

RNA misfolding and the action of chaperones

Rick Russell

Department of Chemistry and Biochemistry, The Institute For Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Properties of RNA that contribute to misfolding
 - 3.1. Stability of RNA secondary structure
 - 3.2. Independent stability of RNA tertiary contacts
 - 3.3. Difficulty in switching between topologies
4. Experimental studies on folding and misfolding of RNA
 - 4.1. Early RNA folding studies: tRNA
 - 4.2. Small RNAs: 5S RNA, small regulatory RNAs, and ribozymes
 - 4.3. Group I intron RNAs
 - 4.4. RNase P RNA
 - 4.5. The 30S ribosomal subunit
5. Chaperone activity of non-specific RNA-binding proteins
 - 5.1. hnRNP proteins
 - 5.2. Viral nucleocapsid proteins
 - 5.3. Ribosomal proteins
 - 5.4. StpA: a bacterial histone-like protein
 - 5.5. Cold-shock proteins
 - 5.6. Hfq: a bacterial Sm-like protein
6. RNA chaperone function and mechanism of DExD/H-box proteins
 - 6.1. Repertoire of structure-disruption activities by DExD/H-box proteins
 - 6.2. Evidence for general RNA chaperone activity
 - 6.3. Mechanism of DExD/H-box proteins as general chaperones
7. Summary and perspective
8. Acknowledgments
9. References

1. ABSTRACT

RNA folds to a myriad of three-dimensional structures and performs an equally diverse set of functions. The ability of RNA to fold and function *in vivo* is all the more remarkable because, *in vitro*, RNA has been shown to have a strong propensity to adopt misfolded, non-functional conformations. A principal factor underlying the dominance of RNA misfolding is that local RNA structure can be quite stable even in the absence of enforcing global tertiary structure. This property allows non-native structure to persist, and it also allows native structure to form and stabilize non-native contacts or non-native topology. In recent years it has become clear that one of the central reasons for the apparent disconnect between the capabilities of RNA *in vivo* and its *in vitro* folding properties is the presence of RNA chaperones, which facilitate conformational transitions of RNA and therefore mitigate the deleterious effects of RNA misfolding. Over the past

two decades, it has been demonstrated that several classes of non-specific RNA binding proteins possess profound RNA chaperone activity *in vitro* and when overexpressed *in vivo*, and at least some of these proteins appear to function as chaperones *in vivo*. More recently, it has been shown that certain DExD/H-box proteins function as general chaperones to facilitate folding of group I and group II introns. These proteins are RNA-dependent ATPases and have RNA helicase activity, and are proposed to function by using energy from ATP binding and hydrolysis to disrupt RNA structure and/or to displace proteins from RNA-protein complexes. This review outlines experimental studies that have led to our current understanding of the range of misfolded RNA structures, the physical origins of RNA misfolding, and the functions and mechanisms of putative RNA chaperone proteins.

2. INTRODUCTION

Structured RNAs play essential roles in metabolic processes such as the processing of tRNAs and mRNAs, translation of mRNAs, translocation of proteins into the endoplasmic reticulum, and maintenance of chromosome ends, all of which are carried out by complex enzymes composed of proteins and at least one structured RNA (1-5). RNA also forms catalytic structures like self-splicing introns and a diverse set of smaller ribozymes. Even mRNA, once regarded as little more than ticker tape encoding the sequences of protein, is replete with structured elements in the 5'- and 3'-untranslated regions including riboswitches, internal ribosome entry sites, frameshift-inducing structures, and various protein-recognition elements (6-8). Although some protein recognition is based principally on the sequence of an RNA, in many other cases the RNA elements must adopt specific three-dimensional structures for specific protein binding (8). Further, *in vitro* selection and rational design schemes have expanded the structural and catalytic repertoire of RNA well beyond what has been found in nature (9).

Although there is great diversity in the shapes and sizes of RNA structures, there are clear themes to the types of structure that RNA forms. Structured RNAs typically include a set of short A-form helices (typically <10 base pairs), which together include the majority of nucleotides within the structure. These helices, referred to as elements of secondary structure, are arranged into a higher-order structure by tertiary contacts, many of which are formed by modular contact-forming motifs that can be identified by sequence (10-12). RNA structure is described as being hierarchical because the secondary structure is largely stable in the absence of the tertiary structure, and is largely left unchanged by the formation of tertiary structure (13), although exceptions to this paradigm exist (14).

The requirement that RNAs fold into their functional three-dimensional structures introduces what is termed the 'RNA folding problem', analogous to the 'protein folding problem', which continues to provide challenges for theoreticians and experimentalists after more than 40 years of research. As previously described by Herschlag, the RNA folding problem is usefully divided into two distinct problems, a thermodynamic one and a kinetic one (15). The thermodynamic problem is that in order to populate a single, functional structure at equilibrium, an RNA chain must specify a native structure that is more stable than all other possible structures. It was pointed out shortly after the first tRNA crystal structures were determined that the independent stability of RNA secondary structure may allow structured RNAs to adopt multiple conformations (16), and a corollary of this property is that it may be difficult for RNAs to specify a single structure. The kinetic problem is that the RNA chain must be able to fold to this single structure on a time scale that allows it to function. Numerous experimental studies on a range of RNAs have shown that a key component of this problem is that RNA has a strong propensity to adopt long-lived misfolded conformations, and it is this problem that is the subject of this review.

The notion that proteins may function as RNA chaperones by facilitating folding transitions was first put forth in the 1970s and early 1980s (17-19), well before many of the structured RNAs listed above were even discovered. Early work focused on the ability of a non-specific RNA-binding protein to facilitate conversion of an inactive conformer of a transfer RNA to its active form (19). Since then, a host of proteins that preferentially bind to single-stranded RNA have been studied and shown to facilitate native folding of a range of RNAs. While some of these effects are probably not of physiological significance, it has become increasingly clear that there is a necessity for proteins to facilitate RNA conformational changes *in vivo*. The clearest example of this is the recent demonstration that inactivation of a mitochondrial DEXD/H-box protein in the fungus *Neurospora crassa* gives a cold-sensitive phenotype, with several group I introns failing to splice properly (20). Here, I review experimental work that has led to the conclusion that RNAs are prone to forming long-lived misfolded conformations, and the work that has led to our current understanding of the functions and mechanisms of proteins that act as RNA chaperones to reduce the deleterious effects of RNA misfolding.

3. PROPERTIES OF RNA THAT CONTRIBUTE TO MISFOLDING

3.1. Stability of RNA secondary structure

Any consideration of RNA folding and misfolding begins naturally with the secondary structure of RNA. It was noted as early as 1975 from an analysis of tRNA structures that structured RNAs differ fundamentally from proteins in that a large fraction of the nucleotides are present in regular base-paired helices, and these helices are stable enough that they form even when individual helices are generated as isolated molecules (16). Since then, this paradigm of extensive and stable secondary structure has proven to be true for a large number of families of structured RNAs. By 1981, the secondary structures of the 16S and 23S rRNAs had been deduced from comparisons of the sequences of multiple organisms (21-23). Approximately half of the nucleotides in these large RNAs form standard Watson-Crick base pairs, and many of the helices are stable enough that they remain formed under conditions that do not support full ribosome structure and activity. Indeed, chemical footprinting experiments under 'inactivating' conditions were used to confirm the accuracy of the secondary structure models (21, 22). More recently, the crystal structures of the ribosomal subunits confirmed that the vast majority of these predicted base pairs are indeed present (24, 25).

The high stability RNA secondary structure is largely independent of all but local interactions, with the stability of each base pair being dominated by the interactions of the two bases with each other and by stacking interactions with the base pairs that are immediately adjacent (26, 27). Indeed, this approximation is accurate enough that the stability of an RNA helix can be

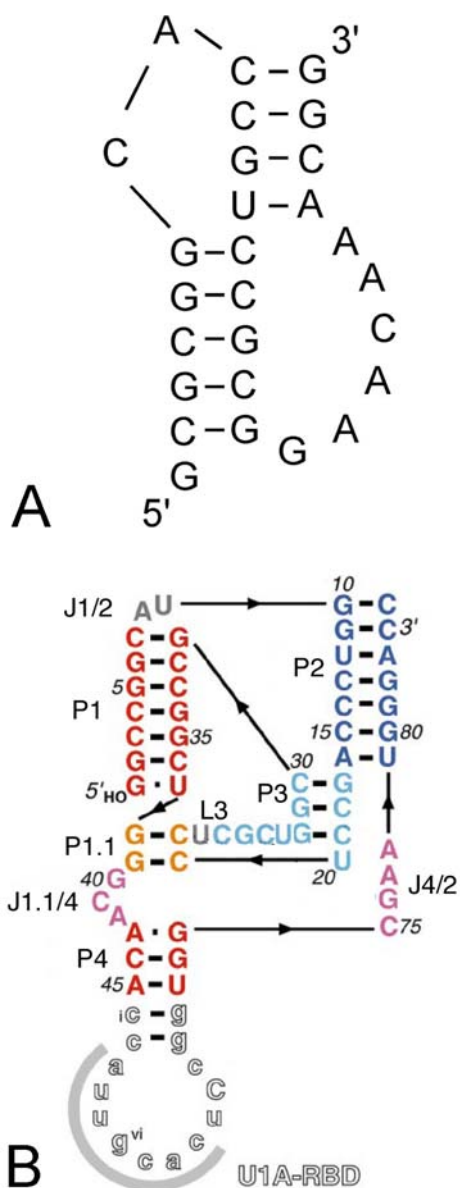


Figure 1. RNA pseudoknots. A, A pseudoknot is formed when loop nucleotides form base pairs with a region outside the loop. Shown is a simple pseudoknot structure from beet western yellow virus (207). B, A double-pseudoknotted RNA structure, the hepatitis delta virus (HDV) ribozyme. Pseudoknots are formed between nucleotides within a hairpin loop and nucleotides outside the hairpin. In the HDV ribozyme, helices P1.1 and P2 are pseudoknots. Outlined nucleotides are non-natural and were introduced to aid in crystallization (reproduced from ref. 208 by permission from Macmillan Publishers Ltd: Nature, copyright 1998).

estimated from its sequence by what has become known as the ‘nearest neighbor’ method (28-31). With a diversity of only four standard nucleotides, an RNA of any significant length has a good chance of having the potential to form at least a few consecutive non-native base pairs. Further,

because of the independence and high stability of RNA secondary structure, non-native helices that do form are often sufficiently long-lived that they interfere with proper folding on a time scale that would present problems for the function of the RNA. Thus, one major origin of misfolding – perhaps the most common one – is the formation of incorrect secondary structure. However, as described below, higher-order components of structure are also likely to play roles, both as origins of misfolding and by providing further stabilization of misfolded species.

3.2. Independent stability of RNA tertiary contacts

The formation of higher-order RNA structure involves the ordering of the secondary elements into defined relative orientations. Much of this ordering is accomplished by formation of contacts between pre-formed elements of secondary structure, and these interactions also have a high degree of energetic independence and stability. Some tertiary contacts are, in fact, composed of standard Watson-Crick base pairs. Known as pseudoknots, these contacts are formed between the terminal loops of two helices or between other single-stranded segments (Figure 1). These interactions, of course, share largely the same thermodynamic parameters as local secondary structure, and are therefore quite stable. Other tertiary contacts are formed by modular structural elements or motifs, which are found in multiple places and in multiple RNAs (12). One of the most common tertiary contact-forming motif is the GNRA tetraloop, which commonly interacts with a ‘receptor’ motif in a neighboring helix to form a tertiary contact (Figure 2). These interactions involve docking of purine bases, most commonly adenine, into the minor groove of the helix, forming an interaction termed an A-minor interaction (32). It has recently been shown that A-minor tertiary contacts can also be formed by motifs within internal loops (33, 34) and are widespread in structured RNAs (33).

For many RNAs, a large fraction of the tertiary contacts form between elements of structure that are found on the surface of the structure and are therefore termed ‘peripheral elements.’ For these RNAs, the tertiary contacts formed from one peripheral element to another and from peripheral elements to the interior of the RNA stabilize the core structure of the RNA. This stabilization from the periphery is quite different from a major strategy of proteins, which are stabilized from the inside by formation of a hydrophobic core.

Analogous to secondary structure, RNA tertiary contacts can form and be stable in the absence of the overall fold of a structured RNA. Independent formation of tertiary contacts is apparent in recent studies that have probed the structures and properties of tetraloop-receptor complexes that are completely removed from any functional RNA (35-37). The independent stability of local tertiary structure is also highlighted by structural studies in which localized regions of the ribosome have been shown to form the same tertiary contacts, and even adopt essentially the same conformations, as they do in the intact ribosomal subunits (24, 25, 38-40), although it should be noted that there are also examples of RNA elements that

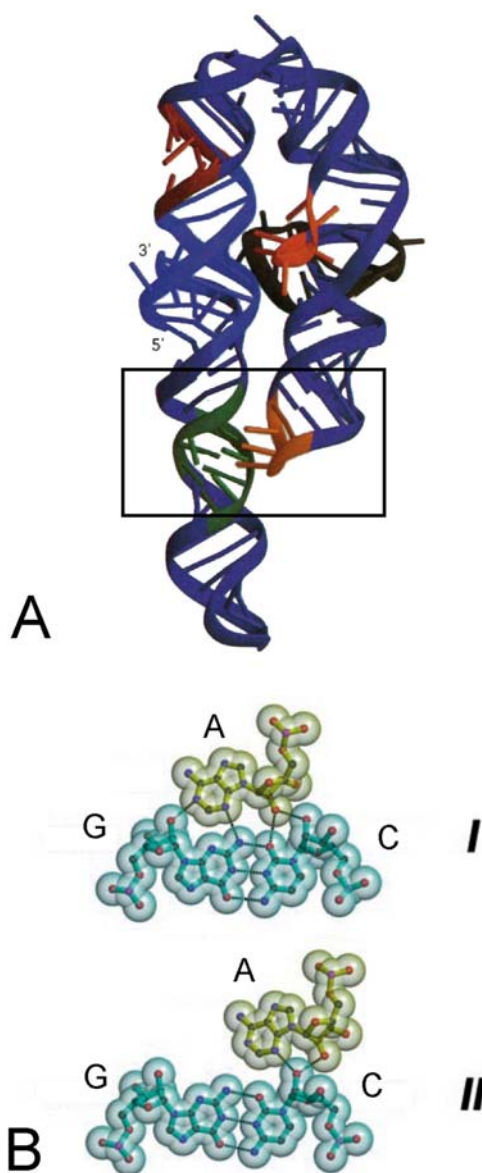


Figure 2. A-minor tertiary interactions. A, Crystal structure of the P4-P6 domain of the *Tetrahymena* group I ribozyme, which includes a tetraloop-receptor interaction (boxed). The GAAA tetraloop sequence (gold) forms tertiary contacts with its receptor sequence (green). The domain also includes an internal loop (orange) that forms A-minor interactions with the adjacent helix. A third helix in the domain (black) forms a tertiary contact composed of base-pairing interactions (the partner was not part of the molecule crystallized and not shown). This panel is reproduced with permission from ref. 209. Copyright 1996, AAAS. B, A-minor tertiary interactions. Adenosine nucleotides form contacts in the minor groove of a helix, most commonly with G-C pairs as shown. The specific contacts for type I and type II interactions are shown. This panel is reprinted from ref. 32 (Copyright 2001 National Academy of Sciences, U.S.A.), and the contacts shown are within the 23S rRNA.

differ substantially in conformation when removed from their natural context (41, 42).

A central reason that isolated tertiary contacts can be stable is probably that the secondary elements are stable and are therefore largely pre-organized. When a tertiary contact is formed between pre-organized secondary elements, there is not a large entropic penalty for structural organization of the secondary elements, as there would be if these elements were not independently stable. This does not imply that the tertiary contacts form without cooperativity, as formation of one tertiary contact might be expected to strengthen others by fixing interacting helices or loops in the correct positions and/or orientations for interaction. Nevertheless, even without cooperative formation of an overall structure, isolated tertiary contacts can be favorable enough to form and accumulate, and thus tertiary contacts can also contribute to the problem of RNA misfolding.

In principle, the ability of tertiary contacts to form independently could allow the formation of non-native tertiary contacts, representing a source of misfolding. However, this type of misfolding has not yet been demonstrated for any RNA. One possibility is that, because the number of tertiary contact-forming regions is relatively small for most structured RNAs and there is significant specificity to the interactions of tertiary partners, non-native tertiary contacts do not represent a major source of misfolding. However, the formation of *native* tertiary contacts has been demonstrated to stabilize misfolded structures (43), presumably because upon formation of the tertiary contact, resolution of the intermediate requires disruption of both the non-native contact that defines the misfolded structure and the stabilizing native contact. Here, the ability of tertiary contacts to form in isolation is clearly a detriment to proper folding because it enables native tertiary contacts to form in the context of existing non-native structure.

3.3. Difficulty in switching between topologies

It has been suggested recently that an additional source of misfolding for RNA may stem from difficulty in obtaining the native topology, where ‘topology’ refers to the spatial arrangement of single- and double-stranded structural elements that cross each other in the context of a folded RNA [(44); see section 4.3 and Figure 4C and 4D). If an incorrect topology is adopted during folding, its resolution can require extensive unfolding. The formation of native tertiary contacts would also be expected to exacerbate this type of barrier by placing further constraints on the positions of portions of the RNA that must move to allow such a rearrangement.

4. EXPERIMENTAL STUDIES ON FOLDING AND MISFOLDING OF RNA

The history of experimental studies in RNA folding is filled with examples of RNA misfolding, so much so that misfolding has largely dominated the field. Although recent examples have been discovered of RNAs that do not appear to misfold, or at least do not accumulate

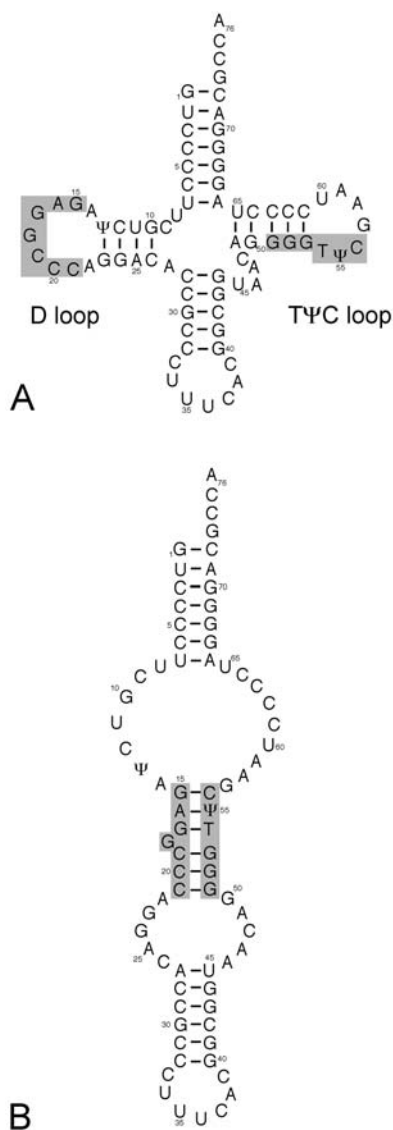


Figure 3. Model for misfolding of tRNA. A, Native secondary structure of *E. coli* tRNA^{Glu}. B, Proposed misfolded secondary structure. Non-native base pairs are formed between nucleotides in the D loop and the TΨC loops, disrupting the corresponding stems. The model is based on data and interpretations in ref. 58.

misfolded species under the conditions they have been studied (45-48), nearly every RNA whose folding has been studied has been found to adopt misfolded conformations at least transiently (49). In this section I review experimental studies that have provided evidence for RNA misfolding, highlighting the types of misfolded structures that have been identified and their lifetimes. The section begins with tRNA, the first RNA that was shown to misfold.

4.1. Early RNA folding studies: tRNA

Studies of tRNA were prominent in development of the early fields of RNA structure, RNA folding and misfolding, and even the influence of chaperones on

folding. At about the same time that the first tRNA sequence was determined and the cloverleaf secondary structure was proposed (50), the groups of Sueoka and Fresco independently reported that tRNAs can adopt two distinct conformations, only one of which is active for aminoacylation (51, 52). The inactive structure, originally referred to as denatured, could be generated by removing or chelating Mg²⁺ ions, and then was long-lived enough that it persisted on the experimental time scale upon subsequent addition of Mg²⁺. An incubation at elevated temperature was shown to restore full activity (> 40 °C), indicating that the RNA had not been irreversibly damaged but rather had folded to an alternative, inactive conformation (52).

A series of subsequent studies of the physical properties of these two conformations showed that the inactive conformation included a somewhat smaller total number of base pairs, was moderately less compact, and was more accessible to binding by certain complementary oligonucleotides (53-55). An extensive change of secondary structure was suggested by the very slow exchange of the inactive and active conformations, and it was suggested that the inactive conformation represented a transition from the cloverleaf secondary structure to an extended secondary structure (56, 57). Although definitive structural studies of the inactive conformation still have not been reported, the conclusion of an extended structure was supported by much more recent work in which chemical and enzymatic footprinting were applied to the inactive conformation (58), and a specific base-pairing pattern was proposed in which nucleotides of the D loop and the T loop form a non-native helix (Figure 3). Although these studies used different tRNAs, the common properties of the inactive conformation suggest a common origin, most likely a significant change in secondary structure.

4.2. Small RNAs: 5S RNA, small regulatory RNAs, and ribozymes

In the years during and immediately after the bulk of the work on tRNAs, interest in RNA structure and folding continued to be centered principally on small RNAs. Crothers and co-workers build on earlier work by Monier by showing that the 5S RNA, a small RNA component of the 50S ribosomal subunit, folds to two distinct conformations that are in slow exchange at physiological temperatures but can be readily exchanged by incubation at elevated temperature (59, 60). Although a specific secondary structure rearrangement was proposed for the exchange of these two structures (61), the nature of the rearrangement was not proven. Other work in this area was on regulatory RNAs in bacteria, principally *cis* elements, that were shown to respond to the nutritional status of the cell by adopting alternative secondary structures, leading to regulated expression of the mRNAs that contain them (62-64). Pioneered by Yanofsky, work in this area showed that exchange between secondary structures can be slow enough that folding is kinetically controlled. More recently, it has been shown that regulation is also achieved by binding of *trans* regulatory RNAs to mRNAs (65), and this area has exploded recently with the discovery of a widespread class of RNA regulators collectively known as microRNAs (66).

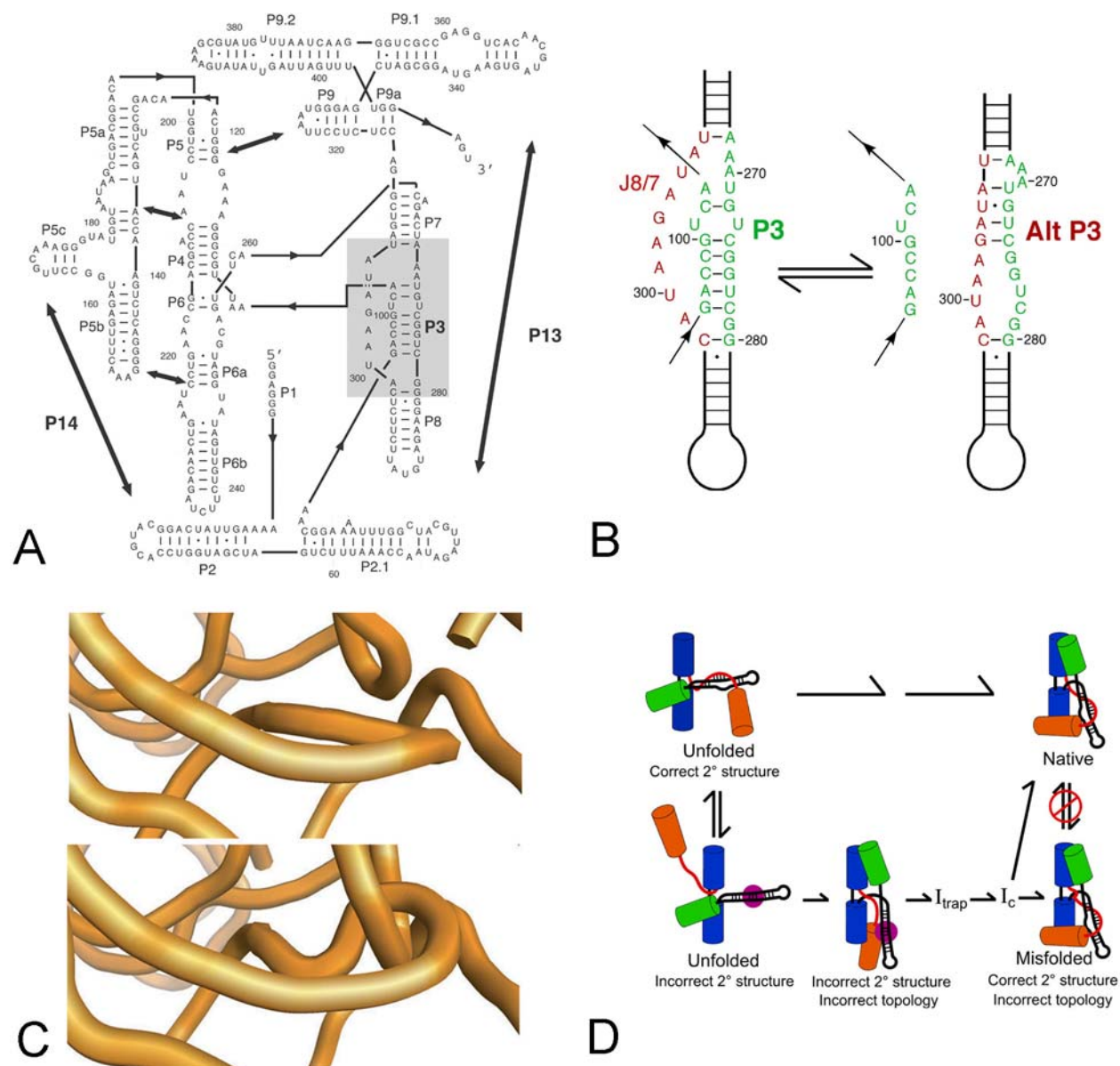


Figure 4. Misfolding of the *Tetrahymena* group I ribozyme. A, Secondary structure of the ribozyme. Long-range tertiary contacts are shown with thick arrows. The P3/altP3 region is highlighted in gray. B, Exchange of the native P3 and non-native alt P3 secondary structures. The structure at the left shows the native P3, as also shown in panel A. The structure at the right shows formation of the non-native alt P3 structure and disruption of P3. C, A topological difference between misfolded and native structures. The top panel shows the native topology within the core of the ribozyme and the bottom panel shows one of several possible incorrect topologies that may be present in the long-lived misfolded structure. D, A model for how a non-native secondary structure can bias an RNA to misfold by giving an incorrect topology, without the non-native secondary structure remaining present in the most stable misfolded form. Formation of the local and non-native structure alt P3 (purple circle) biases the RNA to adopt an incorrect topology early in folding, and this incorrect topology is maintained throughout folding to the misfolded structure even though the secondary structure is exchanged for the native P3 at a later point in folding. Reprinted from ref. 44 with permission from Elsevier.

Several small viral ribozymes have also been the focus of folding studies. The hairpin ribozyme, which adopts a relatively simple structure in which two stem loop structures dock together to form an active site, was observed to give biphasic cleavage kinetics because a fraction of the ribozyme instead adopts an extended

conformation (67, 68). More recent single molecule studies have shown that a more complex mixture of conformational states exist, and that these states fluctuate between docked and undocked conformations with distinct kinetics (69). Additional studies have shown that, in its natural context of a four-helix junction, the importance of this misfolded

conformation is diminished and the ribozyme is considerably more active (70-72). Another small ribozyme, the hepatitis delta virus ribozyme, has also been shown to adopt misfolded conformations, which have been shown through structure mapping and mutagenesis studies to arise from incorrect formation of secondary structure (73, 74).

The recurring theme in misfolding of these small RNAs is that the misfolded and native conformations differ in their secondary structure. Although this could in part reflect the fact that this type of conformational difference is the one most easily probed experimentally (e.g. by site-directed mutagenesis), the importance of secondary structure in misfolding of these RNAs also undoubtedly results in part from their small size. These RNAs have extensive secondary structure but a relatively small number of tertiary contacts, and therefore when they misfold it is most often by forming non-native secondary structures. The larger RNAs, described below, form correspondingly more elaborate tertiary structures and appear to have a more diverse set of misfolded structures and origins of misfolding.

4.3. Group I intron RNAs

Interest in RNA structure was heightened tremendously in 1982, when it was discovered by Cech and co-workers that an RNA intron of a class referred to as 'group I' possessed catalytic activity, splicing itself out of a larger pre-rRNA in the absence of any protein (75, 76). In the nearly 25 years since, this RNA, from the ciliate *Tetrahymena thermophila*, has proven to be a workhorse for studies of RNA structure, function and folding. More recently, studies on folding of other group I introns have appeared, which have broadened our understanding of the complexities of the folding of large RNAs. Indeed, much of what we know about RNA folding processes and misfolded intermediates has come from studies of group I RNAs.

As several groups began to study catalysis by the *Tetrahymena* group I intron and its ribozyme derivatives in detail in the late 1980s and early 1990s, it became apparent that the intron was capable of adopting one or more misfolded structures, which could be resolved by incubating the RNA with Mg^{2+} at elevated temperature (77-79). Physical insight into the nature of one misfolded conformation came from studies of varying the length of the exons; it was shown that a phylogenetically conserved helix can form within the 5'-exon, which competes with the interaction between the 5'-splice site and the 'internal guide sequence' of the intron (80). Subsequent work from Woodson's lab showed that this inhibitory helix is itself in competition with a helix even farther upstream from the intron, and formation of the second upstream helix prevents formation of the inhibitory helix and therefore restores splicing activity (81, 82).

The next few years saw an explosion in the field of RNA folding, and much of the work was on the *Tetrahymena* ribozyme. In pioneering work, Williamson and co-workers developed a time-resolved oligonucleotide hybridization method, in which sequences throughout the ribozyme were monitored for accessibility to a

complementary oligonucleotide at various times following initiation of tertiary folding by addition of Mg^{2+} ions (83). The principal conclusion from initial work was that folding proceeds through at least one stable intermediate. Knowledge of intermediates and their structures was increased tremendously by the development and application of a time-resolved hydroxyl-radical footprinting method to the folding process (84-86). Thus, it became clear that there are several accumulated intermediates, as tertiary structure forms first in one independently folding domain, termed P4-P6, and subsequently spreads throughout the molecule. More recent work from Brenowitz and colleagues has shown that the conclusion of multiple intermediates hold under a range of solution conditions, and even for folding in the presence of monovalent ions alone (87-89).

In addition to discovering folding intermediates, work in the mid- to late-1990s was directed toward determining the roles of these intermediates in folding. Woodson and co-workers used modification-interference, time-resolved footprinting, and an electrophoretic mobility shift assay (EMSA) to determine that the intron adopted a misfolded conformation and that an alternative secondary structure in the core of the intron, termed alt P3, was at least partially responsible for formation of this misfolded intermediate (90, 91) (Figure 4A and 4B). Considerations of the RNA folding process would suggest that formation of the native P3 would be difficult because alt P3, being local in primary sequence, would be imagined to form first, and would then have to be disrupted to form the native P3 base pairs. On the other hand, the alt P3 helix is only four base pairs and would be expected to exchange on the time scale of milliseconds (92), whereas this inactive conformer was resolved to the native state much more slowly, suggesting that resolution of the misfolded intermediate required more extensive rearrangements than just exchange of this alternative secondary structure. Further work from the Williamson lab showed that mutations that weaken native tertiary contacts on the periphery of the molecule accelerated the formation of a protected core (43, 93, 94), as did addition of the denaturant urea, suggesting that the non-native structure present in the misfolded structure was stabilized by native tertiary contacts.

Further insight into structural features of misfolded intermediates and their folding transitions became possible from application by Herschlag and co-workers of a quantitative ribozyme activity assay to follow formation of the native ribozyme. It was shown that what had been viewed as a single misfolded intermediate was actually two distinct intermediates that form in succession during folding. The longer-lived of the two could be populated by up to 90% of the ribozyme for a period of hours (95, 96). Subsequent studies then characterized this misfolded species and its re-folding transition. The misfolded species was shown by small angle x-ray scattering to be nearly as compact as the native ribozyme (97). It was further shown to have each of five long-range peripheral contacts formed and to require disruption of all of them during its re-folding to the native state (44). Intriguingly, this study also indicated that the alt P3 base pairs, despite biasing the RNA toward the long-lived

misfolded conformation, are exchanged to the native P3 base pairs in the misfolded conformer, and a similar proposal was made earlier from oligonucleotide hybridization studies (98). It was suggested that all of the native base pairs are formed in the misfolded conformer, and that the barrier between the native and misfolded species is topological, requiring global unfolding of the RNA to allow one strand or segment to pass to the other side of another within the core (Figure 4C and 4D). While this proposal is intriguing, it will require additional tests to establish whether it is correct.

In recent years, studies on folding of a few other group I introns have broadened the conclusion that misfolded intermediates are common, while also offering some insight into the features that make misfolding more likely. Catalytic activity assays of a ribozyme derived from the phage T4 td intron showed that a non-native extension of the P2 helix can form during folding and inhibit substrate cleavage (99), once again demonstrating a non-native secondary structure. Studies of a *Candida albicans* group IE intron, a subgroup that is closely related to the *Tetrahymena* intron, demonstrate that the RNA is prone to misfolding, and in further analogy with the *Tetrahymena* ribozyme, misfolding can be avoided by pre-incubation with monovalent ions (100-102). Recent studies of a very small group I ribozyme from the bacterium *Azoarcus* suggest that the ribozyme largely avoids stable misfolding, and acquires all of its tertiary structure in milliseconds (103, 104). As this ribozyme lacks peripheral structure almost entirely, this result reinforces the notion that peripheral structure can exacerbate the consequences of misfolding by stabilizing misfolded species, and it introduces the idea that with fewer structural elements in a smaller RNA there is less that can go wrong.

4.4. RNase P RNA

At about the same time that RNA catalysis was discovered in the group I intron, it was shown by Altman, Pace, and co-workers that the catalytic component of the bacterial RNase P enzyme is RNA (105). RNase P functions in tRNA processing by cleaving pre-tRNAs to generate their mature 5'-ends (106). In bacteria, the enzyme is composed of one RNA and one protein, whereas in the archaeal and eukaryotic domains the nuclear-encoded enzyme includes RNA and multiple proteins (107, 108). The bacterial RNA consists of two separable domains; the C- (catalytic) domain contains all of the nucleotides that form the active site and the S- (specificity) domain forms some of the contacts with the tRNA substrate (109, 110). Most of the folding studies to date have focused on the bacterial RNase P RNA, as it is able to fold *in vitro* to a native structure and to function in the absence of accessory proteins.

The first hint that RNase P RNA is prone to misfolding came shortly after the delineation of the RNA as the catalytic component, with the finding that catalysis by the RNA displayed a kinetic lag that could be rescued by pre-incubation of the RNA (111). Detailed characterization of the Mg²⁺-dependent folding of the RNA began a decade later (112). Much of the research on folding of RNase P

RNA has been carried out in a long-standing collaboration between the labs of Pan and Sosnick. In 1997, these groups pioneered the use of circular dichroism and absorbance spectroscopy to study RNA folding, reporting that RNase P RNA from *Bacillus subtilis* requires hundreds of seconds to complete its folding, despite rapid formation of at least one intermediate (113). Further, the folding was found to be strongly accelerated by urea, indicating the presence of kinetically trapped species.

In further work, these groups found that an RNA that was engineered such that one of the connections was broken between the C- and S- domains folded much faster than the wild-type RNA, suggesting that the principal barrier for the wild-type RNA is linked to the interdomain connections (114). When expressed as an independent domain, the C-domain folds much more rapidly and is not accelerated by denaturant, suggesting that its folding is not limited by the resolution of a misfolded structure and reinforcing the notion that the slowest step for folding of the full-length RNA involves resolution of a misfolded structure that involves or is dependent on the connections between the two domains (45, 46). Nevertheless, folding of the isolated C-domain remains complex, as subsequent single molecule fluorescence studies revealed the presence of additional intermediates that are populated at equilibrium (115).

In the last few years, a burst of structural information on RNase P RNA has been obtained from determination of crystal structures, first of the isolated S-domain (116), and then of the entire RNase P RNAs from *Bacillus stearothermophilus* (117) and *Thermotoga maritima* (118) [reviewed in (119)]. The structural information is certain to impact research in folding in the coming years and has already been used to infer structural features of a folding intermediate of the S-domain (120).

4.5. The 30S ribosomal subunit

The ribosome is one of the largest and most complex pieces of machinery that the cell must assemble; the three RNAs that make up the bacterial 30S and 50S must fold and assemble with the more than 50 proteins that make up the functional ribosome. With this complexity in mind, it is not at all surprising that assembly involves multiple steps and appears to include misfolded structures. Research into folding and assembly of the 30S ribosomal subunit is more advanced than for the 50S subunit, and much of the research over the last 35 years has been built on the foundation provided by the seminal work of Nomura and co-workers. A series of studies by this group culminated in delineation of an 'assembly map' for 30S formation, which outlined a crude order of stable protein association with the 16S rRNA (121). Even at this early stage, it was clear that the folding and assembly process was complicated kinetically by the formation of an intermediate containing only a subset of the proteins that was blocked at low temperature from further assembly but could be pushed along the assembly route by incubating at elevated temperature (122, 123). It was also shown that active 30S subunits could be inactivated by a transient decrease in Mg²⁺ concentration or removal of monovalent

ions (124, 125), and that the inactive form was altered in its accessibility to chemical footprinting reagents (126), underscoring the ability of the 30S subunit to adopt alternative conformations.

More recently, detailed physical studies have begun to give physical insights into the nature of the rate-limiting steps for folding and assembly of the 30S subunit, and these studies have been aided tremendously by the advent of high-resolution structural information (25). Holmes and Culver used a chemical footprinting approach to characterize the accessibility of the rRNA bases before and after the principal structural rearrangement, as well as before and after binding of the early and late sets of ribosomal proteins (127, 128). These studies suggested that significant rearrangements of the RNA are involved in the high-temperature conformational change, and these rearrangements may include formation of some of the long-range base pairs present in the functional structure. Development of a powerful method for monitoring the kinetics of protein binding based on mass spectrometry has recently provided a higher-resolution view of the assembly process (129), and has suggested that what has been viewed as a single rate-limiting conformational rearrangement may instead be a series of rearrangements in which the binding of various proteins is coupled to RNA structural changes, some of which may represent disruption of non-native interactions, while others may represent formation of contacts. Further, this work suggested that there is not a single order for completion of these folding steps, but rather a landscape that contains multiple possible routes to the functional structure.

5. CHAPERONE ACTIVITY OF NON-SPECIFIC RNA-BINDING PROTEINS

In several cases, outlined below, it has been shown that proteins that bind single-stranded RNA non-specifically display RNA chaperone activity, facilitating conformational rearrangements of RNA from misfolded to active structures or between functional structures. These proteins are typically basic and are often rich in arginine residues. Presumably, their ability to interact preferentially with single-stranded RNA stabilizes folding intermediates in which RNA-RNA contacts are disrupted, and this stabilization accelerates re-folding transitions. A large excess of these proteins is typically required for activity, consistent with models in which the proteins coat the unstructured RNA intermediates with ill-defined stoichiometry. Although chaperone *activity* has been clearly demonstrated, for many of these proteins it remains to be determined whether they actually function as RNA chaperones *in vivo*. Such demonstration may be difficult because cells may express multiple chaperones with overlapping substrate specificity, such that deletion of any one chaperone would not give an observable phenotype. Nonetheless, it remains uncertain whether the chaperone activities reflect the functions of these proteins or whether they are simply a consequence of their physical properties. Below, the experiments are reviewed that led to identification of the major classes of proteins with

chaperone activity, while some more recently-discovered examples are omitted for brevity.

5.1. hnRNP proteins

The earliest studies of RNA chaperone activity were performed not long after the first RNA misfolded species were identified. These early studies focused on the activity of a protein referred to as UP1, a proteolytic fragment of hnRNP A1 protein isolated from calf thymus (17, 19). The hnRNP proteins are present in mammalian cell nuclei and are thought to associate with RNAs during their biogenesis. The groups of Karpel and Fresco found that the UP1 protein fragment accelerated reactivation of yeast tRNA^{Leu} and *E. coli* 5S RNA from their inactive conformations. The activity of UP1 was inhibited by single-stranded RNA, suggesting that the single-stranded binding activity of the protein was involved in its chaperone activity. The A1 hnRNP protein was subsequently characterized as a protein that facilitates renaturation of DNA as well as RNA, presumably by binding single strands and perhaps bringing them together via multiple sites on the protein or by protein multimerization (130-132).

5.2. Viral nucleocapsid protein

As catalytic RNAs were discovered and characterized in the late 1980s and early 1990s, it became possible to explore whether proteins could increase the activities of these catalytic RNAs by acting as chaperones. In the first such study, Herschlag and co-workers showed that the hammerhead ribozyme was accelerated in its oligonucleotide cleavage reaction by the presence of the HIV p7 nucleocapsid protein (NC) (133). Although the hammerhead ribozyme is not one that is known to adopt long-lived misfolded species during folding (see section 4.2 above), its reaction as a multiple-turnover enzyme requires the formation and dissociation of an RNA helix between the oligonucleotide substrate and the ribozyme, and subsequent work showed that the NC protein accelerated the reaction by facilitating both the formation and dissociation of the helix (134).

At the time this work was performed, the physiological roles of the NC protein were only beginning to emerge, but in the years following it became clear that the NC protein also has substantial chaperone activity for several nucleic acid rearrangements that are vital components of the viral life cycle. These rearrangements include dimerization of the viral RNA (135, 136), annealing of the tRNA primer to the genomic RNA to initiate reverse transcription (135, 137), disruption of template structure during reverse transcription (138, 139), and 'strand transfer' of the newly-generated (-) strand DNA to continue reverse transcription (140, 141) [reviewed in (142)]. Thus, it appears highly likely that RNA chaperoning is a central function of the NC protein.

Physical insight into the mechanism of chaperone action has come from an NMR structure of NC in complex with a viral stem-loop structure (143), which shows an extensive set of electrostatic interactions between arginine residues and the backbone of the RNA. Presumably, similar

contacts are made and even strengthened in complexes with single-stranded RNA, and recent work from Karpel's group has provided further evidence for preferential binding to single-stranded RNA (144). Recently, a single molecule fluorescence approach has been applied to the NC-mediated RNA rearrangements, with the finding that the strand-transfer reaction is catalyzed via multiple pathways by the NC protein (145, 146).

5.3. Ribosomal proteins

As many of the ribosomal proteins are intertwined with segments of the rRNAs in the crystal structures of the ribosomal subunits (24, 25, 147), it is perhaps not surprising that many of the ribosomal proteins possess non-specific RNA binding and RNA chaperone activities. RNA chaperone activity was first characterized for the S12 protein, with lesser activity attributed to several other ribosomal proteins, from experiments in which *E. coli* lysates were fractionated based on their ability to facilitate *in vitro* splicing of group I introns from the phage T4 (148). More recently, a systematic survey of the ability of large subunit ribosomal proteins to chaperone folding of one of these introns revealed that approximately one-third of the 34 large subunit proteins have RNA chaperone activity (149). It remains unclear whether any of these proteins function as RNA chaperones *in vivo*.

As a means of testing proteins for RNA chaperone activity *in vivo*, Schroeder and co-workers have implemented a system in which overexpression in *E. coli* of a protein with RNA chaperone activity can rescue self-splicing of a group I intron in the phage T4 *td* (thymidylate synthase) gene, which is otherwise prevented from splicing by formation of a long-lived misfolded structure that includes a non-native secondary structure [(150); reviewed in (151, 152)]. Significant chaperone activity was detected for the S12 protein, as well as for the NC protein (Section 5.2 above) and the StpA protein (Section 5.4 below), and more recently, chaperone activity has been demonstrated for protein components of Ro ribonucleoprotein complexes (153) and has been confirmed by other methods (154). Although this *in vivo* assay does not establish a chaperone function for these proteins, it establishes that they are capable of RNA chaperone activity inside bacterial cells.

5.4. StpA: a bacterial histone-like protein

Like some of the ribosomal proteins, StpA was shown to function as a chaperone for self-splicing of the group I intron in the phage T4 *td* gene. It was isolated as a multi-copy suppressor of defective splicing by Belfort and co-workers (155). StpA is a small (15 kDa) and highly basic protein that functions as a transcriptional regulator and is quite similar to the highly expressed H-NS nucleoid protein (156). Interestingly, despite the similarity of these two proteins, they regulate transcription of distinct sets of genes, and StpA is much more efficient as an RNA chaperone (156). Further work on the mechanism of chaperone activity by StpA has shown that it increases the exposure of the group I intron core to chemical footprinting reagents (157). Further, the efficiency of this structure disruption by StpA is related to the stability of the RNA structure, such that StpA can be inhibitory for splicing of

variant group I RNAs with decreased stability (158), presumably because it remains associated with the partially unfolded RNA or because it unfolds the RNA faster than it can fully re-fold.

5.5. Cold-shock proteins

Upon temperature downshift, global protein expression in bacteria is decreased, but expression of a relatively small number of proteins (17 in *E. coli*) is substantially increased (159). Several of these 'cold-shock' proteins bind RNA non-specifically and are therefore candidates to function as RNA chaperones. As the stability of RNA secondary structure is strongly temperature-dependent, it would be expected that kinetic traps for RNA would be particularly severe at low temperature, and thus there would be a stringent requirement for RNA chaperones under these conditions (although some of these proteins are also expressed at higher temperature and presumably function similarly). One of the most highly expressed proteins, CspA (70 amino acids), was shown to increase the susceptibility of its own mRNA to cleavage by the single-strand specific RNases A and T1, indicating that CspA reduces the secondary structure content of the mRNA (160). Subsequently, CspA and the related proteins CspE and CspC were shown to function as transcription anti-terminators by preventing the formation of a secondary structure element in nascent RNAs that would otherwise lead to termination (161).

Crystal and solution structures of CspA revealed a five-stranded β -barrel structure with an OB-fold (162, 163), and a more recent co-crystal structure has shown that the structurally related *Bacillus subtilis* CspB protein binds single-stranded nucleic acid in an extended conformation along one surface of the protein (164)(Figure 5). This binding mode, with the nucleic acid extended and single-stranded, is consistent with a role as an RNA chaperone. It remains to be determined whether these proteins function principally as regulators of transcription terminators or whether they have more general roles as RNA chaperones.

5.6. Hfq: a bacterial Sm-like protein

The Hfq protein was originally discovered as a factor required for replication of the phage Q-beta (165), and it was subsequently shown to give wide-ranging phenotypes when inactivated (166) and to be involved in the post-transcriptional regulation of a number of genes [reviewed in (167)]. Hfq shares sequence and structure similarity with the eukaryotic Lsm proteins, forming a hexameric ring structure similar to the hetero-heptameric structures of the Lsm proteins, and binds single-stranded RNA around a central basic cleft (168).

Key insight into the mechanism of regulation by Hfq came from Storz and colleagues, who found in 1998 that Hfq interacts with the small regulatory RNA OxyS (169). At the time, OxyS was one of only a handful of well-characterized, bacterial non-coding RNAs, whereas since then the number has grown substantially [reviewed in (170)]. OxyS was known to activate and repress expression of multiple genes, but the mechanisms of its action were unclear. Subsequent work showed that OxyS binds directly

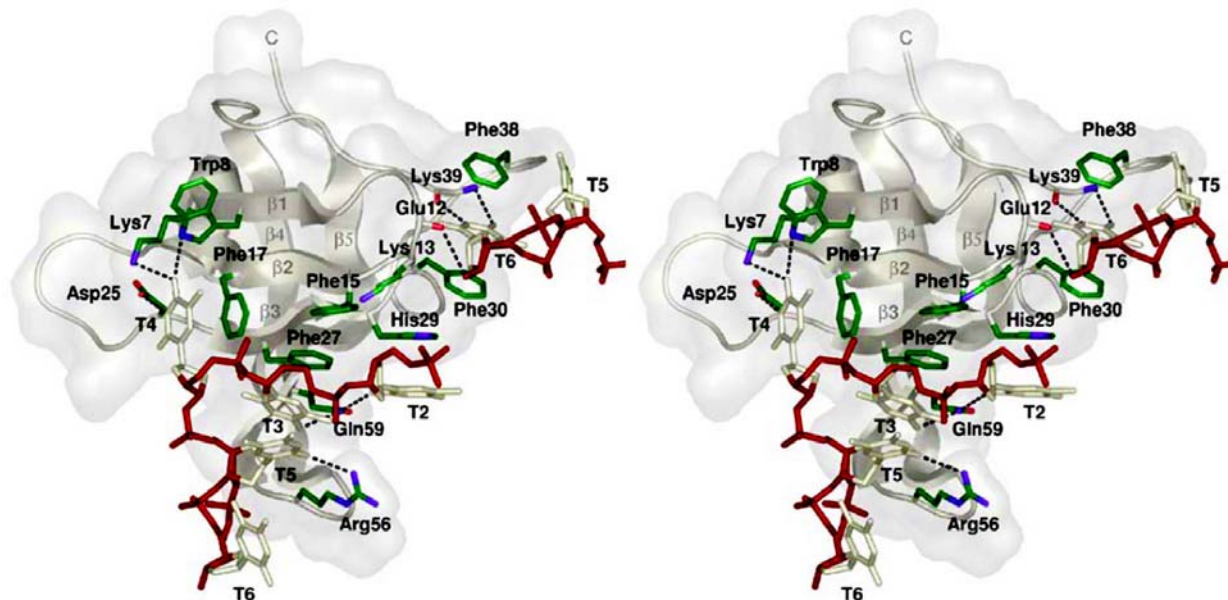


Figure 5. Nucleic acid binding by the cold-shock protein CspB. Two molecular of an oligonucleotide of six thymidine nucleotides (dT6, shown in red) were bound to CspB in the crystal, as shown. The protein is shown as a semi-transparent surface, and carbons of residues on the protein that contact the oligonucleotide are colored green. Reprinted from ref. 164 with permission from Elsevier.

to the mRNAs of at least two genes that it regulates, and that the complex formation of these RNAs depends on the Hfq protein (171). Thus, Hfq was suggested to function as a chaperone to facilitate annealing of the regulatory RNA with its target mRNAs, and evidence for this mode of action has since been accumulated for several additional regulatory RNAs (170, 172). Although it remains likely that Hfq has multiple modes of action, including some that may be unlinked to RNA chaperoning (173), it is clear that facilitating association of RNAs is a central function of the protein.

Although the mechanism by which Hfq facilitates RNA-RNA complex formation remains uncertain, ideas are beginning to emerge from structural and biochemical studies. Focus was initially on the interaction of Hfq with the small regulatory RNAs, but it was subsequently shown that Hfq interacts with the target mRNAs also, raising the possibility that it disrupts structure within the mRNA, allowing access to the regulatory RNA (174). It has recently been suggested that Hfq has distinct interacting surfaces for multiple RNAs and that it can form ternary complexes with two RNAs (175), raising the possibility that, rather than simply facilitating rearrangements by stabilizing single-stranded intermediates, Hfq actively mediates formation of RNA-RNA complexes by facilitating strand displacement reactions (175, 176).

6. RNA CHAPERONE FUNCTION AND MECHANISM OF DEXD/H-BOX PROTEINS

DEXD/H-box proteins are ubiquitous throughout nature and have been implicated in nearly all processes that involve structured RNAs [recently reviewed in (177, 178)]. They are ATPases with significant sequence similarity to

DNA helicases, and are characterized by the presence of several highly conserved sequence motifs, including one of sequence DEXD/H from which they derive their name. At the most general level, DEXD/H-box proteins are thought to function by coupling the energy derived from cycles of ATP binding and hydrolysis to acceleration of RNA structural transitions that would otherwise proceed slowly. A very large number of these proteins appear to function with a high degree of specificity for a particular RNA or RNP substrate, and this specificity is achieved because the DEXD/H-box protein recognizes an RNA or protein that is part of its target complex (179). On the other hand, recent evidence indicates that a subset of DEXD/H-box proteins have more relaxed specificity and apparently function as general RNA chaperones. Unlike the proteins described in the preceding section, which presumably facilitate structural rearrangements of RNA by a relatively passive mechanism in which they bind and stabilize single-stranded intermediates, DEXD/H-box proteins have the potential to actively disrupt contacts in an energy-consuming process. Below, I first review research in which model systems have been used to explore the repertoire of structures that DEXD/H-box proteins are able to disrupt. Then I review recent evidence for the function of some of these proteins as general RNA chaperones and current proposals and evidence for the mechanisms by which they may achieve this general chaperone function.

6.1. Repertoire of structure-disruption activities by DEXD/H-box proteins

Soon after the DEAD box family (a major subset of the DEXD/H-box proteins) was identified by sequence comparisons of several newly-sequenced genes (180), it was shown that one of these proteins, the human p68 protein, was capable of unwinding double-stranded RNA

(181). RNA unwinding had already been suggested for one other DEAD-box protein, the translation initiation factor eIF4A (182, 183), and over the next several years was characterized further for this protein and for several other DExD/H-box proteins (184-189).

However, despite the apparent similarity of the activity with that of DNA helicases, two features of the RNA unwinding activity of the typical DExD/H-box protein are substantially different. First, the processivity of DExD/H-box proteins tends to be very low, such that significant unwinding usually requires a considerable excess of proteins and is only observed for short helices (typically <15 base pairs). Second, whereas most DNA helicases require a single-stranded extension of a defined polarity for efficient loading, some DExD/H-box proteins have been shown not to require a particular directionality for the single-stranded extension (189, 190).

It might be imagined that the low processivity could reflect the genuine function of many DExD/H-box proteins. After all, the structured RNAs that are implicated in the functions of these proteins – precursors of the ribosomal RNAs, the spliceosomal RNAs, and other structured RNAs – have secondary structures that are composed of short helices, not long continuous ones. A protein whose function it is to facilitate rearrangements of these RNAs, or to help them resolve misfolded species, does not need to unwind long, continuous helices. Instead, it may need to disrupt short, non-native helices, and it also may need to disrupt tertiary contacts or even displace proteins. It is interesting in this regard that one DExD/H-box protein that is clearly not ‘typical’, as defined above, but instead has properties closer to those of DNA helicases, is the viral protein NPH-II, which is implicated in viral transcription and may function by unwinding or translocating along long stretches of DNA or RNA (191, 192).

With the idea of defining the catalytic repertoire of DExD/H-box proteins, Jankowsky, Pyle and co-workers tested the ability of these proteins to perform activities other than unwinding simple helices. They first found that the NPH-II protein is capable of displacing the tightly-bound U1A protein from an RNA helix in an ATP-dependent fashion (193). While this displacement presumably occurs during unwinding of the duplex by NPH-II and may result from the unwinding reaction rather than a direct interaction of NPH-II with the protein to be displaced, in subsequent work Jankowsky and co-workers showed that NPH-II can also displace a protein, or even a complex of several proteins, from binding sites on single-stranded RNA, ruling out any involvement of RNA helix unwinding (194). Analogous activities were also observed for the DED1 protein from yeast, which is implicated in translation initiation and behaves in unwinding assays as a more ‘typical’ DExD/H-box protein than NPH-II (188, 195). This group and others have also found that DExD/H-box proteins, including DED1, can facilitate formation of RNA duplexes in addition to unwinding (196, 197), an activity reminiscent of the bacterial Hfq protein. For DED1, the annealing activity was shown to be ATP-independent,

such that the ratio of ATP to ADP defines a steady-state balance between unwinding and annealing activities (196).

6.2. Evidence for general RNA chaperone activity

The first demonstration that a DExD/H-box protein functions as an RNA chaperone came in 2002, when Lambowitz and co-workers showed that mutations in the *Neurospora* CYT-19 DEAD-box protein result in defects in splicing of several mitochondrial group I introns *in vivo* [(20); see also (198)]. Further *in vitro* experiments indicated that CYT-19 binds structured RNAs with little or no specificity for individual group I introns and that it facilitates a structural rearrangement of the related *Tetrahymena* group I intron, indicating that the splicing defect *in vivo* arises from a failure of the introns to fold properly, and that CYT-19 functions as a chaperone in this process.

Subsequent work from the groups of Lambowitz and Perlman suggested further that CYT-19 functions as a broad-specificity or general RNA chaperone, and that a mitochondrial yeast DExD/H-box protein, Mss116p, functions analogously (199). Whereas prior work had detected defects in splicing of only a subset of the group I and group II introns upon mutation of *MSS116* (200), the more recent work gave detectable decreases in splicing for all nine group I introns and all four group II introns (199). Although indirect effects of Mss116p inactivation cannot be excluded, particularly as most of the introns require specific-binding proteins whose expression may be compromised by loss of Mss116p activity, this result and the earlier characterization of the activities of Mss116p (201) suggested that Mss116p functions as a general chaperone for both group I and group II introns. Heterologous expression of CYT-19 in the *MSS-116*-disrupted strain gave partial or complete restoration of splicing for each intron, indicating that CYT-19 is also capable of facilitating folding of a range of structured RNAs that extends beyond group I introns. Further, *in vitro* studies from Lambowitz, Perlman, and co-workers (202) and from my group (203), described below, have directly demonstrated that CYT-19 acts as a chaperone in folding of group I and group II introns and have begun to probe the mechanisms of this chaperone activity.

6.3. Mechanism of DExD/H-box proteins as general chaperones

To explore the mechanism of RNA chaperone activity by CYT-19, my group designed an *in vitro* system to follow re-folding of the long-lived misfolded conformation of the group I RNA from *Tetrahymena* to its native state (203). We found that CYT-19 indeed accelerated re-folding of this misfolded RNA to the native state and could subsequently be removed before the fraction of native ribozyme was determined by activity, indicating that it acts solely to increase the rate of a conformational transition.

Although much about the mechanism by which CYT-19 gives re-folding of this RNA remains to be determined, we made two additional observations that gave insight into the general mechanisms of CYT-19 action.

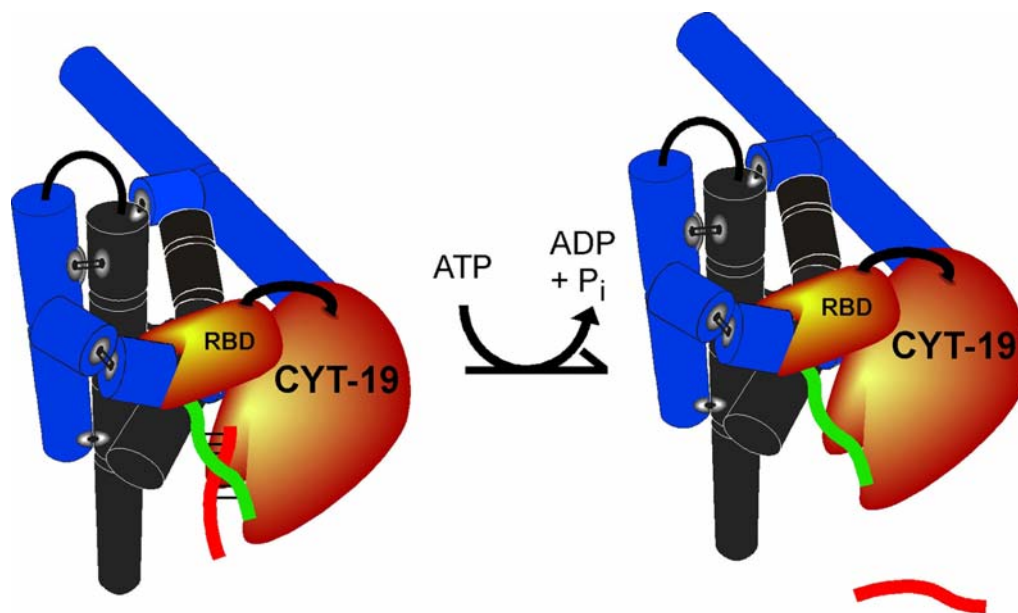


Figure 6. Model for chaperone activity of CYT-19. CYT-19 binds to structured RNA via a site on the protein that is distinct from the site responsible for unwinding helices or performing other structure-disruption activities (labeled RBD). This site remains bound at its attachment point while the ATP-dependent structure-disruption occurs. The figure shows unwinding of the helix of the *Tetrahymena* ribozyme formed between the ribozyme and its oligonucleotide substrate (colored green and red, respectively). The secondary elements of the ribozyme are depicted as black and blue cylinders for core and peripheral elements, respectively. Reprinted from ref. 203 (Copyright 2006 National Academy of Sciences, U.S.A.).

First, we found that CYT-19 could also unwind the six-base-pair helix between the RNA and its oligonucleotide substrate, and it performed this reaction much more efficiently than it unwound the same helix free in solution. This increased efficiency from adjacent secondary and tertiary structure most simply suggested that CYT-19 forms additional interactions with the structured RNA via a site on the protein that is distinct from the one responsible for unwinding, and further work has localized this second site on the protein to an ancillary C-terminal domain (204). The second observation is that this unwinding activity was eliminated by formation of tertiary contacts between the helix and the body of the structured RNA.

These results led to a model for the mechanism of action of CYT-19, and presumably other DExD/H-box proteins that act as general RNA chaperones, that is distinctly different from the mechanism of processive DNA and RNA helicases (Figure 6). In this model, binding via an ancillary domain tethers the unwinding active site of CYT-19 in proximity to its structured RNA substrates. The protein preferentially disrupts structural elements that are not tightly packed to the rest of the RNA structure while remaining bound to the RNA through its ancillary domain. This mechanism accounts for the lack of a polarity requirement for single-stranded extensions in model studies of DExD/H-box proteins (189, 190); presumably the ancillary domain is binding to the extension and remaining bound during unwinding. The mechanism also accounts for the low processivity of unwinding, because continued association by the ancillary domain will most likely prevent significant translocation by the unwinding domain.

One feature of this mechanism, the preference for disruption of loosely-associated structure, may provide a means of targeting CYT-19 to act on misfolded rather than native RNAs, because misfolded RNAs are more likely to have structural elements that are unable to pack correctly. Further, the non-native structural elements are the ones that are less likely to pack, so this mechanism may bias chaperones to disrupt the non-native portions of misfolded structures. Similar mechanisms may apply to DExD/H-box proteins that interact with specific RNAs or RNPs, except that in these cases the ancillary domain binding directs the protein to act on particular substrates.

7. SUMMARY AND PERSPECTIVE

Over the course of nearly forty years, *in vitro* studies of RNA folding have shown time and time again that RNAs are prone to forming misfolded conformations, and considerations of the structures and properties of RNA suggest that misfolded species are likely to be encountered during folding *in vivo* as well. Numerous proteins that bind RNA non-specifically have been shown to possess RNA chaperone activity *in vitro*, and some of these proteins have been shown to function by facilitating specific rearrangements or complex formation of RNAs. Whether any of these proteins function as general chaperones, interacting with and facilitating folding of an assortment of RNAs, remains unclear. In contrast, evidence is currently mounting that a subset of the myriad DExD/H-box proteins function more generally, although the range of RNAs that have been shown to be assisted in folding by these RNA chaperones is currently limited to

group I and group II introns. Proper folding of some RNAs may be so difficult that cells have devoted chaperones specifically to that function; for example, several DExD/H-box proteins appear to function specifically in ribosome biogenesis (205). On the other hand, the large number and diverse structures of RNAs would suggest that there must be some chaperones that are available to function non-specifically, unfolding whatever misfolded RNA structures happen to arise. The need for general chaperones is underscored by recent revelations that the number of non-coding RNAs is much larger than was previously known (206).

Although our knowledge of the involvement and actions of chaperones in RNA folding is increasing, many questions remain. Most centrally, which RNAs are dependent on chaperones for proper folding? Although a dependency is clear for RNAs and RNPs that require a particular chaperone, it is quite possible that many other RNAs are dependent on chaperones that have broad specificities and are functionally redundant with multiple chaperones, such that folding defects do not arise upon inactivation of any single chaperone. Are there different types of chaperone duties for different types of chaperone proteins? It might be imagined that the unfolding required by some misfolded RNAs to allow them to re-fold to their native states would be sufficiently unfavorable as to require energy in the form of ATP, whereas re-folding of others would proceed quite efficiently as long as chaperone proteins were present to stabilize the less structured intermediates and corresponding transitions states. Presumably the first class of misfolded structures would use DExD/H-box proteins, whereas the second would use other chaperones, but whether such a division of labor occurs among chaperone proteins is completely unexplored. Further, even the mechanistic assumptions underlying this expectation – that DExD/H-box proteins actively disengage contacts, whereas other chaperone proteins stabilize incrementally larger segments of RNA as they unfold on their own – are not established. And finally, if there are general chaperones running around the cell unfolding RNAs and RNPs, how does the cell protect its native RNAs from unwanted attention by chaperones?

8. ACKNOWLEDGMENTS

Work in my laboratory is supported by grants from the NIH (R01-GM070456) and the Welch Foundation (F-1563). I am grateful to Pilar Tijerina for help in preparing figures, and to members of my lab for comments on the manuscript.

9. REFERENCES

1. Staley, J.P. and C. Guthrie: Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* 92, 315-326 (1998)
2. Puglisi, J.D., S.C. Blanchard, and R. Green: Approaching translation at atomic resolution. *Nat Struct Biol* 7, 855-861 (2000)

3. Keenan, R.J., D.M. Freymann, R.M. Stroud, and P. Walter: The signal recognition particle. *Annu Rev Biochem* 70, 755-775 (2001)
4. Collins, K.: The biogenesis and regulation of telomerase holoenzymes. *Nat Rev Mol Cell Biol* 7, 484-494 (2006)
5. Evans, D., S.M. Marquez, and N.R. Pace: RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* 31, 333-341 (2006)
6. Tucker, B.J. and R.R. Breaker: Riboswitches as versatile gene control elements. *Curr Opin Struct Biol* 15, 342-348 (2005)
7. Hellen, C.U. and P. Sarnow: Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* 15, 1593-1612 (2001)
8. Batey, R.T.: Structures of regulatory elements in mRNAs. *Curr Opin Struct Biol* 16, 299-306 (2006)
9. Bunka, D.H. and P.G. Stockley: Aptamers come of age - at last. *Nat Rev Microbiol* 4, 588-596 (2006)
10. Gutell, R.R. and C.R. Woese: Higher order structural elements in ribosomal RNAs: pseudo-knots and the use of noncanonical pairs. *Proc Natl Acad Sci USA* 87, 663-667 (1990)
11. Hermann, T. and D.J. Patel: Stitching together RNA tertiary architectures. *J Mol Biol* 294, 829-849 (1999)
12. Leontis, N.B. and E. Westhof: Analysis of RNA motifs. *Curr Opin Struct Biol* 13, 300-308 (2003)
13. Tinoco, I., Jr. and C. Bustamante: How RNA folds. *J Mol Biol* 293, 271-281 (1999)
14. Wu, M. and I. Tinoco, Jr.: RNA folding causes secondary structure rearrangement. *Proc Natl Acad Sci USA* 95, 11555-11560 (1998)
15. Herschlag, D.: RNA chaperones and the RNA folding problem. *J Biol Chem* 270, 20871-20874 (1995)
16. Sigler, P.B.: An analysis of the structure of tRNA. *Annu Rev Biophys Bioeng* 4, 477-527 (1975)
17. Karpel, R.L., D.G. Swistel, N.S. Miller, M.E. Geroch, C. Lu, and J.R. Fresco: Acceleration of RNA renaturation by nucleic acid unwinding proteins. *Brookhaven Symp Biol* 26, 165-174 (1975)
18. Draper, D.E. and P.H. von Hippel: Nucleic acid binding properties of *Escherichia coli* ribosomal protein S1. I. Structure and interactions of binding site I. *J Mol Biol* 122, 321-338 (1978)
19. Karpel, R.L., N.S. Miller, and J.R. Fresco: Mechanistic studies of ribonucleic acid renaturation by a helix-destabilizing protein. *Biochemistry* 21, 2102-2108 (1982)
20. Mohr, S., J.M. Stryker, and A.M. Lambowitz: A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. *Cell* 109, 769-779 (2002)
21. Woese, C.R., L.J. Magrum, R. Gupta, R.B. Siegel, D.A. Stahl, J. Kop, N. Crawford, J. Brosius, R. Gutell, J.J. Hogan, and H.F. Noller: Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res* 8, 2275-2293 (1980)
22. Noller, H.F., J. Kop, V. Wheaton, J. Brosius, R.R. Gutell, A.M. Kopylov, F. Dohme, W. Herr, D.A. Stahl, R. Gupta, and C.R. Woese: Secondary structure model for 23S ribosomal RNA. *Nucleic Acids Res* 9, 6167-6189 (1981)
23. Noller, H.F. and C.R. Woese: Secondary structure of 16S ribosomal RNA. *Science* 212, 403-411 (1981)

24. Ban, N., P. Nissen, J. Hansen, P.B. Moore, and T.A. Steitz: The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905-920 (2000)
25. Wimberly, B.T., D.E. Brodersen, W.M. Clemons, Jr., R.J. Morgan-Warren, A.P. Carter, C. Vornheim, T. Hartsch, and V. Ramakrishnan: Structure of the 30S ribosomal subunit. *Nature* 407, 327-339 (2000)
26. DeVoe, H. and I. Tinoco, Jr.: The stability of helical polynucleotides: base contributions. *J Mol Biol* 4, 500-517 (1962)
27. Turner, D.H.: Thermodynamics of base pairing. *Curr Opin Struct Biol* 6, 299-304 (1996)
28. Tinoco, I., Jr., O.C. Uhlenbeck, and M.D. Levine: Estimation of secondary structure in ribonucleic acids. *Nature* 230, 362-367 (1971)
29. Gralla, J. and D.M. Crothers: Free energy of imperfect nucleic acid helices. II. Small hairpin loops. *J Mol Biol* 73, 497-511 (1973)
30. Borer, P.N., B. Dengler, I. Tinoco, Jr., and O.C. Uhlenbeck: Stability of ribonucleic acid double-stranded helices. *J Mol Biol* 86, 843-853 (1974)
31. Freier, S.M., R. Kierzek, J.A. Jaeger, N. Sugimoto, M.H. Caruthers, T. Neilson, and D.H. Turner: Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA* 83, 9373-9377 (1986)
32. Nissen, P., J.A. Ippolito, N. Ban, P.B. Moore, and T.A. Steitz: RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci USA* 98, 4899-4903 (2001)
33. Doherty, E.A., R.T. Batey, B. Masquida, and J.A. Doudna: A universal mode of helix packing in RNA. *Nat Struct Biol* 8, 339-343 (2001)
34. Lee, J.C., R.R. Gutell, and R. Russell: The UAA/GAN internal loop motif: a new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions. *J Mol Biol* 360, 978-988 (2006)
35. Davis, J.H., M. Tonelli, L.G. Scott, L. Jaeger, J.R. Williamson, and S.E. Butcher: RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J Mol Biol* 351, 371-382 (2005)
36. Hodak, J.H., J.L. Fiore, D.J. Nesbitt, C.D. Downey, and A. Pardi: Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc Natl Acad Sci USA* 102, 10505-10510 (2005)
37. Downey, C.D., J.L. Fiore, C.D. Stoddard, J.H. Hodak, D.J. Nesbitt, and A. Pardi: Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker. *Biochemistry* 45, 3664-3673 (2006)
38. Correll, C.C., A. Munishkin, Y.L. Chan, Z. Ren, I.G. Wool, and T.A. Steitz: Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc Natl Acad Sci U S A* 95, 13436-13441 (1998)
39. Conn, G.L., D.E. Draper, E.E. Lattman, and A.G. Gittis: Crystal structure of a conserved ribosomal protein-RNA complex. *Science* 284, 1171-1174 (1999)
40. Agalarov, S.C., G. Sridhar Prasad, P.M. Funke, C.D. Stout, and J.R. Williamson: Structure of the S15,S6,S18-rRNA complex: assembly of the 30S ribosome central domain. *Science* 288, 107-113 (2000)
41. Correll, C.C., B. Freeborn, P.B. Moore, and T.A. Steitz: Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* 91, 705-712 (1997)
42. Puglisi, E.V., R. Green, H.F. Noller, and J.D. Puglisi: Structure of a conserved RNA component of the peptidyl transferase centre. *Nat Struct Biol* 4, 775-778 (1997)
43. Treiber, D.K., M.S. Rook, P.P. Zarrinkar, and J.R. Williamson: Kinetic intermediates trapped by native interactions in RNA folding. *Science* 279, 1943-1946 (1998)
44. Russell, R., R. Das, H. Suh, K. Travers, A. Laederach, M. Engelhardt, and D. Herschlag: The paradoxical behavior of a highly structured misfolded intermediate in RNA folding. *J Mol Biol* 363, 531-544 (2006)
45. Fang, X.W., T. Pan, and T.R. Sosnick: Mg²⁺-dependent folding of a large ribozyme without kinetic traps. *Nat Struct Biol* 6, 1091-1095 (1999)
46. Fang, X.W., P. Thiyagarajan, T.R. Sosnick, and T. Pan: The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc Natl Acad Sci USA* 99, 8518-8523 (2002)
47. Swisher, J.F., L.J. Su, M. Brenowitz, V.E. Anderson, and A.M. Pyle: Productive folding to the native state by a group II intron ribozyme. *J Mol Biol* 315, 297-310 (2002)
48. Su, L.J., M. Brenowitz, and A.M. Pyle: An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. *J Mol Biol* 334, 639-652 (2003)
49. Treiber, D.K. and J.R. Williamson: Exposing the kinetic traps in RNA folding. *Curr Opin Struct Biol* 9, 339-345 (1999)
50. Holley, R.W., J. Apgar, G.A. Everett, J.T. Madison, M. Marquisee, S.H. Merrill, J.R. Penswick, and A. Zamir: Structure of a ribonucleic acid. *Science* 147, 1462-1465 (1965)
51. Gartland, W.J. and N. Sueoka: Two interconvertible forms of tryptophanyl tRNA in *E. coli*. *Proc Natl Acad Sci USA* 55, 948-956 (1966)
52. Lindahl, T., A. Adams, and J.R. Fresco: Renaturation of transfer ribonucleic acids through site binding of magnesium. *Proc Natl Acad Sci USA* 55, 941-948 (1966)
53. Ishida, T. and N. Sueoka: Rearrangement of the secondary structure of tryptophan sRNA in *Escherichia coli*. *Proc Natl Acad Sci U S A* 58, 1080-1087 (1967)
54. Adams, A., T. Lindahl, and J.R. Fresco: Conformational differences between the biologically active and inactive forms of a transfer ribonucleic acid. *Proc Natl Acad Sci USA* 57, 1684-1691 (1967)
55. Uhlenbeck, O.C., J.G. Chirikjian, and J.R. Fresco: Oligonucleotide binding to the native and denatured conformers of yeast transfer RNA-3 Leu. *J Mol Biol* 89, 495-504 (1974)
56. Cole, P.E., S.K. Yang, and D.M. Crothers: Conformational changes of transfer ribonucleic acid. Equilibrium phase diagrams. *Biochemistry* 11, 4358-4368 (1972)
57. Cole, P.E. and D.M. Crothers: Conformational changes of transfer ribonucleic acid. Relaxation kinetics of the early melting transition of methionine transfer ribonucleic

- acid (*Escherichia coli*). *Biochemistry* 11, 4368-4374 (1972)
58. Madore, E., C. Florentz, R. Giege, and J. Lapointe: Magnesium-dependent alternative foldings of active and inactive *Escherichia coli* tRNA(Glu) revealed by chemical probing. *Nucleic Acids Res* 27, 3583-3588 (1999)
59. Weidner, H. and D.M. Crothers: Pathway-dependent refolding of *E. coli* 5S RNA. *Nucleic Acids Res* 4, 3401-3414 (1977)
60. Aubert, M., J.F. Scott, M. Reynier, and R. Monier: Rearrangement of the conformation of *Escherichia coli* 5S RNA. *Proc Natl Acad Sci U S A* 61, 292-299 (1968)
61. Weidner, H., R. Yuan, and D.M. Crothers: Does 5S RNA function by a switch between two secondary structures? *Nature* 266, 193-194 (1977)
62. Lee, F. and C. Yanofsky: Transcription termination at the trp operon attenuators of *Escherichia coli* and *Salmonella typhimurium*: RNA secondary structure and regulation of termination. *Proc Natl Acad Sci U S A* 74, 4365-4369 (1977)
63. Zurawski, G., D. Elseviers, G.V. Stauffer, and C. Yanofsky: Translational control of transcription termination at the attenuator of the *Escherichia coli* tryptophan operon. *Proc Natl Acad Sci U S A* 75, 5988-5992 (1978)
64. Yanofsky, C.: Attenuation in the control of expression of bacterial operons. *Nature* 289, 751-758 (1981)
65. Wagner, E.G. and R.W. Simons: Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* 48, 713-742 (1994)
66. Ambros, V.: microRNAs: tiny regulators with great potential. *Cell* 107, 823-826 (2001)
67. Esteban, J.A., A.R. Banerjee, and J.M. Burke: Kinetic mechanism of the hairpin ribozyme. Identification and characterization of two nonexchangeable conformations. *J Biol Chem* 272, 13629-13639 (1997)
68. Esteban, J.A., N.G. Walter, G. Kotzorek, J.E. Heckman, and J.M. Burke: Structural basis for heterogeneous kinetics: reengineering the hairpin ribozyme. *Proc Natl Acad Sci U S A* 95, 6091-6096 (1998)
69. Zhuang, X., H. Kim, M.J. Pereira, H.P. Babcock, N.G. Walter, and S. Chu: Correlating structural dynamics and function in single ribozyme molecules. *Science* 296, 1473-1476 (2002)
70. Murchie, A.I., J.B. Thomson, F. Walter, and D.M. Lilley: Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops. *Mol Cell* 1, 873-881 (1998)
71. Walter, N.G., J.M. Burke, and D.P. Millar: Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction. *Nat Struct Biol* 6, 544-549 (1999)
72. Tan, E., T.J. Wilson, M.K. Nahas, R.M. Clegg, D.M. Lilley, and T. Ha: A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc Natl Acad Sci U S A* 100, 9308-9313 (2003)
73. Chadalavada, D.M., S.M. Knudsen, S. Nakano, and P.C. Bevilacqua: A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme. *J Mol Biol* 301, 349-367 (2000)
74. Chadalavada, D.M., S.E. Senchak, and P.C. Bevilacqua: The folding pathway of the genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots. *J Mol Biol* 317, 559-575 (2002)
75. Kruger, K., P.J. Grabowski, A.J. Zaug, J. Sands, D.E. Gottschling, and T.R. Cech: Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* 31, 147-157 (1982)
76. Cech, T.R. and B.L. Bass: Biological catalysis by RNA. *Annu Rev Biochem* 55, 599-629 (1986)
77. Sugimoto, N., R. Kierzek, and D.H. Turner: Kinetics for reaction of a circularized intervening sequence with CU, UCU, CUCU, and CUCUCU: mechanistic implications from the dependence on temperature and on oligomer and Mg²⁺ concentrations. *Biochemistry* 27, 6384-6392 (1988)
78. Walstrum, S.A. and O.C. Uhlenbeck: The self-splicing RNA of *Tetrahymena* is trapped in a less active conformation by gel purification. *Biochemistry* 29, 10573-10576 (1990)
79. Herschlag, D. and T.R. Cech: Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry* 29, 10159-10171 (1990)
80. Woodson, S.A. and T.R. Cech: Alternative secondary structures in the 5' exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry* 30, 2042-2050 (1991)
81. Woodson, S.A.: Exon sequences distant from the splice junction are required for efficient self-splicing of the *Tetrahymena* IVS. *Nucleic Acids Res* 20, 4027-4032 (1992)
82. Woodson, S.A. and V.L. Emerick: An alternative helix in the 26S rRNA promotes excision and integration of the *Tetrahymena* intervening sequence. *Mol Cell Biol* 13, 1137-1145 (1993)
83. Zarrinkar, P.P. and J.R. Williamson: Kinetic intermediates in RNA folding. *Science* 265, 918-924 (1994)
84. Sclavi, B., S. Woodson, M. Sullivan, M.R. Chance, and M. Brenowitz: Time-resolved synchrotron X-ray "footprinting", a new approach to the study of nucleic acid structure and function: application to protein-DNA interactions and RNA folding. *J Mol Biol* 266, 144-159 (1997)
85. Sclavi, B., M. Sullivan, M.R. Chance, M. Brenowitz, and S.A. Woodson: RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science* 279, 1940-1943 (1998)
86. Brenowitz, M., M.R. Chance, G. Dhavan, and K. Takamoto: Probing the structural dynamics of nucleic acids by quantitative time-resolved and equilibrium hydroxyl radical "footprinting". *Curr Opin Struct Biol* 12, 648-653 (2002)
87. Uchida, T., K. Takamoto, Q. He, M.R. Chance, and M. Brenowitz: Multiple monovalent ion-dependent pathways for the folding of the L-21 *Tetrahymena thermophila* ribozyme. *J Mol Biol* 328, 463-478 (2003)
88. Shcherbakova, I., S. Gupta, M.R. Chance, and M. Brenowitz: Monovalent ion-mediated folding of the *Tetrahymena thermophila* ribozyme. *J Mol Biol* 342, 1431-1442 (2004)
89. Laederach, A., I. Shcherbakova, M.P. Liang, M. Brenowitz, and R.B. Altman: Local kinetic measures of

- macromolecular structure reveal partitioning among multiple parallel pathways from the earliest steps in the folding of a large RNA molecule. *J Mol Biol* 358, 1179-1190 (2006)
90. Pan, J. and S.A. Woodson: Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J Mol Biol* 280, 597-609 (1998)
 91. Pan, J., M.L. Deras, and S.A. Woodson: Fast folding of a ribozyme by stabilizing core interactions: Evidence for multiple folding pathways in RNA. *J Mol Biol* 296, 133-144 (2000)
 92. Serra, M.J. and D.H. Turner: Predicting thermodynamic properties of RNA. *Methods Enzymol* 259, 242-261 (1995)
 93. Rook, M.S., D.K. Treiber, and J.R. Williamson: Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. *J Mol Biol* 281, 609-620 (1998)
 94. Treiber, D.K. and J.R. Williamson: Concerted kinetic folding of a multidomain ribozyme with a disrupted loop-receptor interaction. *J Mol Biol* 305, 11-21 (2001)
 95. Russell, R. and D. Herschlag: New pathways in folding of the *Tetrahymena* group I RNA enzyme. *J Mol Biol* 291, 1155-1167 (1999)
 96. Russell, R. and D. Herschlag: Probing the folding landscape of the *Tetrahymena* ribozyme: Commitment to form the native conformation is late in the folding pathway. *J Mol Biol* 308, 839-851 (2001)
 97. Russell, R., I.S. Millett, S. Doniach, and D. Herschlag: Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nat Struct Biol* 7, 367-370 (2000)
 98. Ohki, Y., Y. Ikawa, H. Shiraishi, and T. Inoue: Mispairing P3 region in the hierarchical folding pathway of the *Tetrahymena* ribozyme. *Genes Cells* 7, 851-860 (2002)
 99. Pichler, A. and R. Schroeder: Folding problems of the 5' splice site containing the P1 stem of the group I thymidylate synthase intron: substrate binding inhibition *in vitro* and mis-splicing *in vivo*. *J Biol Chem* 277, 17987-17993 (2002)
 100. Jiang, Y.F., M. Xiao, P. Yin, and Y. Zhang: Monovalent cations use multiple mechanisms to resolve ribozyme misfolding. *RNA* 12, 561-566 (2006)
 101. Russell, R., X. Zhuang, H.P. Babcock, I.S. Millett, S. Doniach, S. Chu, and D. Herschlag: Exploring the folding landscape of a structured RNA. *Proc Natl Acad Sci USA* 99, 155-160 (2002)
 102. Heilman-Miller, S.L., D. Thirumalai, and S.A. Woodson: Role of counterion condensation in folding of the *Tetrahymena* ribozyme. I. Equilibrium stabilization by cations. *J Mol Biol* 306, 1157-1166 (2001)
 103. Rangan, P., B. Masquida, E. Westhof, and S.A. Woodson: Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc Natl Acad Sci USA* 100, 1574-1579 (2003)
 104. Chauhan, S., G. Caliskan, R.M. Briber, U. Perez-Salas, P. Rangan, D. Thirumalai, and S.A. Woodson: RNA tertiary interactions mediate native collapse of a bacterial group I ribozyme. *J Mol Biol* 353, 1199-1209 (2005)
 105. Guerrier-Takada, C., K. Gardiner, T. Marsh, N. Pace, and S. Altman: The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849-857 (1983)
 106. Pace, N.R. and J.W. Brown: Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol* 177, 1919-1928 (1995)
 107. Frank, D.N. and N.R. Pace: Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* 67, 153-180 (1998)
 108. Xiao, S., F. Scott, C.A. Fierke, and D.R. Engelke: Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes. *Annu Rev Biochem* 71, 165-189 (2002)
 109. Loria, A. and T. Pan: Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA* 2, 551-563 (1996)
 110. Pan, T. and M. Jakacka: Multiple substrate binding sites in the ribozyme from *Bacillus subtilis* RNase P. *EMBO J* 15, 2249-2255 (1996)
 111. Altman, S. and C. Guerrier-Takada: M1 RNA, the RNA subunit of *Escherichia coli* ribonuclease P, can undergo a pH-sensitive conformational change. *Biochemistry* 25, 1205-1208 (1986)
 112. Zarrinkar, P.P., J. Wang, and J.R. Williamson: Slow folding kinetics of RNase P RNA. *RNA* 2, 564-573 (1996)
 113. Pan, T. and T.R. Sosnick: Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nat Struct Biol* 4, 931-938 (1997)
 114. Pan, T., X. Fang, and T. Sosnick: Pathway modulation, circular permutation and rapid RNA folding under kinetic control. *J Mol Biol* 286, 721-731 (1999)
 115. Xie, Z., N. Srividya, T.R. Sosnick, T. Pan, and N.F. Scherer: Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc Natl Acad Sci USA* 101, 534-539 (2004)
 116. Krasilnikov, A.S., X. Yang, T. Pan, and A. Mondragon: Crystal structure of the specificity domain of ribonuclease P. *Nature* 421, 760-764 (2003)
 117. Kazantsev, A.V., A.A. Krivenko, D.J. Harrington, S.R. Holbrook, P.D. Adams, and N.R. Pace: Crystal structure of a bacterial ribonuclease P RNA. *Proc Natl Acad Sci USA* 102, 13392-13397 (2005)
 118. Torres-Larios, A., K.K. Swinger, A.S. Krasilnikov, T. Pan, and A. Mondragon: Crystal structure of the RNA component of bacterial ribonuclease P. *Nature* 437, 584-587 (2005)
 119. Torres-Larios, A., K.K. Swinger, T. Pan, and A. Mondragon: Structure of ribonuclease P—a universal ribozyme. *Curr Opin Struct Biol* 16, 327-335 (2006)
 120. Baird, N.J., E. Westhof, H. Qin, T. Pan, and T.R. Sosnick: Structure of a folding intermediate reveals the interplay between core and peripheral elements in RNA folding. *J Mol Biol* 352, 712-722 (2005)
 121. Held, W.A., B. Ballou, S. Mizushima, and M. Nomura: Assembly mapping of 30 S ribosomal proteins from *Escherichia coli*. Further studies. *J Biol Chem* 249, 3103-3111 (1974)
 122. Traub, P. and M. Nomura: Structure and function of *Escherichia coli* ribosomes. VI. Mechanism of assembly of 30 S ribosomes studied *in vitro*. *J Mol Biol* 40, 391-413 (1969)
 123. Held, W.A. and M. Nomura: Rate determining step in the reconstitution of *Escherichia coli* 30S ribosomal subunits. *Biochemistry* 12, 3273-3281 (1973)

124. Zamir, A., R. Miskin, and D. Elson: Inactivation and reactivation of ribosomal subunits: amino acyl-transfer RNA binding activity of the 30 s subunit of *Escherichia coli*. *J Mol Biol* 60, 347-364 (1971)
125. Ginzburg, I., R. Miskin, and A. Zamir: N-ethyl maleimide as a probe for the study of functional sites and conformations of 30 S ribosomal subunits. *J Mol Biol* 79, 481-494 (1973)
126. Hogan, J.J. and H.F. Noller: Altered topography of 16S RNA in the inactive form of *Escherichia coli* 30S ribosomal subunits. *Biochemistry* 17, 587-593 (1978)
127. Holmes, K.L. and G.M. Culver: Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nat Struct Mol Biol* 11, 179-186 (2004)
128. Holmes, K.L. and G.M. Culver: Analysis of conformational changes in 16 S rRNA during the course of 30 S subunit assembly. *J Mol Biol* 354, 340-357 (2005)
129. Talkington, M.W., G. Siuzdak, and J.R. Williamson: An assembly landscape for the 30S ribosomal subunit. *Nature* 438, 628-632 (2005)
130. Pontius, B.W. and P. Berg: Renaturation of complementary DNA strands mediated by purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: implications for a mechanism for rapid molecular assembly. *Proc Natl Acad Sci USA* 87, 8403-8407 (1990)
131. Kumar, A. and S.H. Wilson: Studies of the strand-annealing activity of mammalian hnRNP complex protein A1. *Biochemistry* 29, 10717-10722 (1990)
132. Pontius, B.W. and P. Berg: Rapid assembly and disassembly of complementary DNA strands through an equilibrium intermediate state mediated by A1 hnRNP protein. *J Biol Chem* 267, 13815-13818 (1992)
133. Tsuchihashi, Z., M. Khosla, and D. Herschlag: Protein enhancement of hammerhead ribozyme catalysis. *Science* 262, 99-102 (1993)
134. Herschlag, D., M. Khosla, Z. Tsuchihashi, and R.L. Karpel: An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J* 13, 2913-2924 (1994)
135. Prats, A.C., L. Sarih, C. Gabus, S. Litvak, G. Keith, and J.L. Darlix: Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. *EMBO J* 7, 1777-1783 (1988)
136. Feng, Y.X., T.D. Copeland, L.E. Henderson, R.J. Gorelick, W.J. Bosche, J.G. Levin, and A. Rein: HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA *in vitro*. *Proc Natl Acad Sci USA* 93, 7577-7581 (1996)
137. Feng, Y.X., S. Campbell, D. Harvin, B. Ehresmann, C. Ehresmann, and A. Rein: The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J Virol* 73, 4251-4256 (1999)
138. Tanchou, V., C. Gabus, V. Rogemond, and J.L. Darlix: Formation of stable and functional HIV-1 nucleoprotein complexes *in vitro*. *J Mol Biol* 252, 563-571 (1995)
139. Wu, W., L.E. Henderson, T.D. Copeland, R.J. Gorelick, W.J. Bosche, A. Rein, and J.G. Levin: Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract. *J Virol* 70, 7132-7142 (1996)
140. Allain, B., M. Lapadat-Tapolsky, C. Berlioz, and J.L. Darlix: Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *EMBO J* 13, 973-981 (1994)
141. Guo, J., L.E. Henderson, J. Bess, B. Kane, and J.G. Levin: Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J Virol* 71, 5178-5188 (1997)
142. Rein, A., L.E. Henderson, and J.G. Levin: Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem Sci* 23, 297-301 (1998)
143. De Guzman, R.N., Z.R. Wu, C.C. Stalling, L. Pappalardo, P.N. Borer, and M.F. Summers: Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* 279, 384-388 (1998)
144. Urbaneja, M.A., M. Wu, J.R. Casas-Finet, and R.L. Karpel: HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J Mol Biol* 318, 749-764 (2002)
145. Cosa, G., E.J. Harbron, Y. Zeng, H.W. Liu, D.B. O'Connor, C. Eta-Hosokawa, K. Musier-Forsyth, and P.F. Barbara: Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys J* 87, 2759-2767 (2004)
146. Liu, H.W., G. Cosa, C.F. Landes, Y. Zeng, B.J. Kovalski, D.G. Mullen, G. Barany, K. Musier-Forsyth, and P.F. Barbara: Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription. *Biophys J* 89, 3470-3479 (2005)
147. Yusupov, M.M., G.Z. Yusupova, A. Baucom, K. Lieberman, T.N. Earnest, J.H. Cate, and H.F. Noller: Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883-896 (2001)
148. Coetzee, T., D. Herschlag, and M. Belfort: *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 introns by acting as RNA chaperones. *Genes Dev* 8, 1575-1588 (1994)
149. Semrad, K., R. Green, and R. Schroeder: RNA chaperone activity of large ribosomal subunit proteins from *Escherichia coli*. *RNA* 10, 1855-1860 (2004)
150. Clodi, E., K. Semrad, and R. Schroeder: Assaying RNA chaperone activity *in vivo* using a novel RNA folding trap. *EMBO J* 18, 3776-3782 (1999)
151. Rajkowitsch, L., K. Semrad, O. Mayer, and R. Schroeder: Assays for the RNA chaperone activity of proteins. *Biochem Soc Trans* 33, 450-456 (2005)
152. Schroeder, R., R. Grossberger, A. Pichler, and C. Waldsich: RNA folding *in vivo*. *Curr Opin Struct Biol* 12, 296-300 (2002)
153. Belisova, A., K. Semrad, O. Mayer, G. Kocian, E. Waigmann, R. Schroeder, and G. Steiner: RNA chaperone activity of protein components of human Ro RNPs. *RNA* 11, 1084-1094 (2005)

154. Stein, A.J., G. Fuchs, C. Fu, S.L. Wolin, and K.M. Reinisch: Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity. *Cell* 121, 529-539 (2005)
155. Zhang, A., V. Derbyshire, J.L. Salvo, and M. Belfort: *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA* 1, 783-793 (1995)
156. Zhang, A., S. Rimsky, M.E. Reaban, H. Buc, and M. Belfort: *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J* 15, 1340-1349 (1996)
157. Waldsich, C., R. Grossberger, and R. Schroeder: RNA chaperone StpA loosens interactions of the tertiary structure in the *td* group I intron *in vivo*. *Genes Dev* 16, 2300-2312 (2002)
158. Grossberger, R., O. Mayer, C. Waldsich, K. Semrad, S. Urschitz, and R. Schroeder: Influence of RNA structural stability on the RNA chaperone activity of the *Escherichia coli* protein StpA. *Nucleic Acids Res* 33, 2280-2289 (2005)
159. Ermolenko, D.N. and G.I. Makhatadze: Bacterial cold-shock proteins. *Cell Mol Life Sci* 59, 1902-1913 (2002)
160. Jiang, W., Y. Hou, and M. Inouye: CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *J Biol Chem* 272, 196-202 (1997)
161. Bae, W., B. Xia, M. Inouye, and K. Severinov: *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci USA* 97, 7784-7789 (2000)
162. Schindelin, H., W. Jiang, M. Inouye, and U. Heinemann: Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci USA* 91, 5119-5123 (1994)
163. Newkirk, K., W. Feng, W. Jiang, R. Tejero, S.D. Emerson, M. Inouye, and G.T. Montelione: Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc Natl Acad Sci USA* 91, 5114-5118 (1994)
164. Max, K.E., M. Zeeb, R. Bienert, J. Balbach, and U. Heinemann: T-rich DNA single strands bind to a preformed site on the bacterial cold shock protein Bs-CspB. *J Mol Biol* 360, 702-714 (2006)
165. Franze de Fernandez, M.T., L. Eoyang, and J.T. August: Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* 219, 588-590 (1968)
166. Tsui, H.C., H.C. Leung, and M.E. Winkler: Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol Microbiol* 13, 35-49 (1994)
167. Valentin-Hansen, P., M. Eriksen, and C. Udesen: The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* 51, 1525-1533 (2004)
168. Schumacher, M.A., R.F. Pearson, T. Moller, P. Valentin-Hansen, and R.G. Brennan: Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J* 21, 3546-3556 (2002)
169. Zhang, A., S. Altuvia, A. Tiwari, L. Argaman, R. Hengge-Aronis, and G. Storz: The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. *EMBO J* 17, 6061-6068 (1998)
170. Gottesman, S.: The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu Rev Microbiol* 58, 303-328 (2004)
171. Zhang, A., K.M. Wassarman, J. Ortega, A.C. Steven, and G. Storz: The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* 9, 11-22 (2002)
172. Storz, G., S. Altuvia, and K.M. Wassarman: An abundance of RNA regulators. *Annu Rev Biochem* 74, 199-217 (2005)
173. Wilusz, C.J. and J. Wilusz: Eukaryotic Lsm proteins: lessons from bacteria. *Nat Struct Mol Biol* 12, 1031-1036 (2005)
174. Geissmann, T.A. and D. Touati: Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J* 23, 396-405 (2004)
175. Mikulecky, P.J., M.K. Kaw, C.C. Brescia, J.C. Takach, D.D. Sledjeski, and A.L. Feig: *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol* 11, 1206-1214 (2004)
176. Lease, R.A. and S.A. Woodson: Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* 344, 1211-1223 (2004)
177. Cordin, O., J. Banroques, N.K. Tanner, and P. Linder: The DEAD-box protein family of RNA helicases. *Gene* 367, 17-37 (2006)
178. Linder, P.: DEAD-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res* 34, 4168-4180 (2006)
179. Silverman, E., G. Edwalds-Gilbert, and R.J. Lin: DEXD/H-box proteins and their partners: helping RNA helicases unwind. *Gene* 312, 1-16 (2003)
180. Linder, P., P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier, and P.P. Slonimski: Birth of the D-E-A-D box. *Nature* 337, 121-122 (1989)
181. Hirling, H., M. Scheffner, T. Restle, and H. Stahl: RNA helicase activity associated with the human p68 protein. *Nature* 339, 562-564 (1989)
182. Ray, B.K., T.G. Lawson, J.C. Kramer, M.H. Cladaras, J.A. Grifo, R.D. Abramson, W.C. Merrick, and R.E. Thach: ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. *J Biol Chem* 260, 7651-7658 (1985)
183. Abramson, R.D., T.E. Dever, T.G. Lawson, B.K. Ray, R.E. Thach, and W.C. Merrick: The ATP-dependent interaction of eukaryotic initiation factors with mRNA. *J Biol Chem* 262, 3826-3832 (1987)
184. Shuman, S.: Vaccinia virus RNA helicase: an essential enzyme related to the DE-H family of RNA-dependent NTPases. *Proc Natl Acad Sci USA* 89, 10935-10939 (1992)
185. Schwer, B. and C.H. Gross: Prp22, a DEXH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *EMBO J* 17, 2086-2094 (1998)
186. Kim, D.W., Y. Gwack, J.H. Han, and J. Choe: C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem Biophys Res Commun* 215, 160-166 (1995)
187. Rogers, G.W., Jr., N.J. Richter, and W.C. Merrick: Biochemical and kinetic characterization of the RNA

- helicase activity of eukaryotic initiation factor 4A. *J Biol Chem* 274, 12236-12244 (1999)
188. Iost, I., M. Dreyfus, and P. Linder: Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. *J Biol Chem* 274, 17677-17683 (1999)
189. Rogers, G.W., Jr., W.F. Lima, and W.C. Merrick: Further characterization of the helicase activity of eIF4A. Substrate specificity. *J Biol Chem* 276, 12598-12608 (2001)
190. Bizebard, T., I. Ferlenghi, I. Iost, and M. Dreyfus: Studies on three *E. coli* DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochemistry* 43, 7857-7866 (2004)
191. Gross, C.H. and S. Shuman: Vaccinia virions lacking the RNA helicase nucleoside triphosphate phosphohydrolase II are defective in early transcription. *J Virol* 70, 8549-8557 (1996)
192. Jankowsky, E., C.H. Gross, S. Shuman, and A.M. Pyle: The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature* 403, 447-451 (2000)
193. Jankowsky, E., C.H. Gross, S. Shuman, and A.M. Pyle: Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science* 291, 121-125 (2001)
194. Fairman, M.E., P.A. Maroney, W. Wang, H.A. Bowers, P. Gollnick, T.W. Nilsen, and E. Jankowsky: Protein displacement by DExH/D RNA helicases without duplex unwinding. *Science* 304, 730-734 (2004)
195. Chuang, R.Y., P.L. Weaver, Z. Liu, and T.H. Chang: Requirement of the DEAD-Box protein Ded1p for messenger RNA translation. *Science* 275, 1468-1471 (1997)
196. Yang, Q. and E. Jankowsky: ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry* 44, 13591-13601 (2005)
197. Rossler, O.G., A. Straka, and H. Stahl: Rearrangement of structured RNA via branch migration structures catalysed by the highly related DEAD-box proteins p68 and p72. *Nucleic Acids Res* 29, 2088-2096 (2001)
198. Lorsch, J.R.: RNA chaperones exist and DEAD box proteins get a life. *Cell* 109, 797-800 (2002)
199. Huang, H.R., C.E. Rowe, S. Mohr, Y. Jiang, A.M. Lambowitz, and P.S. Perlman: The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc Natl Acad Sci USA* 102, 163-168 (2005)
200. Seraphin, B., A. Boulet, M. Simon, and G. Faye: Construction of a yeast strain devoid of mitochondrial introns and its use to screen nuclear genes involved in mitochondrial splicing. *Proc Natl Acad Sci USA* 84, 6810-6814 (1987)
201. Seraphin, B., M. Simon, A. Boulet, and G. Faye: Mitochondrial splicing requires a protein from a novel helicase family. *Nature* 337, 84-87 (1989)
202. Mohr, S., M. Matsuura, P.S. Perlman, and A.M. Lambowitz: A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc Natl Acad Sci USA* 103, 3569-3574 (2006)
203. Tijerina, P., H. Bhaskaran, and R. Russell: Non-specific binding to structured RNA and preferential unwinding of an exposed helix by the CYT-19 protein, a DEAD-box RNA chaperone. *Proc Natl Acad Sci USA* 103, 16698-16703 (2006)
204. Grohman, J.K., M.D. Campo, H. Bhaskaran, P. Tijerina, A.M. Lambowitz, and R. Russell: Probing the mechanisms of DEAD-Box proteins as general RNA chaperones: The C-Terminal domain of CYT-19 mediates general recognition of RNA. *Biochemistry*, in press (2007)
205. Tanner, N.K. and P. Linder: DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell* 8, 251-262 (2001)
206. Claverie, J.M.: Fewer genes, more noncoding RNA. *Science* 309, 1529-1530 (2005)
207. Su, L., L. Chen, M. Egli, J.M. Berger, and A. Rich: Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat Struct Biol* 6, 285-292 (1999)
208. Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna: Crystal structure of a hepatitis delta virus ribozyme. *Nature* 395, 567-574 (1998)
209. Cate, J.H., A.R. Gooding, E. Podell, K. Zhou, B.L. Golden, C.E. Kundrot, T.R. Cech, and J.A. Doudna: Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 273, 1678-1685 (1996)

Key Words: RNA folding, DEAD-box Proteins, Cold-Shock Proteins, tRNA, group I intron, Review

Send correspondence to: Dr Rick Russell, Department of Chemistry and Biochemistry, The Institute For Cellular and Molecular Biology, University of Texas at Austin, Austin, TX 78712, Tel: 512-471-1514, Fax: 512-232-3432, E-mail: rick_russell@mail.utexas.edu

<http://www.bioscience.org/current/vol13.htm>