PERSPECTIVES: STRUCTURAL BIOLOGY

SRP—Where the RNA and Membrane Worlds Meet

Peter Walter, Robert Keenan, Uli Schmitz

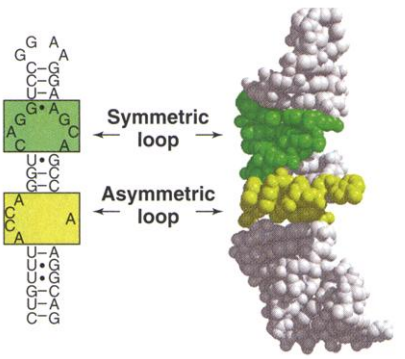
Life is thought to have originated in an RNA world where RNA molecules were responsible for both catalysis and information storage. Historic records of ancient RNA machines are found in the conserved inner workings of today's ribonucleoprotein complexes. In all cells these structures of RNA and protein—which include spliceosomes, ribosomes, and the signal recognition particle (SRP)—mediate essential steps in the processing of genetic information encoded in the DNA and its translation into protein. To perform their myriad tasks, newly synthesized proteins have to be directed to the correct location—to the cell cytoplasm, the plasma membrane, or an extracellular destination. It is the job of the SRP to select proteins destined to be secreted or integrated into the plasma membrane and to target them to the endoplasmic reticulum in eukaryotic cells (or to the plasma membrane in prokaryotic cells).

All cells analyzed so far have SRPs. In its most primitive incarnation (in prokaryotes), the SRP consists of a single protein (termed Ffh) and a small RNA molecule (4.5*S* RNA in *Escherichia coli*). Other cells have SRPs that contain additional proteins and larger RNAs, yet all share the same conserved and, in many cases, functionally interchangeable core. But the mystery remains as to why the SRP requires an RNA component. Clearly, as it is highly conserved within the three kingdoms of life, the RNA subunit is essential. A major step toward understanding this fundamental paradox has been taken by Batey *et al.* (*1*) on page 1232 of this issue. They describe the crystal structure of the conserved ribonucleoprotein core of a bacterial SRP and in so doing cut to the heart of the debate about the significance of the RNA subunit.

The investigators included the M domain (so-called because of the abundance of methionines) of Ffh in their crystals. The M domain of the SRP binds to the signal sequences of newly synthesized peptide chains as they emerge from ribosomes in the cytoplasm. In eukaryotic cells, the SRP then directs the ribosome to the endoplasmic reticulum (ER) membrane by binding to its receptor. The peptide chain is then released into the lumen of the ER and from there is directed to the cell surface where it is either inserted into the plasma membrane or secreted. In prokaryotes, proteins are targeted for transport directly to the plasma membrane by the SRP and SRP receptor. Thus, the M domain performs the first task of the SRP, that is, to distinguish secretory and plasma membrane proteins from those destined to remain in the cell cytoplasm.

The structure of the Ffh M domain was first solved as part of the intact protein in the absence of RNA (*2*). This domain features a large hydrophobic pocket (presumed to be the signal sequence binding pocket) that is flanked by a flexible loop and lined with a large abundance of methionine side chains. The methionines are thought to contribute to the plasticity of the binding pocket (*3*). Owing to the intrinsic flexibility of the unbranched methionine side chains, signal sequences composed of a wide variety of different amino acids can be accommodated as long as they are sufficiently hydrophobic and can adopt an α-helical conformation. Adjacent to the putative signal sequence binding pocket, the M domain folds into a well-ordered small globular structure in which two helices are arranged in a classical helix-turn-helix fold (*2*)—an arrangement found in many DNA binding proteins. This motif was shown by a variety of experimental techniques to form the RNA binding site.

The binding site for Ffh on the 4.5*S* RNA (and on SRP RNA in general) is a ~50-nucleotide stem termed domain IV. It is bounded on one end by a tetraloop and contains two short internal loops with highly conserved sequences (see figure, this page). Indeed, it was the realization of this sequence conservation that first suggested that the SRP is conserved in prokaryotes (*4*). The solution nuclear magnetic resonance structure of domain IV showed that the first symmetric loop (which is most proximal to the tetraloop) is highly ordered and contains a number of unusual base interactions including an interstrand stacking interaction that significantly distorts the RNA backbone. In this structure, the minor groove disappears, creating a distinct flat face on the RNA surface. The second, asymmetric loop, in contrast, is disordered in solution and confers significant conformational flexibility to that region of the RNA (*5*).



Yet meet we shall... The symmetric (dark green) and asymmetric (yellow-green) internal loops of the RNA subunit of the signal recognition particle (SRP) form a continuous surface (right) when interacting with the Ffh M domain of the protein component.

Many of the exciting conclusions of the Batey study stem from the synthesis of structural insights gained from the complex with those gained previously from isolated protein and RNA. The new structure reveals surprising details of the molecular interface where 4.5*S* RNA and Ffh meet. To begin with, it turns out that the M domain structure in the complex can be almost completely superimposed onto that solved for the protein alone. This is in contrast to the RNA subunit, which undergoes major rearrangements. Thus, it is the RNA rather than the protein component, as one—perhaps naïvely—would expect, that changes its conformation upon formation of the SRP complex. The RNA subunit of SRP, therefore, must have a much more profound part to play than simply providing a template to define or refine the protein fold.

In the RNA-protein complex, the asymmetric loop of RNA is held in an ordered and highly unusual conformation with four of the five bases that make up the loop fully exposed on the outside of the RNA molecule. Bases in the symmetric loop also undergo rearrangements relative to the solution structure; in particular, the cytosine at position 62 flips dramatically from an exposed position to bury itself in the RNA-protein interface of the complex. However, the overall distortion of the RNA backbone with its flattened minor groove is rather similar between the free and bound structures. In this way, bases from both loops form a virtually contiguous surface to which the protein binds,

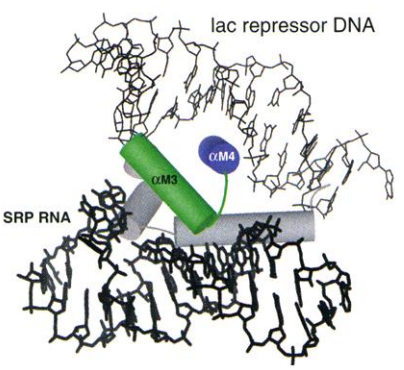
P. Walter is at the Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA 94143, USA. E-mail: walter@cgl.ucsf.edu R. Keenan is at Maxygen, Redwood City, CA 94063, USA. U. Schmitz is at Genelab Technologies, Redwood City, CA 94063, USA.

even though they are separated in the secondary structure by three conventional base pairs (see figure, previous page).

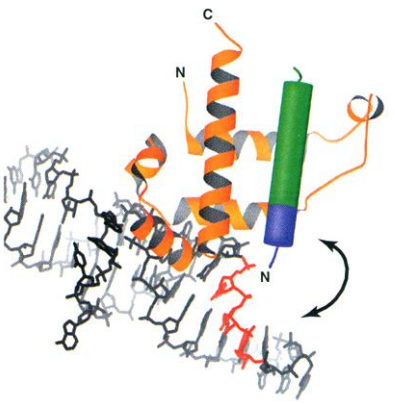
This unusual surface presented to Ffh requires the helix-turn-helix motif to contact the RNA in a manner very different from that observed for the classical helix-turn-helix proteins that bind DNA. These proteins project amino acid side chains from the "recognition helix" (see bottom figure, blue) deep into the major groove of DNA to make base-specific contacts. In contrast, the helix-turn-helix fold of Ffh interacts with the RNA surface through the many contacts of the protein's backbone (see bottom figure, green).

Beyond the differences between the free and protein-bound RNA structures, there are further hints of an intrinsic conformational flexibility in SRP RNA. These are provided by the observation that an extensive network of highly ordered water molecules, and magnesium and potassium ions, contributes substantially to the structure that RNA assumes in the complex, as well as to the RNA-protein interface. We can speculate that using such "wet," or hydrated, macromolecular interfaces containing covalently unrestrained moieties allows for large degrees of structural plasticity. Through rearrangement of water molecules and ions, it may be possible to achieve and accommodate different conformational states more readily than is possible with functional groups restrained by the protein or RNA backbone (*6*).

In the crystal structure of the complex, the regions that form the proposed signal sequence binding pocket are disordered, yet the relative position of the binding site can be inferred from superposition with the known domain structure. This marriage of data reveals that the hydrophobic floor of the signal sequence binding pocket extends onto a ledge of RNA backbone formed by a helical segment between the symmetric loop and the terminal tetraloop of domain IV (see top figure, this page; red). This juxtaposition leads to the provocative model that signal sequences may bind to both protein and RNA, their hydrophobic core snuggling into the methionine-lined pocket of the M domain and their positively charged region (which usually precedes the hydrophobic core) contacting the RNA directly. We surmise that such contacts take advantage of the flexibility built into the RNA and RNA-protein interface to trigger a conformational change. This change could be usurped to target the forming peptide chain (and the ribosome in which it sits) to the ER membrane. In this way the RNA would directly and indispensably participate in determining the state of the SRP and possibly its ability to interact with upstream (ribosome) or downstream (the SRP receptor in the ER membrane) components. So, the structure provides us with new hints as to why evolution may have selected these particular molecular design principles for the SRP.



...And part... DNA and RNA binding helix-turn-helix (HTH) proteins use different surfaces within the HTH motif to interact with their corresponding nucleic acid. The RNA of the SRP (thick lines) interacts primarily with the first helix (green) of the Ffh HTH motif, whereas lac repressor DNA (thin lines, shown here modeled onto the Ffh HTH motif) interacts primarily with the "recognition helix" (blue) of the lac repressor.



...And meet again (*8*). A model for the binding of the signal sequences of newly synthesized peptides to the SRP core. A signal sequence (blue and green cylinder) bound in the hydrophobic groove of the M domain of Ffh (orange) positions its positively charged amino (N)-terminal region (blue) close to a region of negatively charged backbone within SRP RNA (red). Binding in this manner may induce a conformational change in the RNA or protein-RNA interface that is an integral step in the targeting of the newly synthesized protein.

The reaction catalyzed by SRP must have been of great importance early in evolution as phylogenetic evidence clearly points to an ancient function for SRP (one that perhaps reaches as far back as the era when RNA ruled the world). For a collection of RNA molecules to have evolved from self-replicating catalysts toward systems of higher complexity, they must have become encapsulated in membranes. Only then could beneficial catalysts be selected for (because, by being sequestered in membrane-bound compartments, these enzymes were assured of benefitting from the products of the reactions they catalyzed). Thus, membranes that surrounded (and hence defined) the first cells would have been an essential feature of the earliest steps in the evolution of life. Membranes, however, pose new problems—lipids self-assemble into bilayers that spontaneously form tight barriers and, therefore, even the most primitive cells must modulate membrane permeability to permit uptake of nutrients and secretion of waste products. But RNA, because it is very hydrophilic, seems entirely unsuited for this purpose (although, surprisingly, RNA molecules that increase ion permeability in vitro have recently been identified) (*7*). One can surmise that the need to invent proteins was strongly driven by the need to build defined hydrophobic structures that when integrated into membranes could confer selective permeability.

The SRP may therefore have evolved as an early means for RNA-based primordial ribosomes to deal with greasy polypeptide chains. A small protein domain (with the ability to bind to a large variety of hydrophobic peptides) intimately linked to an RNA molecule could have evolved to chaperone newly formed membrane proteins into the plasma membrane. This primitive SRP could have developed more specific targeting functions later on. The dynamic features of RNA would then have become crucial, permitting the loading and unloading of the new signal peptide emerging from the ribosome or providing a dynamic link between the signal peptide and the ribosome itself. Thus, the elegant structure described by Batey *et al.* provides sustenance to nurture the debate about how SRP structure relates to its function. Furthermore, it rekindles the excitement generated through understanding the similar patterns of structure-function relationships in many different phylogenetic groups.

References

1. R. T. Batey *et al.*, *Science* **287**, 1232 (2000).
2. R. J. Keenan *et al.*, *Cell* **94**, 181 (1998).
3. H. D. Bernstein *et al.*, *Nature* **340**, 482 (1989).
4. M. A. Poritz *et al.*, *Cell* **55**, 4 (1988).
5. U. Schmitz *et al.*, *RNA* **5**, 1419 (1999); U. Schmitz *et al.*, *Nature Struct. Biol.* **6**, 634 (1999).
6. N. K. Sauter *et al.*, *Nature Struct. Biol.* **5**, 945 (1998); D. T. Gewirth *et al.*, *Nature Struct. Biol.* **2**, 386 (1995); T. Schirmer *et al.*, *Nature* **343**, 140 (1990); W. E. Royer Jr. *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14526 (1996).
7. A. Khvorova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10649 (1999).
8. S. Butler, *Not on Sad Stygian Shore* (1904).