside antibiotics are highly charged at physiological pH values, relatively impermeable, and are known to interfere with translation. Werstuck and Green have therefore selected RNA aptamers that specifically bind the benzimidazole-containing Hoechst dye H33258, a known DNA minor groove binder (Figure 3). Two aptamers were then inserted

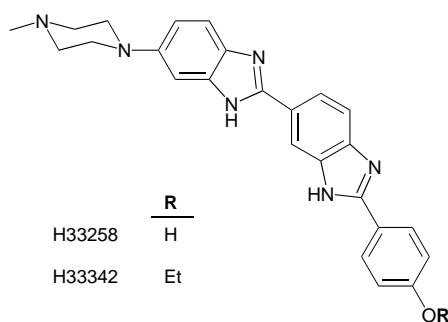


Figure 3. Structure of Hoechst dyes H33258 and H33342. H33258 was used to generate RNA aptamers, and the more soluble derivative H33342 was used to regulate gene expression in living cells that express the corresponding aptamer.^[48]

into the 5'-UTR of a mammalian β -galactosidase expression plasmid, and Chinese hamster ovary cells were cotransfected with the engineered vector and a luciferase reporter gene. To test for the in vivo function of these RNA aptamers, Werstuck and Green have analyzed for β -galactosidase and luciferase activities after translation took place, as an indication for the successful expression of these genes. In the absence of a Hoechst dye, the incorporated aptamers did not influence gene expression, which indicated that the translation machinery had not been affected by the presence of these “foreign” constructs within the 5'-UTR. In contrast, in the presence of increasing concentrations of H33342 (a closely related more soluble Hoechst dye; Figure 3), the β -galactosidase activity had dropped substantially with no effect on the luciferase control activity.^[48]

This set of experiments illustrates how small organic molecules can, in principle, be used to control gene expression in living cells. The ability to regulate the translation of exogenous genes with small molecules that act as a switch opens up exciting opportunities. It may be beneficial for gene therapy applications, where genes containing a built-in switch are introduced and their expression is kept under the control of the cognate small molecule. It may also be useful for manipulating specific genes as a tool to investigate the biological outcome of their expression.

The elegant use of RNA aptamers to create controllable genes circumvents a major problem of RNA binding specificity. RNA aptamers, selected under highly stringent conditions, typically show extremely high affinity and specificity for their

cognate ligands. This is not the case with “naturally occurring” RNA sequences and their known ligands. Aminoglycoside antibiotics, for example, are rather nonspecific RNA binders. They bind numerous RNA sequences with similar affinity. Their promiscuous binding is likely to result from their recognition of similar RNA folds rather than sequences.^[5] In certain cases aminoglycoside binding is accommodated by a conformational change in the RNA host.^[35c] These issues, together with the minute cellular amount of mRNA relative to rRNA and tRNA, makes the de novo design of highly selective mRNA binders a formidable task. The in vivo utilization of small molecules as regulators of endogenous gene expression remains a major challenge. It is likely to get resolved as our knowledge of RNA structure, folding, and recognition advances.

The area of small molecule recognition by RNA is growing rapidly as fast screening methods for evaluating RNA affinity and specificity of potential binders are developed.^[51] Studies of structure–activity relationships of combinatorial libraries complemented by novel naturally occurring products will further advance our understanding of the interactions between RNA and small molecules. As our knowledge of RNA–protein interactions evolves, new RNA targets will be identified and targeted by small molecules. We are likely to encounter small molecules that thrive as specific RNA binders.

The last two decades have brought bacterial and viral infections to center stage. Numerous new viruses have appeared. Their effect may be inapparent (for example, hepatitis G), yet in other cases it might have devastating consequences (for example, HIV).^[52] Concurrently, there has been a resurgence of previously known viral and bacterial diseases, as well as an emergence of resistant mutants of well-studied pathogens.^[53] These changes have produced new targets for antiviral and antibacterial therapy and present a major challenge to the scientific community. The thrilling days of antibiotics discovery, less than 60 years ago, have now been replaced by a deep concern for the future, and the necessity to uncover new approaches for antiviral and antibacterial therapy. Since RNA molecules play key roles in protein synthesis, transcriptional regulation, and viral replication novel antibiotic and antiviral agents may well be found within small molecules that target RNA sites.

German version: *Angew. Chem.* **1999**, *111*, 1681–1685

Keywords: aminoglycosides • antibiotics • gene technology • RNA • translation

- [1] For a wonderful monograph discussing the evolution of molecular biology, its major discoveries and makers, see H. F. Judson, *The Eighth Day of Creation*, Simon & Schuster, New York, **1980**.
- [2] A. Rich in *Horizons in Biochemistry* (Eds.: M. Kasha, B. Pullman), Academic Press, New York, **1962**, pp. 103–126.
- [3] Recent reviews: a) G. Varani, K. Nagai, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 407–445; b) N. K. Gray, M. Wickens, *Annu. Rev. Cell. Dev. Biol.* **1998**, *14*, 399–458; c) H. Siomi, G. Dreyfuss, *Curr. Opin. Gen. Dev.* **1997**, *7*, 345–353.
- [4] a) T. Hermann, E. Westhof, *Curr. Opin. Biotechnol.* **1998**, *9*, 66–73; b) N. D. Pearson, C. D. Prescott, *Chem. Biol.* **1997**, *4*, 409–414.

- [5] K. Michael, Y. Tor, *Chem. Eur. J.* **1998**, *4*, 2091–2098.
- [6] It is impossible to capture all the exciting developments and their contributors within such a short account. The interested reader is referred to ref. [1] and to an excellent textbook: J. D. Watson, N. H. Hopkins, J. W. Roberts, J. A. Steitz, A. M. Weiner, *Molecular Biology of the Gene*, 4th ed., Benjamin/Cummings, Menlo Park, **1987**.
- [7] T. Caspersson, J. Schultz, *Nature* **1939**, *143*, 602–603. In those days the term pentose nucleic acids was used in reference to RNA. See T. Caspersson, *Chromosoma* **1940**, *1*, 605–619.
- [8] a) A. Claude, *Science* **1940**, *91*, 77–78; b) A. Claude, *Cold Spring Harbor Symp. Quant. Biol.* **1941**, *9*, 263–271.
- [9] J. Brachet, *Symp. Soc. Exp. Biol.* **1947**, *1*, 207–224.
- [10] G. E. Palade, *J. Biophys. Biochem. Cytol.* **1955**, *1*, 59–68.
- [11] J. W. Littlefield, E. B. Keller, J. Gross, P. C. Zamecnik, *J. Biol. Chem.* **1955**, *217*, 111–123.
- [12] For a historical account, see A. Tissières in *Ribosomes* (Eds.: M. Nomura, A. Tissières), Cold Spring Harbor Laboratory Press, New York, **1974**, pp. 3–12.
- [13] A. Boivin, R. Vendrely, *Experientia* **1947**, *3*, 32–34.
- [14] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.
- [15] F. Jacob, J. Monod, *J. Mol. Biol.* **1961**, *3*, 318–356.
- [16] a) S. Brenner, F. Jacob, M. Meselson, *Nature* **1961**, *190*, 576–581; b) F. Gros, H. Hiatt, W. Gilbert, C. G. Kurland, R. W. Risebrough, J. D. Watson, *Nature* **1961**, *190*, 581–585.
- [17] M. B. Hoagland, M. L. Stephenson, J. F. Scott, L. I. Hecht, P. C. Zamecnik, *J. Biol. Chem.* **1958**, *231*, 241–257. See also M. Hoagland, *Trends Biochem. Sci.* **1996**, *21*, 77–80.
- [18] F. Chapeville, F. Lipman, G. von Ehrenstein, B. Weisblum, W. J. Ray, S. Benzer, *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 1086–1092.
- [19] F. H. C. Crick, an unpublished note to the RNA Tie Club, **1955**. See also ref. [1], and F. H. C. Crick, *What Mad Pursuit*, Basic Books, New York, **1988**.
- [20] A key experiment: M. W. Nirenberg, J. H. Matthaei, *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1588–1602. See also “The Genetic Code”: *Cold Spring Harbor Symp. Quant. Biol.* **1966**, *31*.
- [21] R. Green, H. F. Noller, *Annu. Rev. Biochem.* **1997**, *66*, 679–716.
- [22] H. F. Noller, V. Hoffarth, L. Zimniak, *Science* **1992**, *256*, 1416–1419.
- [23] I. Nitta, Y. Kamada, H. Noda, T. Ueda, K. Watanabe, *Science* **1998**, *281*, 666–669.
- [24] T. R. Cech, *Science* **1987**, *236*, 1532–1539.
- [25] P. Schimmel, R. Alexander, *Science* **1998**, *281*, 658–659.
- [26] a) M. Chastain, I. Tinoco, Jr., *Prog. Nucl. Acid Res. Mol. Biol.* **1991**, *41*, 131–177; b) J. Doudna, R. T. Batey, R. P. Rambo, *Angew. Chem.* **1999**, in press.
- [27] C. S. Chow, F. M. Bogdan, *Chem. Rev.* **1997**, *97*, 1489–1513.
- [28] a) B. A. Newton, P. E. Reynolds, *Biochemical Studies of Antimicrobial Drugs*, University Press, Cambridge, **1966**; b) D. Vázquez, *Inhibitors of Protein Biosynthesis*, Springer, Berlin, **1979**; c) E. F. Gale, E. Cundliffe, P. E. Reynolds, M. H. Richmond, M. J. Waring, *The Molecular Basis of Antibiotics Action*, 2nd ed., Wiley, Chichester, **1981**.
- [29] D. Moazed, H. F. Noller, *Nature* **1987**, *327*, 389–394.
- [30] a) U. von Ahsen, J. Davies, R. Schroeder, *Nature* **1991**, *353*, 368–370; b) U. von Ahsen, J. Davies, R. Schroeder, *J. Mol. Biol.* **1992**, *226*, 935–941; c) J. Davies, U. von Ahsen, R. Schroeder in *The RNA World* (Eds.: R. F. Gesteland, J. F. Atkins), Cold Spring Harbor Laboratory Press, New York, **1993**, pp. 185–204.
- [31] T. K. Stage, K. J. Hertel, O. C. Uhlenbeck, *RNA* **1995**, *1*, 95–101.
- [32] a) J. Rogers, A. H. Chang, U. von Ahsen, R. Schroeder, J. Davies, *J. Mol. Biol.* **1996**, *259*, 916–925; b) J.-S. Chia, H.-L. Wu, H.-W. Wang, D.-S. Chen, P.-J. Chen, *J. Biomed. Sci.* **1997**, *4*, 208–216.
- [33] P. Purohit, S. Stern, *Nature* **1994**, *370*, 659–662.
- [34] H. Miyaguchi, H. Narita, K. Sakamoto, S. Yokoyama, *Nucleic Acids Res.* **1996**, *24*, 3700–3706.
- [35] a) M. I. Recht, D. Fourmy, S. C. Blanchard, K. D. Dahlquist, J. D. Puglisi, *J. Mol. Biol.* **1996**, *262*, 421–436; b) D. Fourmy, M. I. Recht, S. C. Blanchard, J. D. Puglisi, *Science* **1996**, *274*, 1367–1371; c) D. Fourmy, S. Yoshizawa, J. D. Puglisi, *J. Mol. Biol.* **1998**, *277*, 333–345.
- [36] S. R. Kirk, Y. Tor, *Bioorg. Med. Chem.*, in press.
- [37] a) Y. Wang, R. R. Rando, *Chem. Biol.* **1995**, *2*, 281–290; b) S. M. Lato, A. R. Boles, A. D. Ellington, *Chem. Biol.* **1995**, *2*, 291–303; c) M. G. Wallis, U. von Ahsen, R. Schroeder, M. Famulok, *Chem. Biol.* **1995**, *2*, 543–552; d) M. Famulok, A. Hüttenhofer, *Biochemistry* **1996**, *35*, 4265–4270; e) S. T. Wallace, R. Schroeder, *RNA* **1998**, *4*, 112–123.
- [38] a) L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, *Chem. Biol.* **1997**, *4*, 35–50. b) L. C. Jiang, D. J. Patel, *Nat. Struct. Biol.* **1998**, *5*, 769–774.
- [39] M. L. Zapp, S.