RNA Chaperones and the RNA Folding Problem*

Daniel Herschlag†

From the Department of Biochemistry, Stanford University, Stanford, California 94305-5307

Functional and structural inter-relationships of RNA and proteins in the execution and control of biological processes such as RNA processing, RNA splicing, and translation are increasingly apparent. In this minireview, I present an RNA chaperone hypothesis, which fosters the view that constraints imposed by fundamental problems in the folding of RNA have profoundly influenced the nature of RNA/protein interactions in biology. The origin of this view is outlined as follows. RNA has two fundamental folding problems: a tendency to fold into and become kinetically trapped in alternative conformations and a difficulty in specifying a single tertiary structure that is thermodynamically strongly favored over competing structures. RNA-binding proteins can help solve both RNA folding problems. Nonspecific RNA-binding proteins1 solve the kinetic folding problem in vivo by acting as RNA chaperones that prevent RNA misfolding and resolve misfolded RNAs, thereby ensuring that RNA is accessible for its biological function. In addition, specific RNA-binding proteins can solve the thermodynamic folding problem by stabilizing a specific tertiary structure. The emergence of nonspecific RNA-binding peptides with chaperone-type activities may have been an early step in the transition from the RNA world to the RNA/protein world. Specific RNA-binding proteins may also have RNA chaperone activities that help prevent misfolding of their cognate RNAs. RNA-dependent ATPases may act as RNA chaperones that spatially and temporally control RNA conformational rearrangements.

"RNA chaperone" refers to proteins that aid in RNA folding and is not meant to refer to chaperones made of RNA.2 For clarity, the classical chaperones that aid protein folding are referred to as "protein chaperones." In keeping with the accepted definition of protein chaperones, RNA chaperones are defined as proteins that aid in the process of RNA folding by preventing misfolding or by resolving misfolded species. This is in contrast to proteins that help protein or RNA folding by catalyzing steps along the folding pathway or by stabilizing the final folded protein or RNA structure.2

There are no established examples of RNA chaperones that act in vivo. This hypothesis is presented because the in vitro data reviewed herein provide support for the hypothesis and this view provides a conceptual framework for RNA folding and RNA/protein interactions. The kinetic problem in RNA folding is emphasized, while space constraints have greatly limited discussion of the thermodynamic problem.

The Two Fundamental Folding Problems of RNA

Many of the examples of RNA misfolding in vitro suggest that the inactive or alternative conformer is kinetically trapped such that it does not revert to the active conformation even after long periods of time. Early work showed that several tRNAs were isolated in two conformations, only one of which could be charged by the cognate aminocyl-tRNA synthetase (11–14). An inactive tRNA was stable on the hour time scale in the presence of Mg2+, but was converted to an active conformation upon heating in the presence of Mg2+ (12). These inactive tRNAs apparently adopt stable alternative secondary structures (15–19).

Larger RNAs provide much additional evidence for a kinetic folding problem. For example, in vitro self-splicing reactions of group I introns, which are >200 nucleotides, typically do not proceed to completion. This suggests the presence of kinetically trapped, alternatively folded conformers (see also Refs. 20–26).

The RNA folding problems observed in vitro could be relevant to the in vivo behavior of RNA or could instead arise as an artifact of in vitro handling of RNA, as RNA is typically purified under denaturing conditions and then renatured. A comparison of the primary, secondary, and tertiary structure of RNA and proteins, based in part on an insightful analysis of tRNA structure (27), suggests that the kinetic folding problems described above and additional thermodynamic folding problems are intrinsic to RNA (summarized in Fig. 1 and Table I).

Primary Structure—RNA has a paucity of primary structure diversity compared with proteins, with just 4 side chains instead of 20. Furthermore, the 4 RNA side chains are more similar to one another than the protein side chains. The RNA side chains come in only two "sizes," purine and pyrimidine, and each is a planar group decorated with hydrogen bond donors and acceptors, whereas the protein side chains comprise hydrophobic, hydrophilic, and charged groups of varying sizes and shapes. The dearth of primary structure diversity, or low "information content," of an RNA polymer (relative to a protein polymer) would be expected to render it more difficult for an RNA sequence to specify a unique tertiary structure.

Secondary Structure—The high thermodynamic stability of RNA duplexes is expected to result in kinetic folding problems. The most stable protein α-helices dissociate on the sub-microsecond time scale (28). In contrast, an RNA duplex of 10 base pairs has a half-time for dissociation of ~30 min, and G/C-rich duplexes of 10 base pairs have half-times up to ~100 years at 30°C. Thus, RNA can get stuck in the wrong conformation (Fig. 1). This kinetic problem could prevent a structured RNA from adopting the correct conformation, could prevent access to mRNA, and could even prevent turnover of an RNA subsequent to correct folding.

The potential for alternative folds appears to be a common property of RNAs. Even random RNAs are predicted to have structures with about half of the residues base-paired, consistent with the estimated helical content of randomly associated RNAs (35, 36).

Tertiary Structure—The problem of stable alternative secondary folds is exacerbated by fortuitous tertiary interactions with 2'-hydroxyls, phosphoryl groups, and metal ions and by the formation of nonstandard base/base interactions that can further stabilize incorrect RNA conformers. Even after RNA adopts the correct secondary structure, it is not yet "out of the woods." The low information content of RNA primary structure is further decreased by sequestering the base-pairing faces of residues in the interior of duplexed regions, while the side chains of proteins face outward in α-helices and β-sheets. Each RNA secondary structure element thus has a strong resemblance to others, so that RNA can have a difficult time specifying a unique tertiary structure. For example, a duplex of the Tetrahymena group I ribozyme docks into tertiary interactions incorrectly approximately 1/1000 of the time, and mu-

---

* This minireview will be reprinted in the 1995 Minireview Compendium, which will be available in December, 1995. This work was supported by National Institutes of Health Grant GM40924.

† Lucille P. Markey Scholar in Biomedical Sciences and a Searle Scholar.

‡ Non-specific is used for simplicity, even though there is presumably no truly nonspecific RNA-binding protein. It refers to RNA-binding proteins with low or wide binding specificities. However, RNA-binding proteins, even those that bind a particular target RNA in vivo, bind other RNAs with reasonably high affinity. The difference between specific and nonspecific or widely specific proteins is quantitative rather than qualitative, so that an absolute distinction is not possible.

1 The term RNA chaperone is already in use by some in the field (1–8). I suggest that only proteins with demonstrated biological roles as chaperones in RNA folding be referred to as RNA chaperones, while the ability to facilitate folding be referred to as "RNA chaperone activity." The definition of RNA chaperones is further honed in the text, and some possible ambiguities are addressed. For example, specific RNA-binding proteins can exhibit biological or nonbiological RNA chaperone activity in the folding of cognate or noncognate RNAs (see Fig. 2 and text). I suggest that these proteins not be referred to as RNA chaperones, in deference to their other functions.
molecular chaperones full circle, as early speculation about the involvement of chaperones in protein folding, which is now well established, was framed by analogy to the ability of single-strand nucleic acid-binding proteins to catalyze nucleic acid duplex formation: both facilitate correct folding by preventing misfolding (42).

The energetics of RNA chaperone action are depicted schematically in Fig. 1A, and one physical model is portrayed in Fig. 1B.

Several recent experiments strongly support such an in vitro RNA chaperone activity of RNA-binding proteins. Slow physical steps in the reaction of a hammerhead ribozyme limit turnover and specificity (43) so this system provides an intermolecular model for the kinetic problems in RNA folding, i.e., dissociation of intermolecular duplexes is crucial for turnover and for discrimination against incorrect (mis-paired) RNA substrates (44). This can be likened to the unraveling of RNA that have adopted incorrect, and ideally structures during folding. Proteins such as the NC protein from HIV-1 and the hnrNP A1 protein were shown to facilitate these physical steps and thereby enhance the ribozyme reaction. In addition, the NC protein resolved a kinetically trapped misfolded complex with HH16 (4, 6, 46, 47).

As mentioned above, the self-splicing of group I introns in vitro is often slow and inefficient, whereas splicing in vivo appears to be fast and efficient (48). In some cases, proteins facilitate splicing in vivo by binding specifically to and stabilizing the catalytically active conformation of the intron (49–51). In contrast, the E. coli S12 ribosomal protein facilitates proper folding of group I introns by nonspecific binding, suggesting a second mechanism for aiding group I intron splicing in vivo (5). Characterization of the S12 protein facilitation further strengthened the analogy between RNA chaperones and protein chaperones. (i) The S12 protein shows no preferential binding to group I introns over exons or other RNAs, suggesting that S12 does not act by specifically stabilizing the intron's catalytic conformation. (ii) The S12 protein is also able to facilitate a hammerhead ribozyme reaction, analogous to the NC and hnrNP A1 proteins, further suggesting a nonspecific rather than specific mode of action. (iii) The S12 protein promotes splicing of a population of kinetically trapped, unreactive precursor RNA, suggesting an ability to resolve misfolded RNAs. (iv) Protein chaperones function solely during a folding step and are not present in the final active species. The same stimulatory effect on group I self-splicing was observed whether or not S12 was removed by protease, prior to initiation of the self-splicing reaction. Thus, the S12 protein is required solely for folding and can act as a true chaperone.

Why a Specific RNA-binding Protein Would Also Act As An RNA Chaperone “Preassociation” Binding Mechanism—Even if a specific RNA-binding protein solves RNA’s thermodynamic folding problem by stabilizing the correct RNA conformation, kinetic problems of attaining that conformation remain, as depicted in the bottom pathway of Fig. 2. (i) The RNA can be kinetically trapped in misfolded conformations (k\text{trap}), and (ii) the correctly folded RNA may lack the thermodynamic stability to exist long enough to be trapped efficiently by its cognate protein (k\text{trap} versus k\text{fold}). The ability of the S12 ribosomal protein to act as an RNA chaperone in the folding of group I introns (5) raises the intriguing possibility that this chaperone activity also solves these kinetic folding problems. Both problems could be avoided by following a preassociation binding mechanism (Fig. 2, top pathway), in which the protein initially uses nonspecific interactions and/or a subset of specific interactions to bind the unfolded RNA and prevent misfolding. The high levels of nonspecific binding exhibited by many specific RNA-binding proteins could allow this chaperone activity. Subsequently, the RNA might undergo conformational rearrangements within the
RNA Chaperones and the RNA Folding Problem

RNA chaperones facilitate proper folding of its cognate RNA. The nonspecific and specific interactions are represented schematically by the absence and presence, respectively, of charge and shape complementarity within the complexes.

The Analogy between Protein and RNA Chaperones—Protein chaperones have suggested that they facilitate the process of protein folding by preventing misfolding (52, 53). As described above, proteins also appear to facilitate RNA folding by preventing misfolding. The following comparisons between RNA and protein chaperones may help elucidate properties that are unique to RNA chaperones as well as concepts that are fundamental to both RNA and protein chaperones.

Protein chaperones appear to be a distinct class of molecules designed to facilitate protein folding. In contrast, the significant extent of nonspecific binding by RNA-binding proteins suggests that many RNA-binding proteins may exhibit RNA chaperone activity in vitro. Over 20 different proteins from E. coli extracts were able to facilitate group I intron splicing in vitro (5), but it is not known which proteins, if any, act as cellular RNA chaperones. The hnrRNPs proteins, which coat pre-mRNA as it is transcribed, represent the most obvious candidate class for cellular RNA chaperones (see also Ref. 54).

Protein chaperones facilitate folding but do not remain bound to the final folded protein product, whereas RNA chaperones may facilitate the folding process and subsequently remain bound because of high levels of nonspecific binding affinity. This may represent a basic difference in the primary recognition element for the two classes of chaperones. Protein chaperones appear to recognize unfolded proteins because of exposed hydrophobic residues; when the protein folds and these residues are buried, the chaperone no longer binds strongly (52). In contrast, the charged phosphodiester backbone and their bases are likely to be at least partially exposed in unfolded or misfolded RNA, allowing nonspecific binding, especially via electrostatic interactions.

Mechanistic studies of protein chaperones have suggested that they prevent misfolding by sequestering unfolded forms so that they cannot aggregate (52, 55, 56). RNA chaperones may act similarly by binding to regions of an RNA and preventing or slowing formation of certain intramolecular structures. The RNA chaperones have also been shown to resolve RNAs that have already misfolded (see above), whereas the protein chaperones that have been best characterized can bind and sequester unfolded proteins but appear unable to bind efficiently to and resolve protein aggregates. The high nonspecific binding activity of RNA-binding proteins may account for this difference by allowing RNA chaperones to bind and subsequently to resolve misfolded RNA conformers (Fig. 1B). However, recent in vivo characterization of the Hsp104 protein has suggested that it actively resolubilizes protein aggregates (57, 58), although the molecular basis for this is not known.

There are proteins other than the chaperones referred to above that aid proper protein folding such as prolyl isomerases and protein disulfide-isomerase (55). Specific RNA-binding proteins can exhibit RNA chaperone activity by helping to prevent and resolve misfolding of both cognate and noncognate RNA (Fig. 2). Specific RNA-binding proteins could also aid the process of folding for the cognate RNA by acting as “guides” in the folding process, i.e. by trapping correctly folded domains or subdomains to help bias the RNA to follow along the folding path toward the final correctly folded structure. In addition, protein/protein interactions can bring together two RNAs (or two regions of one RNA), thereby increasing the probability of duplex formation or other interactions.

Proteins that do this might be referred to as matchmakers, rather than chaperones (7). Such proteins may be involved in spliceosome assembly. There is evidence that the hnrRNP A1 protein can act as both a chaperone and matchmaker (2, 6, 7). RNA chaperones, matchmakers, and guides each can increase the observed rate of RNA/RNA assembly, so that it often may be difficult to distinguish these mechanistically.

RNAs could also act as RNA chaperones to assist in the folding of other RNAs. “Facilitators” are RNAs that base-pair to a ribozyme adjacent to the substrate (61); they presumably prevent the ribozyme from folding up upon itself, thereby increasing access for base-pairing to the substrate. This is analogous to the facilitation of duplex formation by single-strand binding proteins. There are several examples of intramolecular changes that either introduce or resolve problems in folding of an RNA (e.g. Refs. 62 and 63). This might be likened to the role of the presence in reducing a kinetic barrier to the folding of certain bacterial RNAPs (32).

Extending the RNA/Protein Folding Analogy to RNA-dependent ATPases—Most or all of the known protein chaperones use ATP (52), in contrast to the RNA chaperones discussed above. The RNA-dependent ATPases (Rd-ATPases), which constitute a large family of proteins (64),6 may be more akin to the protein chaperones as they use the energy of ATP hydrolysis to facilitate structural transitions. However, despite this gross similarity, there appear to be mechanistic distinctions. The GroEL/GroES chaperonin appears to use ATP to stochastically move the ribozyme and allow the unfolded protein the opportunity to fold in solution versus sequestering it to prevent aggregation with other unfolded proteins (65, 53, 66). In contrast, Rd-ATPases are presumably more akin to helicases, using ATP to disrupt duplex and other structured regions in vivo in the folding and unfolding of RNA.

The use of energy by Rd-ATPases could also allow RNA folding and unfolding steps to be integrated and regulated within complex biological phenomena. For example, an Rd-ATPase may be used to dissociate the U4-U6 snRNP complex at just the right time in spliceosomal assembly, facilitating assembly of a catalytically active spliceosome and/or preventing inappropriate or premature splicing (68). U4 may act as an RNA chaperone made of that prevents misfolding of U6.) Rd-ATPases could also help select between alternative splice sites and prevent inaccurate splicing via a proofreading function that limits the time allotted for individual assembly and catalytic steps (69). Rd-ATPases have also been implicated in ribosomal assembly and translational initiation.

An Evolutionary Perspective

The above ideas can be placed within a unifying but speculative evolutionary context in which an early step in the transition from the RNA world to the RNA/protein world was the emergence of nonspecific RNA-binding peptides with chaperone-type activities. These peptides could have provided a selective advantage in a primitive RNA-dominated world by rescuing RNAs from kinetic traps, aiding in the structural transition of a postreplicative duplex to a folded, functional single-stranded RNA, and helping RNAs more broadly explore structural alternatives. The appearance of a functional nonspecific RNA-binding peptide is expected to be more probable than the appearance of a specific RNA binder because there are more solutions to the problem of nonspecific binding and because a non-

The term RNA-dependent ATPase is used instead of the more common term RNA helicase, as it is likely that some Rd-ATPases facilitate structural transitions other than duplex disruption. The Rd-ATPases have been identified predominantly based on sequence similarities; there is little mechanistic information available concerning helicase or other activities.

6 The term RNA-dependent ATPase is used instead of the more common term RNA helicase, as it is likely that some Rd-ATPases facilitate structural transitions other than duplex disruption. The Rd-ATPases have been identified predominantly based on sequence similarities; there is little mechanistic information available concerning helicase or other activities.
specific binder would have many potential functional targets.

Later in evolution, the problems in folding RNA could have been parlayed into new opportunities for biological systems via cooperation between RNA and proteins, with nonspecific RNA-binding proteins with RNA chaperone functions developing binding preferences and ATP-dependent activities for use in control and regulation. For example, the hnrnp A1 protein has RNA chaperone activity (6, 7, 46) and also appears to be involved in splice site selection (70), and the NC protein from HIV has chaperone activity and appears to bind viral RNA specifically during packaging (4).

Orcistraion of RNA Chaperone Activity in Vivo

The nonspecific RNA-binding proteins that enhance RNA function can also shut down RNA function at higher concentrations, so that there is a limited “window of opportunity” for each protein to be functional (1, 6, 7). How then can a cell orchestrate the function of a large subset of such proteins amidst a pool of near-random RNA without merely binding to and obscuring the function of a large subset of the RNAs? How are the chaperones removed to allow the RNA to function? How does a specific RNA-binding protein find its cognate RNA?

The answers to these questions are not known. Although the concentrations of the various RNA and protein components and their affinities can be regulated to influence RNA processing and function (e.g., Refs. 31, 71, 72), it is not clear that affinities can be tuned and concentrations regulated precisely enough to fully avoid problems of inappropriate RNA/protein pairings and proteins obscuring RNA function. Higher order temporal and spatial cellular organization could be used to avoid these problems and to integrate RNA/protein interactions into other cellular processes. RNA could be “handed off” from one protein to another, with hnrnp proteins binding pre-mRNA as it is transcribed, perhaps followed by slower binding of more strongly bound proteins and other bacterial operons that are regulated via a trp

Minireview: RNA Chaperones and the RNA Folding Problem

Orchestration of RNA Chaperone Activity in Vivo

The nonspecific RNA-binding proteins that enhance RNA function can also shut down RNA function at higher concentrations, so that there is a limited “window of opportunity” for each protein to be functional (1, 6, 7). How then can a cell orchestrate the function of a large subset of such proteins amidst a pool of near-random RNA without merely binding to and obscuring the function of a large subset of the RNAs? How are the chaperones removed to allow the RNA to function? How does a specific RNA-binding protein find its cognate RNA?

The answers to these questions are not known. Although the concentrations of the various RNA and protein components and their affinities can be regulated to influence RNA processing and function (e.g., Refs. 31, 71, 72), it is not clear that affinities can be tuned and concentrations regulated precisely enough to fully avoid problems of inappropriate RNA/protein pairings and proteins obscuring RNA function. Higher order temporal and spatial cellular organization could be used to avoid these problems and to integrate RNA/protein interactions into other cellular processes. RNA could be “handed off” from one protein to another, with hnrnp proteins binding pre-mRNA as it is transcribed, perhaps followed by slower binding of more strongly bound proteins and other bacterial operons that are regulated via a trp operon.

Minireview: RNA Chaperones and the RNA Folding Problem

Orchestration of RNA Chaperone Activity in Vivo

The nonspecific RNA-binding proteins that enhance RNA function can also shut down RNA function at higher concentrations, so that there is a limited “window of opportunity” for each protein to be functional (1, 6, 7). How then can a cell orchestrate the function of a large subset of such proteins amidst a pool of near-random RNA without merely binding to and obscuring the function of a large subset of the RNAs? How are the chaperones removed to allow the RNA to function? How does a specific RNA-binding protein find its cognate RNA?

The answers to these questions are not known. Although the concentrations of the various RNA and protein components and their affinities can be regulated to influence RNA processing and function (e.g., Refs. 31, 71, 72), it is not clear that affinities can be tuned and concentrations regulated precisely enough to fully avoid problems of inappropriate RNA/protein pairings and proteins obscuring RNA function. Higher order temporal and spatial cellular organization could be used to avoid these problems and to integrate RNA/protein interactions into other cellular processes. RNA could be “handed off” from one protein to another, with hnrnp proteins binding pre-mRNA as it is transcribed, perhaps followed by slower binding of more strongly bound proteins and other bacterial operons that are regulated via a trp operon.

Minireview: RNA Chaperones and the RNA Folding Problem

Orchestration of RNA Chaperone Activity in Vivo

The nonspecific RNA-binding proteins that enhance RNA function can also shut down RNA function at higher concentrations, so that there is a limited “window of opportunity” for each protein to be functional (1, 6, 7). How then can a cell orchestrate the function of a large subset of such proteins amidst a pool of near-random RNA without merely binding to and obscuring the function of a large subset of the RNAs? How are the chaperones removed to allow the RNA to function? How does a specific RNA-binding protein find its cognate RNA?

The answers to these questions are not known. Although the concentrations of the various RNA and protein components and their affinities can be regulated to influence RNA processing and function (e.g., Refs. 31, 71, 72), it is not clear that affinities can be tuned and concentrations regulated precisely enough to fully avoid problems of inappropriate RNA/protein pairings and proteins obscuring RNA function. Higher order temporal and spatial cellular organization could be used to avoid these problems and to integrate RNA/protein interactions into other cellular processes. RNA could be “handed off” from one protein to another, with hnrnp proteins binding pre-mRNA as it is transcribed, perhaps followed by slower binding of more strongly bound proteins and other bacterial operons that are regulated via a trp operon.

Minireview: RNA Chaperones and the RNA Folding Problem

Orchestration of RNA Chaperone Activity in Vivo

The nonspecific RNA-binding proteins that enhance RNA function can also shut down RNA function at higher concentrations, so that there is a limited “window of opportunity” for each protein to be functional (1, 6, 7). How then can a cell orchestrate the function of a large subset of such proteins amidst a pool of near-random RNA without merely binding to and obscuring the function of a large subset of the RNAs? How are the chaperones removed to allow the RNA to function? How does a specific RNA-binding protein find its cognate RNA?

The answers to these questions are not known. Although the concentrations of the various RNA and protein components and their affinities can be regulated to influence RNA processing and function (e.g., Refs. 31, 71, 72), it is not clear that affinities can be tuned and concentrations regulated precisely enough to fully avoid problems of inappropriate RNA/protein pairings and proteins obscuring RNA function. Higher order temporal and spatial cellular organization could be used to avoid these problems and to integrate RNA/protein interactions into other cellular processes. RNA could be “handed off” from one protein to another, with hnrnp proteins binding pre-mRNA as it is transcribed, perhaps followed by slower binding of more strongly bound proteins and other bacterial operons that are regulated via a trp operon.