Nucleic acids RNA: from architecture to recognition and catalysis

Editorial overview
Eric Westhof and Dinshaw J Patel

The emphasis of this section of *Current Opinion in Structural Biology* is on RNA architecture, recognition and catalysis. The topics covered range from ribozymes to riboswitches, as well as molecular machines such as the ribosome and the machinery for RNA-silencing-mediated cleavage of mRNA. We now realize that small 21–25-nucleotide RNAs control gene regulation and are central to eukaryotic complexity. Such evolving intricate regulatory networks are fundamentally controlled by the dynamics and the energetics of molecular recognition between RNA molecules, as well as between RNA molecules and their various partners. Molecular recognition is therefore a recurring theme in this overview, with the emphasis on complexes of RNA with metabolites, antibiotics and proteins.

For those interested in catalytic RNAs or ribozymes, it is a fascinating time. Indeed, we now have at our disposal several X-ray crystallographic structures at a good resolution, and sophisticated new biophysical techniques to study the kinetics and equilibria present in these complex molecules. At the same time, new progress in the study of the chemistry of nucleic acids has led to an increase in the availability of natural and modified RNA molecules. Two reviews focus on ribozymes. David Lilley reviews overall progress concerning the chemistry of catalytic RNA, from small to large ribozymes. Sarah Woodson focuses her review on the archetype of the large ribozymes, the family of group I introns.

Except for the catalysis of peptide bond formation by the RNA components of the ribosome, all the reactions catalyzed by natural RNAs involve cleavage and ligation of the phosphodiester backbone. David Lilley surveys the various contributions to the catalytic rate enhancement intervening in nucleolytic ribozymes. General acid-base catalysis mediated by metal ions has been, to date, the most investigated mechanism. But recent efforts have demonstrated that general acid-base catalysis mediated by cytosine and adenine bases occurs in both the hairpin and hepatitis delta virus ribozymes. In contrast, but still debated, orientation and proximity of substrates are the main factors in rate enhancement of peptide bond formation in the ribosome. Another central connection between RNA catalysis and folding has been brought to the forefront recently; architecture, implying long-range RNA–RNA contacts, is important not only in the large ribozymes, such as the group I introns, the group II introns and RNase P, but also in the small-sized ribozymes, such as the hammerhead and hairpin ribozymes. Interactions between auxiliary elements, critical for the folding process and the maintenance of the fold, contribute very significantly to the rate enhancements and the requirements for divalent ions.
Since their discovery more than 20 years ago by Tom Cech and colleagues, the group I introns have been the subject of intense investigation and, thus, have served as models for understanding the relationships between sequence, structure and catalysis in RNA. Striking crystal structures of three types of group I introns have appeared in the past few months. Sarah Woodson highlights the structural breakthroughs these crystallographic feats have brought and compares the architectures. Group I introns are classified into various subgroups by the type and diversity of secondary structure elements grafted onto the invariant core. These elements contact each other in diverse topological fashions by loop–loop or GNRA–receptor interactions embracing the core. A surprising feature of the new structures is the sandwiching of the G-cofactor by an ACA sequence: the G-cofactor stacks between the first adenine and the cytosine, which pairs with a guanine of another strand to which the second adenine binds as a sheared A$^\text{c}$G pair. Although all the long-range contacts had been previously predicted on the basis of sequence comparisons and modeling, these amazing crystal structures have unlocked several atomic mysteries and, importantly, raise new challenges to our understanding of these subtle molecular machines.

It is becoming increasingly apparent that RNA is a key player in the regulation of gene expression. Two reviews address recent advances on this topic. One of these outlines the explosive advances in RNA interference, spanning post-transcriptional regulation of mRNA to chromatin regulation through modification and silencing. The other introduces riboswitches, mRNA control elements that undergo metabolite-induced conformational transitions, thereby regulating metabolite gene expression.

Witold Filipowicz et al. outline the tremendous mechanistic and structural advances that have begun to clarify how small interfering RNAs (siRNAs) and micro RNAs (miRNAs) regulate post-transcriptional gene silencing. The review outlines our current understanding of the pathways leading to siRNA generation, the pathways responsible for selection and incorporation of the siRNA guide strand into the Argonaute-containing RNA-induced silencing complex (RISC), and how subsequent pairing with complementary mRNA leads to cleavage or translational inhibition. Key biochemical developments include identification of steps leading to the formation of the RISC loading complex and the sensing apparatus responsible for strand selection into the RISC. Structural biology, when complemented by biochemical experiments, has provided molecular insights into the key interaction events governing steps within the catalytic cycle. These advances include the demonstration that the two RNase III domains of Dicer fold intramolecularly to form a single processing center, which cuts the double-stranded RNA on both strands to generate siRNAs with 2′-nucleotide 3′ overhangs. The structure of the human PAZ domain, a nucleic-acid-binding module common to both Dicer and Argonaute proteins, bound to a mini-siRNA has established how PAZ anchors the 2-nucleotide 3′ overhangs and contacts the anchored siRNA strand in a sequence-independent manner. Structural studies of P. furiosus Argonaute and A. fulgidus Piwi proteins, both in the free state, have shown that the Piwi domain adopts an RNase H fold, which harbors the elusive slicer activity responsible for mRNA cleavage, a conclusion supported by mutation studies of corresponding catalytic cleavage residues in human Ago2. Finally, crystal structures of the viral suppressor p19 bound to siRNAs establish how a dimeric p19 uses a caliper-like end-bracketing mechanism to sequester siRNA duplexes of defined length. These successes will undoubtedly be expanded to elucidate, in the near future, how 5′-phosphate-containing ends of siRNAs are recognized and the nature of the intermolecular contacts responsible for alignment of single-stranded RNAs within the Argonaute scaffold. The pace of fundamental discoveries in the RNA silencing field remains unabated and the recent structures of protein–siRNA complexes attest to their functional impact on this field.

The review by Brian Tucker and Ronald Breaker introduces the concept of riboswitches — metabolite-sensing mRNAs that control gene expression, through capitalization of the molecular recognition and allosteric properties of RNA. Riboswitches, located in the 5′-untranslated segments of mRNAs, are structured elements composed of conserved metabolite-binding aptamer domains and expression platforms, which modulate gene expression through control of transcriptional termination or translational initiation. Such regulation is achieved through formation of intrinsic transcriptional terminators, or through sequestration of Shine–Delgarno sequences and ribosome-binding sites. The structural basis of the exquisite specificity of metabolite recognition by aptamer folds has emerged from crystallographic structures of purine-sensing aptamers. Both guanine- and adenine-sensing aptamers, which differ significantly in sequence, adopt common tuning-fork-like architectures, in which complementary hairpin loop–loop interactions constrain the tertiary fold of the metabolite-binding pocket, generated through formation of a set of stacked triples. Discrimination between purines is based on a network of hydrogen-bonding interactions, such that a single base change associated with Watson–Crick recognition alters specificity. Two newly discovered riboswitches expand the current repertoire of these novel RNA-based gene regulators in new directions. The glucosamine-6-phosphate (GlcN6P) riboswitch functions as a self-cleaving ribozyme upon binding of the metabolite. The glycine riboswitch is unusual in having tandem metabolite-binding domains, which bind this smallest of amino acids in a cooperative manner, thereby establishing pronounced sensitivity to small changes in metabolite concentration.
Structural biology has a key role to play in defining the molecular architectures of the binding pockets that allow discrimination between closely related metabolites. In the longer term, the challenge will be to define the structural linkages between aptamer domains and expression platforms, thereby addressing mechanistic issues related to riboswitch-based control of gene expression.

The molecular machine that is structurally the best defined is the ribosome. Two reviews highlight the role of elongation factors in coordinating the molecular events occurring during the ribosomal cycle and the targeting of ribosomal subunits by clinically relevant antibiotics.

The various states of a ribosomal cycle can be trapped either by specific antibiotics or by non-hydrolyzable GTP analogs. In order to achieve fidelity in the translation process, equilibrium states are separated by irreversible GTP hydrolysis on G-protein elongation factors. The hydrolysis of EF-Tu occurs between the initial selection of charged tRNA and proofreading, preceding full accommodation in the A-site. The hydrolysis of EF-G precedes the simultaneous translocation of the elongated peptidyl-tRNA from the A-site to the P-site, and the deacylated tRNA from the P-site to the E-site. Jakob Nilsson and Poul Nissen review new structural and biochemical data on these two irreversible steps of the ribosome cycle. Their contribution is enlightening because they pinpoint the still largely unknown molecular links. For example, cognate codon-anticodon recognition between aminoacyl-tRNA and mRNA occurs on the 30S particle, but the GTPase stimulating activity is located on the 50S particle. EF-G hydrolysis is sensitive to the presence of a peptidyl chain on the P-site tRNA; how does EF-G sense the presence of the peptidyl chain? The authors remark importantly that these two G proteins operate in opposite ways. Whereas the action of EF-Tu turns off GTPase activity, the opposite is observed for EF-G. Thus, we still need high-resolution crystal structures of functional complexes of ribosomes with the various cofactors, ideally in the pre- and post-hydrolysis states, in order to delineate the coordinated molecular events occurring during a ribosomal cycle.

Thomas Hermann provides a clear and up-to-date overview of the various classes of antibiotics that target primarily the RNA component of the ribosome, with the emphasis on the decoding site, the peptidyl transferase center and the peptide exit tunnel. Crystal structures are now available for a majority of these antibiotics bound to ribosomal subunits, setting the stage for the design of novel semi-synthetic analogs to counter emerging drug-resistant strains. Initial success was achieved for the polycationic aminoglycoside antibiotics, which target the bacterial decoding site and lock a pair of key adenosines in a looped-out conformation. This has set the stage for the synthetic design of hybrid ligands that can span the decoding site and an adjacent binding site for the antibiotic hygromycin. The RNA component of the peptidyl transferase center is targeted by a chemically diverse set of semi-synthetic and natural antibiotics, which range from the recently developed and clinically approved oxazolidones to macrocyclic streptogramins, which trigger conformational changes. The cyclic 14- to 16-membered macrolide antibiotics function by blocking the progression of the nascent peptide chain through the exit tunnel. The antibiotic erythromycin is a member of this class, as is tylosin, which functions through the formation of a reversible covalent bond, and troleandomycin, which triggers a protein conformational change, thereby remodeling the exit tunnel geometry and interfering with a gating switch. Future challenges will harness this wealth of structural information to improve our current understanding of drug selectivity and develop approaches to overcoming drug resistance, while taking into account the conformational flexibility of the antibiotics and the dynamic nature of the ribosome as it proceeds through its catalytic cycle.

The underlying theme of all the reviews is the molecular recognition of RNA. The final review, by Traci Tanaka Hall, addresses aspects of molecular recognition from the perspective of zinc fingers, versatile modules that target both DNA and RNA.

Zinc fingers are small globular domains containing zinc-coordinated CCHH, CCCH and CCCC elements. A large database of information is available on how multiple zinc fingers target DNA, with the best-defined example associated with regulation of the expression of the 5S rRNA gene by the nine CCHH-coordinated zinc fingers of TFIIIA. Several zinc fingers are involved in recognition of the DNA major groove, whereas others function as spacer elements by traversing the minor groove. By contrast, fingers 4–6 of TFIIIA use different recognition principles to target a minimal 5S RNA construct. All three fingers are involved in molecular recognition of the 5S RNA, with fingers 4 and 6 targeting extrahelical bases, whereas finger 5 is bound less tightly. Mutation studies highlight the importance of the relative positioning of the zinc finger elements in the 5S RNA complex. The protein tristetraprolin (TTP) contains a pair of tandem CCCH zinc fingers that target adjacent UAUU sites. These AU-rich elements are found in the 3′-untranslated regions of tumor necrosis factor and granulocyte-macrophage colony stimulating factor mRNAs, with TTP binding inducing mRNA degradation. Remarkably, sequence-specific modular recognition involves intermolecular hydrogen bonding between protein mainchain atoms and the Watson–Crick edges of the RNA. These examples highlight how zinc fingers can be targeted to RNA scaffolds and, in the longer term, attempts will undoubtedly be made to design zinc fingers targeted to specific RNA topologies.
This overview highlights the increasingly critical role played by RNA in cellular regulation. Structural biology approaches have provided unique insights into scaffold topology, recognition principles, and the fundamentals of RNA-mediated catalysis and gene silencing. The functional switch between two RNA conformers upon the addition of a small cofactor, as in a riboswitch, is based on very precise recognition between the RNA cavities and the small molecule. It is now becoming apparent that antibiotics act in a similar fashion. More than 50% of all antibiotics target the ribosome. Recent X-ray data have shown that most antibiotics bind to the RNA elements of the ribosome. Furthermore, the biological effects induced by antibiotics often come from antibiotic binding to RNA, which alters the equilibrium between alternative states. The new structural data have yielded extensive views on atomic recognition between RNA and small molecules. However, we still do not understand the physico-chemical parameters controlling the adoption of alternative RNA conformers. Single-molecule biophysics is providing precious information on the energetics of such equilibria and how they can be influenced by small molecules. The readership can be reassured that the pace of discovery will accelerate as additional biophysical techniques are brought to bear in order to characterize the structure, dynamics and energetics of RNA-mediated processes of increasing complexity, such as that of the spliceosome.