RNA Folding

I. Tinoco, Jr., J.D. Puglisi and J.R. Wyatt[1]

1 Introduction

The biological function of RNA determines its structure. Here we will review the structural elements that are known, or that have been proposed, to explain the biological functions. The emphasis will be on the conformations of the structural elements that make up folded RNA, and on their thermodynamic stability relative to the unfolded single strand.

Crystal structures of transfer RNAs have provided most of the structural information for folded RNA (Saenger 1984). The structures showed that duplex RNA is A-form with about 11 base pairs per turn, and they provided the anticodon loop as the prototypical hairpin loop. The interactions of the TΨC loop with the dihydrouracil loop have revealed different types of three-base interactions. The stacking of the stems of the tRNA cloverleaf suggests how stems can be arranged at junctions of two, three or four stems. The common conformation (three-dimensional structure) of the folded structure is crucial to the interactions of tRNA with messenger RNAs and with ribosomes. The specificity of interaction with aminoacyl synthetases depends on a few bases held at the right positions (Hou and Schimmel 1988; McClain et al. 1988; Sampson et al. 1989).

A type of tertiary structure not found in tRNAs was first proposed to explain why some plant viral RNAs were aminoacylated at their 3′-end. In order to mimic a tRNA the bases in the loop of a hairpin near the 3′-end of the viral RNA pair to a single-stranded region adjacent to the stem of the hairpin (Pleij et al. 1985). This produces a pseudoknot (Studnicka et al. 1989) consisting of two stems and two loops. NMR studies (Puglisi 1989) have shown that the two stems stack to form a continuous helix, which is only slightly perturbed by the attached loops. The helical region of the pseudoknot structure in turnip yellow mosaic virus is recognized by valyl-tRNA synthetase (Florentz and Giegé 1986). Pseudoknots have now been shown to be required for the correct function of several RNAs. Schimmel (1989) has recently reviewed their involvement in regulation of translation.

Unusual conformations have been discovered with as yet no known biological function. These structures indicate the conformational flexibility of RNA and they challenge one to find a function. Left-handed Z-RNA has a structure

[1] Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, CA 94720, USA

Nucleic Acids and Molecular Biology, Vol. 4
ed. by F. Eckstein and D.M.J. Lilley

similar to left-handed Z-DNA (Davis et al. 1990); it occurs in sequences of alternating r(CG) in very high salt concentrations (Hall et al. 1984). In DNA, triple-strand structures seem to be involved in gene regulation (Wells et al. 1988), and DNA duplexes with Watson-Crick-bonded parallel strands instead of anti-parallel strands have been found (van de Sande et al. 1988). Interactions among three bases stabilize the tertiary structure of tRNA and a few of the base pairs in tRNA involve parallel strands. Whether left-handed duplexes, parallel-

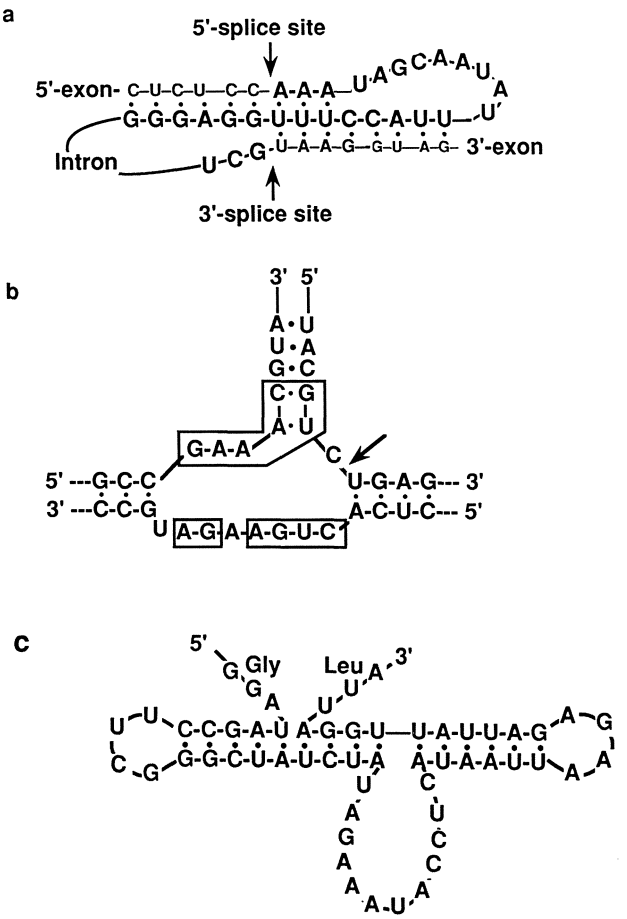

Fig. 1a–c. Folding of RNA proposed for different biochemical functions. **a** The internal guide sequence of type I introns which places the 5′-exon and 3′-exon close together for cleavage and ligation (Waring and Davies 1984). The sequence shown is for the intervening sequence of *Tetrahymena* ribosomal RNA. Note that two types of tertiary structure are proposed: a pseudoknot and a triple-strand interaction. **b** The hammerhead secondary structure for self-cleaving RNAs (Hutchins et al. 1986); conserved sequences are *boxed,* and the *arrow* indicates the cleavage site. Each stem may either be closed by a loop or may continue as a duplex; the sequence shown is from lucerne transient streak virus. The actual three-dimensional structure formed by the interaction of the three stems and loop is of great interest. **c** The proposed secondary structure of an untranslated messenger RNA sequence (Huang et al. 1988). The ribosome reads the Gly codon, skips 50 nucleotides, then reads the Leu codon

stranded duplexes, or extended triplexes occur in general in RNAs and what their biological significance is, if any, is not known.

The previous paragraph represents structures looking for function; there are many examples of functions looking for structures. The catalytic activity of RNA must reside in its correct folded conformation. What is this, at present magical, structure? Waring and Davies (1984) provide a general model for bringing the 5′ and 3′ splice sites together in class I self-splicing introns. The secondary structure of self-cleaving viroid RNAs (Hutchins et al. 1986) and their analogous RNA enzymes (Uhlenbeck 1987) is a hammerhead structure with a three-way junction. A puzzling finding is an "intron" which is not actually removed from the RNA but instead is just ignored during translation (Huang et al. 1988). All these proposed structures (Fig. 1) are based mainly on Watson-Crick base-pairing possibilities and some chemical reactivity data. Their three-dimensional structures are not known. There are many other examples of conserved secondary structures which correlate with some biological activity. Conformational questions include: Why does a hairpin with the sequence (-UUCG-) in a loop closed by a C·G base pair act as a stop signal for reverse transcriptase (Tuerk et al. 1988)? Are single-nucleotide bulges in RNA intra- or extra-helical? Conserved bulges which are required for protein binding occur for example in viroid RNAs (Keese and Symons 1985), 5S RNA (Romby et al. 1988) and R17 RNA (Wu and Uhlenbeck 1987).

It is useful to categorize the elements of the folded conformation as follows. Secondary-structure elements are double-stranded regions, hairpin loops, internal loops, bulges and junctions; these elements are connected by single-stranded regions. Tertiary structures are formed by the interaction of the secondary-structure elements and by the interaction of the secondary-structure elements and single-stranded regions. By understanding the conformations and stabilities of these elements, we hope to understand how RNA folds into its functional forms.

2 Secondary-Structure Elements

The folded conformations of RNA molecules are largely stabilized by the interactions that form double-stranded helical regions. Duplexes of antiparallel strands are stabilized both by hydrogen bonding between the strands and by intra- and inter-strand stacking of the aromatic bases. Since a folded RNA is not completely double-stranded, base-paired regions are separated by regions of unpaired bases, either various types of loops or single strands. In the following sections, the secondary-structure elements will be defined and their conformations discussed.

2.1 Duplexes

The base-paired regions of RNA form right-handed A-form double helices. The A-form helix found in RNA fibers and crystals (Saenger 1984) has a helical

repeat of about 11 base pairs per turn, one more base pair per turn than B-DNA. The main global difference between A-RNA and B-DNA is the displacement of the base pairs from the helix axis. In B-DNA, the displacement is zero; the base pairs stack on top of each other in the center of the helix. In contrast, the base pairs in an A-RNA duplex are displaced about 4 Å towards the minor groove. This causes a large difference in size between the two grooves of an A-RNA helix; the major groove is deep and narrow and the minor groove is shallow and wide. In the A-RNA helix, the sugars adopt a 3′-*endo* conformation; sugars are 2′-*endo* puckered in B-form DNA. The differences in sugar conformation result in a variation in distance between the phosphates of the polynucleotide chain, 5.9 Å in an A-RNA helix and 7 Å in the less tightly wound B-DNA duplex.

RNA duplexes in crystals are not uniform; the conformation varies with sequence. The number of base pairs per turn of the helices of $tRNA^{Phe}$ varies from about 10 in the stem of the D loop to 11 in the anticodon stem (Holbrook et al. 1978). Despite this variation, right-handed RNA helices have not been observed to fall into distinct families analogous to A-DNA and B-DNA. Even in the D stem of $tRNA^{Phe}$, which has a helical repeat approximately that of B-form DNA, the displacement of the bases away from the helical axis remains large and the sugar puckers are 3′-*endo* (Holbrook et al. 1978). As observed for DNA duplexes, global distortions of RNA helices are possible. A recent crystal structure of $U(UA)_6A$ shows that the oligonucleotide forms a duplex with close to canonical A-form geometry. However, the helix is divided into three regions by two kinks that are stabilized by interactions involving 2′-OH groups (Dock-Bregeon et al. 1988).

The conformations of duplex RNA in solution are only now being investigated. Proton NMR studies are consistent with A-form geometry (Chou et al. 1989; Varani et al. 1989; Puglisi 1989). Scalar coupling constants show that the ribose rings are 3′-*endo*. Dipolar (NOE) interactions indicate that the glycosidic angles, and the intra-strand and cross-strand distances are typical of A-form duplexes. The helical repeat for double-stranded RNA in solution has recently been determined to be 11.3 base pairs per turn (Bhattacharyya et al. 1990). The local conformation of RNA in solution may change with each base-pair step, and presumably the number of base pairs per turn in solution is different from that in crystals.

Other duplex structures have been found in addition to the common right-handed double strands with Watson-Crick base pairs. RNA oligonucleotides and polynucleotides containing alternating cytosine and guanine can adopt a left-handed double-stranded conformation structurally similar to Z-form DNA (Hall et al. 1984; Davis et al. 1990). The Z-RNA is favored by high salt, high temperature and low water activity. Anti-Z-RNA antibodies have been used to detect Z-RNA conformations in alcohol-fixed cells (Zarling et al. 1987). DNA can form stable parallel-stranded duplexes of $d(A)_n \cdot d(T)_n$; chemical probing suggests that the base pairs are of reverse Watson-Crick type (van de Sande et al. 1988). Interactions involving bases on strands running in parallel directions stabilize the tertiary structure at the D stem of $tRNA^{Phe}$ (Saenger 1984).

2.2 Hairpin Loops

Of the three hairpin loops in tRNA, only the anticodon loop is not involved in tertiary interactions. The anticodon loop of $tRNA^{Phe}$ contains seven nucleotides; the first two nucleotides from the 5′-end in the anticodon loop make a tight turn while the five nucleotides on the 3′-end are stacked and continue the helical geometry of the stem into the loop. The anticodon loop of tRNA may not be a good general model. In the secondary structure deduced for 16S ribosomal RNA most of the hairpin loops have four unbonded bases (Gutell et al. 1985). Preliminary NMR evidence from this laboratory suggests that the extra stable four-base loop, UUCG, closed by a C·G base pair (Tuerk et al. 1988) adopts a unique conformation; the loop G is in the *syn* conformation instead of the usual *anti* conformation. Smaller loops are also stable. A hairpin containing both AH^+·C and G·U mismatches has only three bases in its loop (Puglisi 1989). Each of the nucleotides in the loop adopts a 2′-*endo* sugar pucker. The first base on the 3′-side of the loop stacks weakly on the closing base pair. The other two nucleotides stack on one another; three nucleotides easily bridge the A-form stem.

In the hairpin loops we have studied in this laboratory containing three or four unbonded bases in the loop, the loop nucleotides adopt a 2′-*endo* sugar conformation. This allows the distances between phosphates to reach their maximum values; we expect 2′-*endo* sugars to be a general feature of small hairpin loops.

2.3 Bulges

Single-base bulges appear often in proposed secondary structures. For example, a strictly conserved helix with a bulged U forms the central core domain in viroids (Keese and Symons 1985) and a bulge, either A or U, is phylogenetically defined at the contact site between *E. coli* 5S ribosomal RNA and protein L18 (Peattie et al. 1981). There is convincing evidence that a required bulged A in the coat protein-binding site of the R17 bacteriophage intercalates into the helix (Wu and Uhlenbeck 1987). In contrast, NMR studies show that a bulged U at the center of an eight base-pair duplex is out of the helix (van den Hoogen et al. 1988). Helices containing bulged nucleotides preferentially bind intercalating dyes. Ethidium binds tightly to a helix in 16S rRNA containing a bulged A; without the bulge the same sequence has a much lower affinity for the dye. Double-strand specific enzymatic probing of hairpins, with and without a bulge, indicates that the presence of the bulged nucleotide affects the helical structure of the entire stem (White and Draper 1987). Bulged nucleotides may be important in protein recognition sites either because of their special conformation, or as preferred sites of intercalation of aromatic amino acids. Bulges of one or more nucleotides distort stacking of the bases in the duplex and induce a bend in the RNA in analogy to what happens in DNA (Woodson and Crothers 1989; Bhattacharyya et al. 1990; Draper, D, personal communication).

2.4 *Internal Loops*

Mismatches occur often in RNA duplex regions. All non-Watson-Crick appositions are classified as internal loops of two nucleotides, but their conformations and their effects on duplex stability differ widely. G·U pairs are often observed; they are found, for example, in the amino acid acceptor stems of $tRNA^{Phe}$ and $tRNA^{Ala}$ (Holbrook et al. 1978; McClain et al. 1988). An A·C base pair, which is seen at pH 7 but is stabilized by lowering the pH and presumably involves a protonated adenine, occurs in a hairpin studied in our laboratory (Puglisi 1989). Protonation of adenine at the imino nitrogen produces an AH^{+}·C pair with geometry similar to a G·U pair. The purine of the base pair is shifted towards the major groove, whereas the pyrimidine shifts towards the minor groove. These non-canonical appositions do not conform to the regular helical geometry and are less stable than Watson-Crick base pairs, but they can be incorporated within an RNA helix. Although the conformations of purine-purine base pairs are not known in general, an m_2^2G·A base pair at the end of the anticodon stem in $tRNA^{Phe}$ is significantly wider than a normal Watson-Crick base pair (Saenger 1984). A double-stranded RNA-dependent unwinding activity has been detected in several organisms (Bass and Weintraub 1988). This activity permanently denatures RNA duplexes by conversion of approximately 50% of the adenosines to inosine. The biological role of this function is unclear. The resulting base–base mismatch, I·U, is one of the most destabilizing by analogy to DNA mismatches (Martin et al. 1985).

Large internal loops which are purine-rich have been implicated as protein recognition sites. Lack of reactivity toward chemical reagents of most of the bases of an internal loop of 5S ribosomal RNA from spinach has led Ehresmann and co-workers to propose unusual base pairing that allows a helical geometry to continue through the loop region (Romby et al. 1988). NMR evidence also supports extensive stacking and hydrogen bonding in the analogous loop from *E. coli* 5S rRNA (Gewirth 1988). Several of the internal loops in the 5′-domain of 16S ribosomal RNA show resistance to chemical modification comparable to that of the most stable helices (Moazed et al. 1986). An oligonucleotide duplex containing an internal loop of eight nucleotides has been studied using proton NMR (Varani et al. 1989). The loop region, which has two G·A appositions, is characterized by extensive stacking in A-like geometry. Selective broadening of some aromatic and sugar resonances suggests that the loop nucleotides move relative to the flanking stems on the millisecond time-scale. Thus the major differences between the loop and stem regions in this duplex are dynamic rather than structural. Protein recognition of internal loops may be facilitated by the instability of these regions relative to a completely Watson-Crick base-paired duplex.

2.5 *Junctions*

Junctions are regions which connect three or more stems; the connecting region for two stems is an internal loop. Junctions commonly appear in the proposed

secondary-structure models of large RNAs. The tRNA cloverleaf is a four-stem junction; in the three-dimensional structure, the acceptor stem and the T stem stack coaxially as do the anticodon stem and the D stem. Several biochemically well characterized RNAs form three-stem junctions; junctions with as many as five stems are observed in 16S ribosomal RNA (Gutell et al. 1985). At any junction, as with tRNA, the arms have the potential to stack coaxially in pairs. In the absence of crystal structures, it is difficult to determine the orientation of the junction arms with respect to each other. In 5S ribosomal RNA, two different orientations of the helices at a three-stem junction have been proposed based on chemical reactivity of unbonded bases in the junction region (Christiansen et al. 1987; Romby et al. 1988). A junction of three stems forms the hammerhead self-cleaving domain (Hutchins et al. 1986; Uhlenbeck 1987) which specifically hydrolyzes a phosphodiester to form a 2′,3′-cyclic phosphate and a 5′-hydroxyl. The conserved nucleotides required for catalytic function are mainly in regions which are not base paired in the proposed secondary structure (Fig. 1b). Despite extensive study of the sequence and the solution conditions required for the cleavage reaction, nothing is known about the three-dimensional structure except that the three-stem regions are required.

2.6 Single-Stranded Regions

Single-stranded regions may occur at the 5′- or 3′-termini of RNA molecules or between secondary-structure regions. These single-stranded regions are different from loop regions in that their ends are not constrained; they should be similar in conformation to isolated single strands. Bulges and hairpin loops have considerably less conformational freedom than single strands, no matter how many unpaired bases they have, because the ends of the loop must be close together. Though single-stranded regions are less conformationally constrained than nucleotides involved in secondary or tertiary interactions, these regions will have partially ordered structures. At neutral pH, homopolynucleotides poly(rA) and poly(rC) form single-stranded stacked helices (Saenger 1984). Unpaired nucleotides often stack at the ends of duplexes (Varani et al. 1989; Puglisi 1989).

3 Tertiary-Structure Elements

Tertiary structure is formed by the interaction of secondary-structure elements with each other and with single-stranded regions. The interactions mainly involve base pairing, either Watson-Crick or other base–base hydrogen bonding; however, hydrogen bonding between bases and ribose or phosphates also occur. Stacking of bases, including intercalation, is also important; in $tRNA^{Phe}$ only 5 of the 76 bases are not stacked.

Interactions of hairpin loops with single-stranded regions produce pseudoknots. Internal loops, bulges, and junctions can interact with each other and with single-stranded regions. Interactions of any non-paired bases (in loops, bulges,

junctions or single-stranded regions) with duplex regions produce triple strands. Duplex–duplex interactions produce quadruple strands.

3.1 Pseudoknots

Pseudoknots were originally defined as base pairing of nucleotides separated by regions of secondary structure; this was done in order to explicitly exclude them from the algorithms to calculate secondary structure (Studnicka et al. 1978). Interest in a special type of pseudoknot (Fig. 2a) increased when it was proposed to explain the aminoacylation of the 3′-end of plant viral RNAs (Pleij et al. 1985). This type of pseudoknot is formed by pairing bases in a hairpin loop with a single-stranded region to form two stem and loop regions. Conservation of possible base-pairing regions in different plant viral RNAs, reactivity with single-strand and double-strand specific enzymes, and chemical reactivity provided compelling evidence for the pseudoknot. Other examples of pseudoknots include the internal guide sequence which positions the 5′-splice site near the 3′-splice site of class I introns (Waring and Davies 1984), a nucleation site for gene *32* protein binding to the gene *32* message in bacteriophage T4 (McPheeters et al. 1988), the core of the three-dimensional model for 16S ribosomal RNA (Stern et al. 1988), the recognition site for ribosomal protein S4 at a regulatory site in *E. coli* α-mRNA (Tang and Draper 1989) and a signal that causes frameshifting during the translation of a coronavirus RNA (Brierley et al. 1989). Although the original definition of pseudoknots was general, we will limit the term pseudoknot to the type found by Pleij et al. (1985).

We have studied a series of oligonucleotides that can fold to form pseudoknots with contiguous stems (Puglisi et al. 1988; Wyatt et al. 1989; Puglisi 1989; Wyatt 1990) in order to establish the three-dimensional structure, the minimum loop lengths required and the thermodynamic stability of pseudoknots. Single-strand (S_1) and double-strand specific (V_1) enzymes, and diethylpyrocarbonate were used to probe the structures of the sequences shown in Table 1. In 5 mM $MgCl_2$, 60 mM Na^+, pH 6.4 at 4°C only pseudoknots are present if both loops

Table 1. Sequences tested for their ability to form pseudoknots in 5 mM Mg^{2+} at 4°C. Eq'm means that a measurable amount of hairpin is in equilibrium with the pseudoknot

stem 1	loop 1	stem 2	stem 1	loop 2	stem 2	Formation
5′GCG	AUUU	CUGAC	CGC	UUUUUU	GUCAG3′	yes
5′GCG	AUUU	CUGAC	CGC	UUUUU	GUCAG3′	yes
5′GCG	AUUU	CUGAC	CGC	UUUU	GUCAG3′	yes
5′GCG	AUUU	CUGAC	CGC	UUU	GUCAG3′	yes
5′GCG	AUUU	CUGAC	CGC	UU	GUCAG3′	eq'm
5′GCG	AUU	CUGAC	CGC	UUUUUU	GUCAG3′	yes
5′GCG	UUU	CUGAC	CGC	UUUUUU	GUCAG3	yes
5′GCG	UUU	CUGAC	CGC	UCA	GUCAG3′	eq'm
5′GCG	UUU	CUGAC	CGC	UUUUCA	GUCAG3′	yes

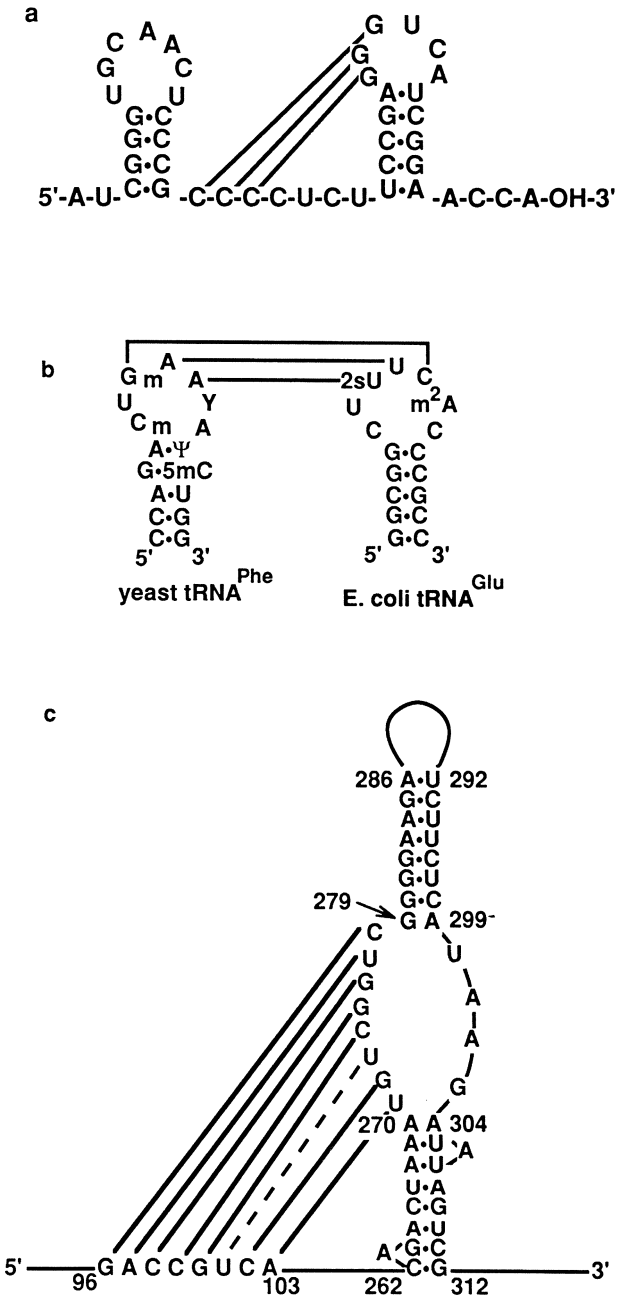

Fig. 2a–c. Pseudoknotted folding and other tertiary interactions in RNA. **a** The pseudoknot (hairpin-loop-to-single-strand) at the 3′-end of turnip yellow mosaic virus (Rietveld et al. 1982). **b** Hairpin-loop-to-hairpin-loop interaction between two complementary anticodons on two tRNAs. The complex is about 10^6 times more stable than the corresponding trinucleotides (Grosjean et al. 1976). **c** Internal-loop-to-single-strand-tertiary interaction proposed for the intervening sequence of *Tetrahymena* ribosomal RNA (Kim and Cech 1987)

have three or more nucleotides; however stability does depend on sequence. If loop 2 contains only two nucleotides, the pseudoknot is destabilized; pseudoknot and hairpins are present in measurable quantities. The minimum loop size in a pseudoknot depends on the number of base pairs in the stems (Pleij et al. 1985). One loop (loop 1) crosses the major groove of stem 2; the other (loop 2) crosses the minor groove of stem 1. In A-form RNA the minimum distance across the major groove is about 10 Å for six or seven base pairs; the distance remains below 20 Å for stems of from two to nine base pairs. The minimum distance across the minor groove is about 18 Å for two base pairs but the distance increases sharply with increasing number of base pairs; it is already 28 Å for five base pairs. As the extended phosphate–phosphate distance in a 2′-*endo* nucleotide is 7 Å, two nucleotides should suffice for loop 1, which crosses five base pairs as in our pseudoknot. For loop 2, which crosses three base pairs, the minimum number of loop nucleotides expected is three. Of course, distortion of the stems could change these requirements. In addition to a sufficient number of nucleotides to span the grooves in a pseudoknot, mM Mg^{2+} ion concentration or 500 mM Na^+ ion concentration is required for pseudoknot formation.

Proton NMR was used to determine the conformation of the pseudoknot shown in Fig. 3. The imino proton spectrum reveals the imino protons that exchange slowly with water, either because they are base paired, or because they are otherwise protected from solvent. At 4° C in Mg^{2+}, resonances corresponding to the expected eight base pairs are seen. The distances measured by NOESY and the torsion angles measured by COSY clearly reveal that the two stems are in A-form geometry. The ribose rings in the stem nucleotides are in 3′-*endo* conformations; the distances from the H1′-sugar proton to the aromatic base proton (H8 for purines and H6 for pyrimidines) of the next nucleotide on the same strand is about 4.5 Å instead of the 3.5 Å of B-DNA, and the characteristic short A-form distance from H2 of adenine to an H1′ on the opposite strand is

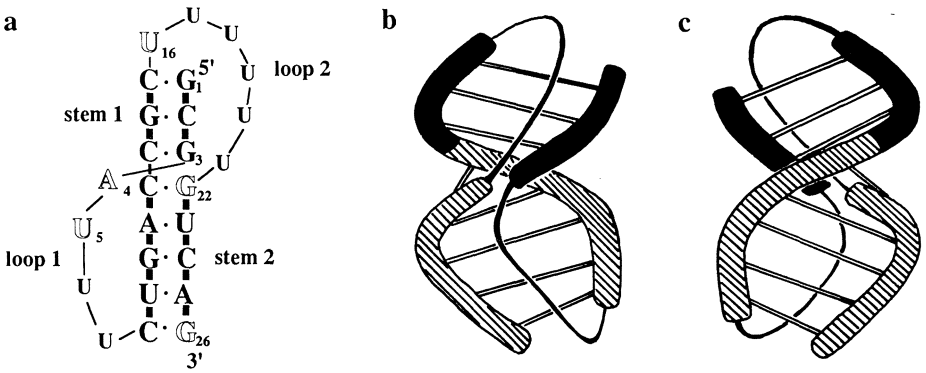

Fig. 3a–c. The conformation of a pseudoknot determined by proton NMR (Puglisi 1989). **a** Two contiguous stems stack coaxially to form a continuous A-form helix with two loops attached on one side. The nucleotides depicted in outline have 2′-*endo* sugars; the stem nucleotides are 3′-*endo* as expected for A-RNA. The U's in the loops have not been individually assigned, but most are 2′-*endo*. **b, c** The two faces of the pseudoknot are shown. One face is a continuous helix; the other face is crossed by the two loops

seen. The two stems of the pseudoknot stack on one another to form a continuous helix as determined by the aromatic proton to neighboring H1′ distances and by imino–imino proton distances. The only distortion is between the nucleotides (G_{22} and G_3 in Fig. 3) where the two loops join the duplex; the imino–imino distance is the same as the others, but the H1′ to aromatic proton connectivity is lost. The loop conformations are not well defined except for the first uracil of loop 2 which stacks on the 5′-terminal base pair. The loop nucleotides that can be assigned are in 2′-*endo* conformations and the majority of the other loop nucleotides are also 2′-*endo*. The stem nucleotide (G_{22}) at the junction to loop 2 is also 2′-*endo*. All the other stem nucleotides are 3′-*endo* except for the terminal 3′-nucleotide which is about half of each conformation, as usually found for terminal nucleotides. The structural conclusions deduced from the NMR measurements are summarized in Fig. 3.

A pseudoknot forms a continuous A-form duplex by coaxial stacking of the two contiguous stems. This pseudoknot duplex has two faces; one face is a regular A-form helix, the other face contains both loops (Fig. 3). The minor distortion at the loop–stem junctions is probably due to the strong electrostatic repulsion of the phosphate groups. This presumably accounts for the Mg^{2+} or equivalent high salt concentration requirement for the formation of the pseudoknot; it also may explain why one of the nucleotides at the loop–stem junctions is 2′-*endo*. The structure determined by the NMR study explains well why the two stems of a pseudoknot can be recognized by protein as a coaxially stacked helix.

3.2 Interactions of Internal Loops, Bulges and Junctions

Base pairing between a hairpin loop and a single-strand region forms a pseudoknot; this is only one out of the 14 interactions possible between loops, bulges, junctions and single-strand regions. Yeast $tRNA^{Phe}$ illustrates some of these different types of tertiary interactions (Saenger 1984). The D hairpin loop interacts with the four-stem junction by base pairing of $A_{14}{\cdot}U_8$ and $G_{15}{\cdot}C_{48}$; the latter base pair involves parallel strands. The D hairpin loop also forms two parallel-stranded base pairs with the T hairpin loop ($G_{18}{\cdot}\Psi_{55}$ and $G_{19}{\cdot}C_{56}$). Base pairing with antiparallel strands occurs between two tRNAs with complementary anticodons (Fig. 2b); these very stable complexes (Grosjean et al. 1976) may be good general models for hairpin-loop–hairpin-loop tertiary interactions in RNA.

A combination of conserved sequences and chemical reactivity was used to propose an internal-loop-to-single-strand tertiary interaction in the catalytic center of the intervening sequence in the ribosomal RNA of *Tetrahymena* (Fig. 2c) (Kim and Cech 1987). The role of this type of tertiary interaction in the precursor RNA of 5S rRNA from *B. subtilis* has been described by Stahl et al. (1979). Tertiary base pairing between two junctions has been proposed for the RNA component of RNase P (James et al. 1988). Examples of different types of tertiary interactions involving bases that are unpaired in the secondary structure should appear as the three-dimensional structures of more RNA molecules are determined.

3.3 Three-Base Interactions

Hydrogen-bonded planar structures of three bases occur in the folded tRNA structure (Saenger 1984); two bases are Watson-Crick base paired and the third bonds to the N7 and either the 6-keto or 6-amino of the purine base. In yeast $tRNA^{Phe}$ $A_9 \cdot (A_{23} \cdot U_{12})$ and $m^7G_{46} \cdot (G_{22} \cdot C_{13})$ occur but they can be replaced by $G_9 \cdot (G_{23} \cdot C_{12})$ and $A_{46} \cdot (A_{22} \cdot U_{13})$ respectively without loss of aminoacyl accepting ability (Sampson et al. 1989). If the bases are changed to prevent the three-base bonding, the three-dimensional structure and the activity of the tRNA are lost. Thus the three-base structure is required but the identity of the bases is less important. Triple strands with A·U·A, A·U·U or G·C·G bonding have been

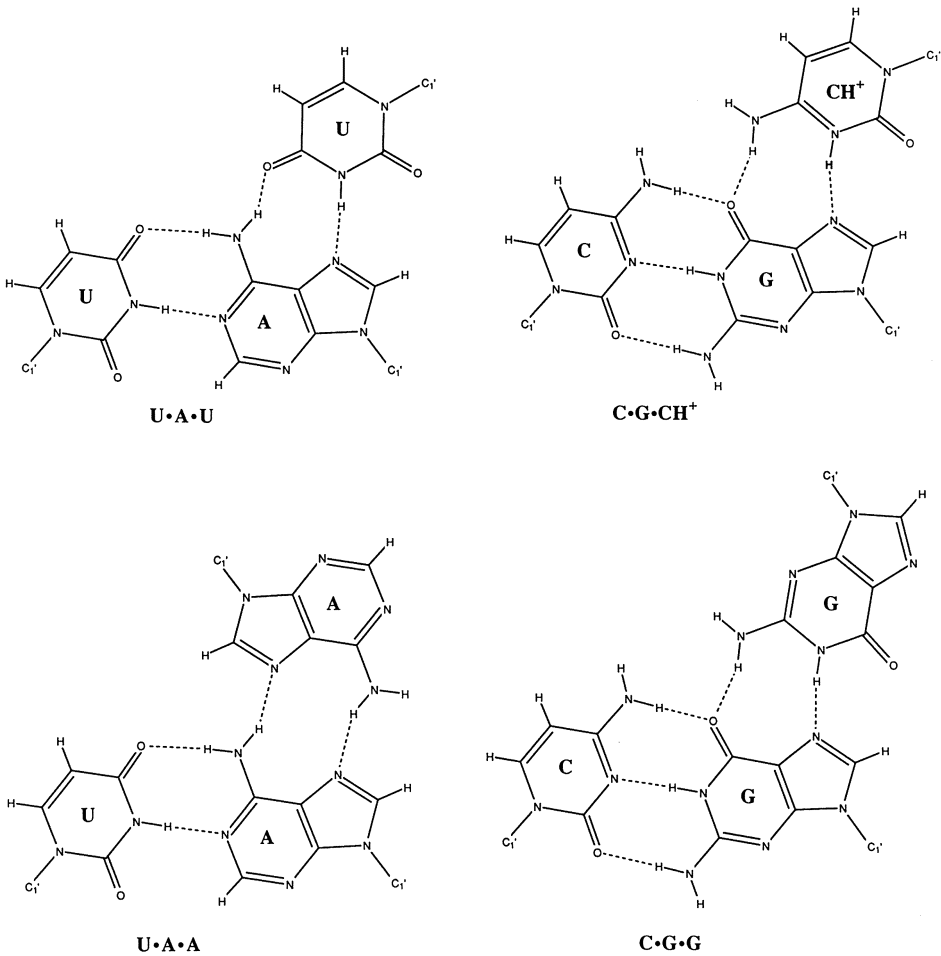

Fig. 4. Three-base hydrogen bonding found in tRNA and in synthetic polynucleotides. The U·A·U and C·G·CH^+ triples occur in homopolyribonucleotides; the third polypyrimidine strand is parallel to the polypurine strand of the Watson-Crick duplex. The other examples are from $tRNA^{Phe}$: $U_{12} \cdot A_{23} \cdot A_9$ and $C_{13} \cdot G_{22} \cdot$7-methyl G_{46}. The A_9 nucleotide is parallel to the A_{23} strand and the 7-methyl G strand is antiparallel to the G_{22} strand

suggested for aligning exons by class I introns (Waring and Davies 1984). As the likelihood of finding three-base interactions in folded RNA will increase with our knowledge of their properties, it is worthwhile to review the little that is known.

Homopolyribonucleotides poly(rA) and poly(rU) form antiparallel Watson-Crick duplexes of poly(rA)·poly(rU) below 0.1 M Na^+, but at higher salt concentrations or in Mg^{2+}, a triplex of poly(rU)·poly(rA)·poly(rU) is formed (Massoulié 1968). The poly(rU) strand is parallel to the poly(rA) and forms Hoogsteen A·U pairs in the major groove of the Watson-Crick helix (Arnott and Bond 1973). Triplexes of poly(rA)·poly(rA)·poly(rU) are formed at 2:1 poly(rA):poly(rU) ratios, if the poly(rA) is less than about 100 nucleotides in length (Broitman et al. 1987). Poly(rC) and poly(rG) can form triplexes of poly(rCH^+)·poly(rG)·poly(rC) at pH 6 with one proton per two C's (Thiele and Guschlbauer 1971). Evidence for poly(rG)·poly(rG)·poly(rC) triplexes has been found using agarose-linked polyribonucleotide affinity columns (Letai et al. 1988). The base-pairing schemes seen in tRNA and in the homopolyribonucleotides are shown in Fig. 4. It will be useful to learn what sequences can form triple strands and what the thermodynamic stability of each triplex is relative to the duplex and single strand.

Poly(rG), guanosine or guanylic acid can form structures with four equivalent hydrogen-bonded bases in a plane; the strands are all parallel (Zimmerman et al. 1975). Whether this has any relevance to RNA folding remains to be seen.

4 Thermodynamics and Kinetics of Structural Elements

RNA molecules adopt the thermodynamically most stable conformation in solution unless kinetic barriers trap an unstable form. The free energy of forming a folded conformation from a single strand consists of contributions from the various structural elements and from their interactions. Thus, in order to predict the correct folded conformation of an RNA molecule from its sequence, the thermodynamic contributions of both secondary and tertiary interactions need to be assessed. A more tractable problem is the prediction of only the secondary structure; the free energy of forming a secondary structure is simply the sum of the free energies of forming the constituent secondary structural elements. The following section will discuss present knowledge of the thermodynamic stability of duplex regions, single-strand regions, loops and bulges. Much less is known about the thermodynamic contributions of tertiary interactions; however, studies on pseudoknots provide support for the idea that tertiary interactions will not significantly disrupt secondary structure.

4.1 Thermodynamics of Secondary-Structure Elements

The free energy of a nucleic acid conformation is a complex superposition of contributions of interactions among bases, sugars, phosphates and solvent. Van

der Waals interactions, hydrogen bonding, electrostatic interactions and hydrophobic effects are involved. Quantitative calculation of these effects is not yet possible, instead, oligonucleotides are used to measure the free energy difference of a chosen structure from the single-stranded state. The best determined thermodynamic parameters are those for the formation of double-helical regions of RNA. The major contributions to the relative thermodynamic stability of a duplex are base-pairing and stacking interactions between neighboring base pairs. For the calculation of duplex stability, a nearest-neighbor model is approximately valid. The free energy of a duplex is the free energy of forming the initial base pair (initiation ΔG°, ca. +3.4 kcal/mol at 37°C) plus the sum of nearest-neighbor free energies. Turner and co-workers (Freier et al. 1986; Turner et al. 1988) have published the standard free energies, enthalpies, and entropies for forming all ten possible nearest neighbors involving Watson-Crick base pairing. The standard free energy decrease at 37°C for adding a base pair to a duplex ranges from −3.4 kcal/mol for neighboring G·C base pairs (5′ G_pC) to −0.9 kcal/mol for two A·U pairs ($A_pA{\cdot}U_pU$). The majority of work on the thermodynamics of model compounds has been done in a standard solvent of 1 M NaCl, 10 mM Na phosphate, pH 7. The published parameters can be used to calculate the expected thermodynamic behavior – T_m, ΔG°, ΔS°, and ΔH° – of any RNA duplex to within about 10% of the experimental value.

The favorable free energy of duplex formation comes mainly from the stacking interactions between neighboring bases; these interactions manifest themselves particularly in the favorable enthalpy of stacking. Nucleotides that are stacked within a structure can contribute significantly to a favorable free energy even if they are not base paired. Turner and co-workers (Freier et al. 1986) measured the free energy contributions of dangling nucleotides on the 5′- and 3′-sides of Watson-Crick base pairs. 5′-dangling nucleotides contribute little to the free energy at 37°C: only 0 to −0.5 kcal/mol. 3′-dangling nucleotides can make a significant contribution to free energy. 3′-dangling purines generally contribute over −1 kcal/mol. 3′-dangling pyrimidines contribute less. However, certain sequences have a high stability; a uridine on the 3′-side of a cytidine in a G·C pair contributes −1.2 kcal/mol. These results may help explain the particular stability of certain loops in RNA.

Very little is known about the thermodynamic stability of mispaired nucleotides. G·U base pairs are the only extensively characterized mispair; they are only slightly less stable than A·U pairs in similar sequence contexts (Sugimoto et al. 1986). In DNA duplexes, the thermodynamic stabilities of 12 mispairs surrounded by A·T pairs have been determined (Aboul-ela et al. 1985); G·T mispairs are the most stable in DNA. We have studied an A·C mispair within the stem of an RNA hairpin loop (Puglisi 1989). The stability of this mispair increases with decreasing pH; the proposed hydrogen bonding involves a protonated adenine. At pH 6.3 and 37°C, the hairpin with an $AH^+{\cdot}C$ mispair is 2 kcal/mol less stable in free energy than the same hairpin with an A·U base pair. This is roughly the destabilization observed by Aboul-ela et al. (1985) for A·C mispairs in DNA. A comprehensive study of mispair stability is needed.

Base–base mispairs produce the smallest internal loop, one containing two bases. The only systematic study of larger loops is that of Gralla and Crothers

(1973), in which a series of oligonucleotides were synthesized with internal loops consisting entirely of cytidines. These data have been modified using the improved duplex parameters and tabulated by Freier et al. (1986). The internal loop causes a destabilization of the duplex, with increasing destabilization with increasing size; the values ΔG° (37° C) run from +1.7 kcal/mol for an internal loop of four to +3.6 kcal/mol for an internal loop of ten. These parameters are very approximate, and certainly do not take into account sequence effects on stability, such as non-Watson-Crick base pair formation. In addition, the data were obtained for symmetric internal loops, with the same number of non-paired nucleotides on either strand. Asymmetric loops should also be studied.

Groebe and Uhlenbeck (1989) have studied the effects of single bulged nucleotides in the stem of a hairpin loop with four unpaired bases. The free energy of destabilization of the single base bulge was about +3 kcal/mol at 37° C. The destabilization depends on the identity of bulged nucleotide; U was least destabilizing (+2.7 kcal/mol) and G was most destabilizing (+3.5 kcal/mol).

The stability of RNA hairpin loops has been investigated by Groebe and Uhlenbeck (1988); loop sequences consisted of A_n, C_n, or U_n, with the stem sequence held constant. Loops of four to five nucleotides were found to be most stable, roughly 2 kcal/mol more stable than loops of nine nucleotides. Loops consisting of A or U had approximately the same free energy contribution at 37° C (ca. +3 to +4 kcal/mol), whereas loops with C were less stable (+ 6.5 to +7 kcal/mol). Specific sequence effects can be very important. If the loop sequence of the stable hairpin found in T4 phage is changed from -UUCG- to -UUUG-, the T_m drops 5° C (Tuerk et al. 1988). The closing base pair of a loop can also affect loop stability. If the closing pair in the above loop is changed from C·G to G·C, the T_m drops 6° C.

Experiments in our laboratory have further shown the sequence dependence of the stability of hairpin loops. Formation of non-Watson-Crick pairs can considerably change loop structure and stability. We studied a hairpin loop that was expected to form a stem of three base pairs and a loop of nine nucleotides. Instead, the sequence formed three extra base pairs, including A·U, A^+·C and G·U, such that a loop of only three nucleotides is formed. The stability of the smaller loop (three vs nine nucleotides worth about 2–3 kcal/mol) offset the destabilization due to the A^+·C mispair. The G·U base pair serves as the closing pair for this loop. In general, specific stacking or mispairing within a loop region will contribute a favorable free energy to loop formation. Current parameters do not account for these sequence effects.

4.2 Thermodynamics of Tertiary Structure

A number of methods, including absorbance melting, NMR and calorimetry have been used to monitor the unfolding of tRNA (Cole et al. 1972; Riesner et al. 1973; Crothers et al. 1974; Privalov and Filimonov 1978). These experiments demonstrate that tRNA melts in a complex, multiphasic manner. A phase diagram has been constructed as a function of ionic strength (Na^+) and temperature (Cole et al. 1972). At high ionic strength and low temperature, the molecule is in

its native tertiary structure; as the temperature is raised and/or the ionic strength is lowered, various partially denatured forms of the tRNA become stable. It appears that the tertiary structure melts first, and then separate regions of secondary structure subsequently denature (Crothers et al. 1974). Site-specific binding of Mg^{2+} stabilizes the tertiary structure of tRNA; in the crystal structure of $tRNA^{Phe}$ four hydrated magnesium ions are bound to nonhelical regions (Holbrook et al. 1977; Jack et al. 1977).

Oligonucleotides that form pseudoknots represent a simple model system for studying tertiary interactions. To use the thermodynamics of tertiary interactions to predict RNA structure one must know the stability of the tertiary structure relative to that of the secondary structure from which it is formed. The stability of a pseudoknot should be compared to that of the two hairpins that any pseudoknot sequence can form. The hairpins correspond to formation of either stem 1 (the 5′-hairpin) or stem 2 (the 3′-hairpin). In 5 mM Mg^{2+} at 37°C, the pseudoknot structure shown in Fig. 3 is only 0.5 to 1 kcal/mol more stable (a decrease in free energy) than either of the two hairpins. The enthalpy of the pseudoknot is only decreased by about 10 to 15 kcal/mol relative to the 3′-hairpin; this is less than the addition of three base pairs would produce in an RNA duplex as calculated from the nearest-neighbor parameters. The change in thermodynamic values on forming a duplex with bases in a hairpin loop is different from a duplex formed from two single strands.

Pseudoknots are stabilized through binding of Mg^{2+}. In the absence of Mg^{2+}, the pseudoknot structure shown in Fig. 3 is destabilized relative to the 5′-hairpin conformation. Small changes in sequence can also affect the structural equilibrium. Removal of a single A at position 4 in loop 1 destabilizes the 5′-hairpin structure, so that in the absence of Mg^{2+}, this sequence forms a 3′-hairpin conformation. If the loop 2 sequence is changed from $-U_6-$ to -UCA-, the 3′-hairpin is significantly stabilized relative to the pseudoknot even in Mg^{2+}. Since the structures only differ in free energies by 1 to 2 kcal/mol, these equilibria are easily manipulated.

The secondary structure of an RNA is composed of a folded group of stems, loops and single-stranded regions. Tertiary interactions between these secondary structure elements produce the correct three-dimensional structure. This implies a thermodynamic hierarchy for RNA folding:

$$\text{Primary (bonding)} > \text{Secondary} > \text{Tertiary}$$

The largest contribution to the total energy of an RNA molecule is the energy of covalent bonding. It is defined to be the same for folded and unfolded forms; any change in energy results from conformational changes. The secondary structure is the framework for the formation of the folded three-dimensional structure of an RNA. In the case of pseudoknot formation, tertiary interactions do not offer sufficient stabilization to compete with a region of base-paired secondary structure. The formation of the three base pair stem 1 results in a smaller decrease in free energy than if the same three base pairs formed a Watson-Crick stem elsewhere. An example of such a structural competition is shown in Fig. 5; the pseudoknot would not form, since the pairing to form the second hairpin is more favorable. A pseudoknot is the unimolecular equivalent of the bimolecular

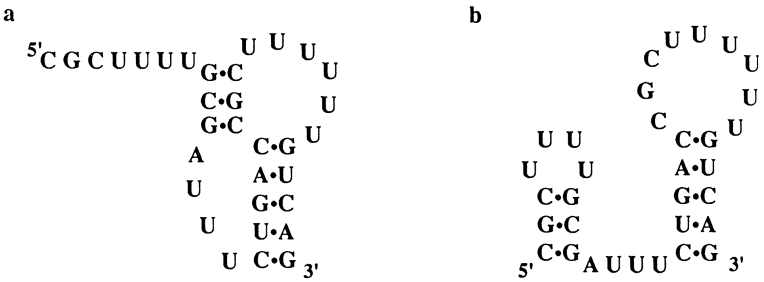

Fig. 5a, b. Competition between forming a tertiary structure pseudoknot and an additional secondary structure hairpin involving the same base pairs. Comparison of the measured free energy of formation of the pseudoknot with the predicted free energy for formation of the additional hairpin suggests that the hairpin would form

codon–anticodon (single strand to loop) interaction. It is interesting to note that the codon–anticodon interaction is quite stable (enthalpies equal to that predicted from nearest neighbors) (Eisinger et al. 1971; Uhlenbeck 1972). This may reflect the increased structural freedom of the bimolecular complex compared to that of the unimolecular pseudoknot interaction.

4.3 Prediction of Secondary Structure

An RNA strand can form a large number of possible Watson-Crick base-paired conformations. Prediction of the lowest free energy secondary structure is a conceptually simple, but for large RNA molecules, a computationally demanding task. A number of computer algorithms have been published which systematically search all base pairs possible to calculate the lowest free energy secondary structure (Nussinov and Jacobson 1980; Zuker and Stiegler 1981). More recently programs have been designed to calculate significantly different secondary structures near the lowest free energy; not only the optimal structure but also the suboptimal structures are found (Williams and Tinoco 1986; Zuker 1989).

Calculations are only as good as the available thermodynamic parameters. The parameters discussed above were all measured in 1 M NaCl, 10 mM Na phosphate, pH 7 (Turner et al. 1988). For duplex regions, the contributions of nearest-neighbor base pairs are well defined. However, thermodynamic parameters for RNA loops and other structural elements are poorly defined. Furthermore, specific binding of ligands like Mg^{2+} or proteins can affect relative stabilities of RNA secondary structures. For these reasons, the algorithms that calculate suboptimal structures (Williams and Tinoco 1986; Zuker 1989) are very useful. These families of structures can be used with experimental data or phylogenetic comparisons (Gutell et al. 1985) to deduce the actual secondary structure or structures.

Current computer algorithms do not consider the contribution of tertiary interactions. Any interaction between separated regions of secondary structure

is defined as a tertiary interaction. An example of this is the formation of a pseudoknot. This definition is convenient from the standpoint of structure prediction for several reasons. First, there are essentially no data on the free energy contributions from tertiary structures. Second, the inclusion of tertiary interactions makes the free energy of an RNA structure no longer a simple sum of contributions from separate elements. However, the present computer algorithms can identify many low free energy secondary structures. Since secondary structure should dominate the conformational free energy, tertiary interactions can be added to a possible secondary structure to predict a three-dimensional folded RNA.

4.4 Kinetics of Structural Transitions

Structural transitions probably play a crucial role in the function of RNA molecules: it is important to have a qualitative knowledge of the rates involved. Simple unstacking reactions of a dinucleotide are rapid and occur on a nanosecond time-scale, whereas base-pair formation occurs more slowly with rates in the millisecond (msec) range at μM concentrations. There have been a number of kinetic studies of helix-to-single-strand transitions in simple RNA duplexes. The association reaction of the two strands is bimolecular. Generally, the association rates observed for oligonucleotides are independent of chain length and are 10^5–$10^6\,M^{-1}\,s^{-1}$. Often negative activation energies are measured, indicating a complex mechanism, with an initial rapid pre-equilibrium step involving formation of a few base pairs (Riesner and Römer 1971). In contrast, the rate constants for the unimolecular helix disruption reaction are strongly dependent on chain length. For example the duplex $(A_4U_4)_2$ has a dissociation time constant of less than 1 ms at 21°C, whereas $(A_7U_7)_2$ has a dissociation time constant of about 1 s at 21°C. Activation energies are large and positive, and increase with increasing chain length. For shorter oligonucleotides, the activation energy is often approximately equal to the enthalpy of melting the helix, since these base pairs must be disrupted. For the unimolecular helix-to-single-strand transition of hairpin loops, the helix association reaction is unimolecular and is thus much more rapid (much lower activation entropy). Time constants of less than 10^{-4} s are observed near the T_m.

The kinetics of structural transitions in tRNA have been investigated in detail; various transitions are observed with increasing temperature that correspond to melting of different portions of the tRNA. In summary, relatively fast transitions, with time constants of less than ms involve the individual hairpin stem regions; slower transitions (> ms) involve more global changes of the structure (Crothers et al. 1974). A transition between an extended form of the tRNA with alternate base pairing stable at low salt, and the native structure stable at higher salt can be induced by a salt jump. The subsequent rearrangement of structure is extremely slow (time constant ca. 1000 s). Pörschke found that Mg^{2+} can induce a transition between different conformations of the anticodon loop. The slow time constant (ms time-scale) indicates more global changes than simple unstacking of nucleotides (Labuda and Pörschke 1980). A

mechanism was proposed that involves different stacked conformations of the anticodon loop. Such a transition may be important for codon recognition.

We have examined a similar transition between two nucleic acid conformations, that between pseudoknot and 5′-hairpin (corresponding to formation of only stem 1 in Fig. 3). The kinetics for the forward transition are quite slow, with rate constants on the order of 1 s^{-1}. An activation energy for the transition was determined as +42 kcal/mol; this is equal to the enthalpy of melting the 5 base pairs in stem 2. This is consistent with a mechanism in which the base pairs in stem 2 are broken to form an intermediate, with subsequent reformation of additional base pairs (i.e. $A_4 \cdot C_{12}$, $U_5 \cdot A_{11}$ and $G_6 \cdot U_{10}$). Since this transition occurs significantly below the melting temperature of stem 2, it is quite slow. As in tRNA, this example shows the considerable kinetic barriers that can be involved in reorganization of base pairing.

5 Conclusion

RNA interacts specifically with itself and other RNA molecules, with DNA and with proteins. To be able to understand all this requires knowledge of the range of conformations that an RNA can adopt as temperature, ionic conditions or ligand concentrations are varied. The time-scale of the folding and unfolding is also relevant. Instead of studying each biologically important RNA individually, one can study oligonucleotides that are prototypes for important elements in RNA structure. Characterization of the recurring RNA structural motifs will greatly simplify the study of the naturally occurring molecules. As RNA structures must have evolved to perform vital functions, knowledge of these structures should further the understanding of all the varied functions performed by RNA.

References

Aboul-ela F, Koh D, Tinoco I Jr, Martin F (1985) Base-base mismatches. Thermodynamics of double helix formation for $dCA_3XA_3G + dCT_3YT_3G$ (X,Y = A,C,G,T). Nucleic Acids Res 13:4811–4824

Arnott S, Bond PJ (1973) Structures for poly(U)·poly(A)·poly(U) triple stranded polynucleotides. Nat New Biol 244:99–101

Bass BL, Weintraub H (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55:1089–1098

Bhattacharyya A, Murchie AIH, Lilley DMJ (1990) RNA bulges and the helical periodicity of double stranded RNA. Nature 343:484–485

Brierley I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus frameshifting signal: requirement for an RNA pseudoknot. Cell 57:537–547

Broitman SL, Im DD, Fresco JR (1987) Formation of the triple-stranded polynucleotide helix, poly (A·A·U). Proc Natl Acad Sci USA 84:5120–5124

Chou S-H, Flynn P, Reid B (1989) Solid-phase synthesis and high-resolution NMR studies of two synthetic double-helical RNA dodecamers: r(CGCGAAUUCGCG) and r(CGCGUAUACGCG). Biochemistry 28:2422–2435

Christiansen J, Brown RS, Sproat BS, Garrett RA (1987) Xenopus transcription factor IIIA

binds primarily at junctions between helical stems and internal loops in oocyte 5S RNA. EMBO J 6:453–460
Cole PE, Yang SK, Crothers DM (1972) Conformational changes of transfer ribonucleic acid. Equilibrium phase diagram. Biochemistry 11:4358–4368
Crothers DM, Cole PE, Hilbers CW, Schulman RG (1974) The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. J Mol Biol 87:63–88
Davis PW, Adamiak RW, Tinoco I Jr (1990) Z-RNA: The solution NMR structure of r(CGCGCG). Biopolymers 29:109–121
Dock-Bregeon AC, Chevrier B, Podjarny A, Moras D, deBear JS, Gough GR, Gilham PT, Johnson JE (1988) High resolution structure of the RNA duplex $[U(U\text{-}A)_6A]_2$. Nature 335:375–378
Eisinger J, Feuer B, Yamane T (1971) Codon-anticodon binding in $tRNA^{Phe}$. Nat New Biol 231:126–128
Florentz C, Giegé R (1986) Contact areas of the turnip yellow mosaic virus tRNA-like structure interacting with yeast valyl-tRNA synthetase. J Mol Biol 191:117–130
Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T, Turner DH (1986) Improved free-energy parameters for predictions of RNA duplex stability. Proc Natl Acad Sci USA 83:9373–9377
Gewirth DT (1988) New approaches to structural studies of *E. coli* 5S RNA and its complexes with ribosomal proteins by NMR. Ph.D. Thesis, Yale University
Gralla J, Crothers DM (1973) Free energy of imperfect nucleic acid helices III. Small internal loops resulting from mismatches. J Mol Biol 78:301–319
Groebe DR, Uhlenbeck OC (1988) Characterization of RNA hairpin loop stability. Nucleic Acids Res 16:11725–11735
Groebe DR, Uhlenbeck OC (1989) Thermal stability of RNA hairpins containing a four-membered loop and a bulge nucleotide. Biochemistry 28:742–747
Grosjean H, Söll D, Crothers DM (1976) Studies of the complex between transfer RNAs with complementary anticodons. I. Origins of enhanced affinity between complementary triplets. J Mol Biol 103:499–519
Gutell RG, Weiser B, Woese CR, Noller HF (1985) Comparative anatomy of 16S-like ribosomal RNA. Prog Nucleic Acid Res Mol Biol 32:155–216
Hall K, Cruz P, Tinoco I Jr, Jovin TM, van de Sande JH (1984) 'Z-RNA'--a left-Handed RNA double helix. Nature 311:584–586
Holbrook SR, Sussman JL, Warrant RW, Church GM, Kim S-H (1977) RNA ligand interactions: (I) Magnesium binding sites in yeast $tRNA^{Phe}$. Nucleic Acids Res 4:2811–2820
Holbrook SR, Sussman JL, Warrant RW, Kim S-H (1978) Crystal structure of yeast phenylalanine transfer RNA. II. Structural features and functional implications. J Mol Biol 123:631–660
Hou Y-M, Schimmel P (1988) A simple structural feature is a major determinant of the identity of a transfer RNA. Nature 333:140–145
Huang WM, Ao S-Z, Casjeans S, Orlandi R, Zeikus R, Weiss R, Winge D, Fang M (1988) A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. Science 239:1005–1012
Hutchins CJ, Rathjen PD, Forster AC, Symons RH (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. Nucleic Acids Res 14:3627–3640
Jack A, Ladner JE, Rhodes D, Brown RS, Klug A (1977) A crystallographic study of metal-binding to yeast phenylalanine transfer RNA. J Mol Biol 111:315–328
James BD, Olsen GJ, Liu J, Pace NR (1988) The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. Cell 52:19–26
Keese P, Symons RH (1985) Domains in viroids: Evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. Proc Natl Acad Sci USA 82:4582–4586
Kim S-H, Cech TR (1987) Three-dimensional model of the active site of the self-splicing rRNA precursor of *Tetrahymena*. Proc Natl Acad Sci USA 84:8788–8792
Labuda D, Pörschke D (1980) Multistep mechanism of codon recognition by transfer ribonucleic acid. Biochemistry 19:3799–3805

Letai AG, Palldino MA, Fromm E, Rizzo V, Fresco JR (1988) Specificity in formation of triple-stranded nucleic acid helical complexes: Studies with agarose-linked polyribonucleotide affinity columns. Biochemistry 27:9108–9112
Martin FH, Castro MM, Aboul-ela F, Tinoco I Jr (1985) Base pairing involving deoxyinosine: implications for probe design. Nucleic Acid Res 13:8927–8938
Massoulié J (1968) Thermodynamique des associations de poly A et poly U en milieu neutre et alcalin. Eur J Biochem 3:428–438
McClain WH, Chen Y-M, Foss K, Schneider J (1988) Association of transfer RNA acceptor identity with a helical irregularity. Science 242:1681–1684
McPheeters DS, Stormo GD, Gold L (1988) Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA. J Mol Biol 201:517–535
Moazed D, Stern S, Noller HF (1986) Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. J Mol Biol 187:399–416
Nussinov R, Jacobson A (1980) Fast algorithm for predicting the secondary structure of single stranded RNA. Proc Natl Acad Sci 77:6309–6313
Peattie DA, Douthwaite S, Garrett RA, Noller HF (1981) A "bulged" double helix in a RNA-protein contact site. Proc Natl Acad Sci USA 78:7331–7335
Pleij CWA, Rietveld K, Bosch L (1985) A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res 13:1717–1731
Privalov PL, Filimonov VV (1978) Thermodynamic analysis of transfer RNA unfolding. J Mol Biol 122:447–464
Puglisi JD (1989) RNA folding: structure and conformational equilibria of RNA pseudoknots. Ph.D. Thesis, University of California, Berkeley
Puglisi JD, Wyatt JR, Tinoco I Jr (1988) A pseudoknotted RNA oligonucleotide. Nature 321:283–286
Riesner D, Römer R (1971) Thermodynamics and kinetics of conformational transitions in oligonucleotides and tRNA. In: Duchesne J (ed) Physico-chemical properties of nucleic acids. Academic Press, New York, pp 237–318
Riesner D, Maass G, Thiebe R, Philippsen P, Zachau HG (1973) Conformational transitions in yeast $tRNA^{Phe}$ as studied with tRNA fragments. Eur J Biochem 36:76–88
Rietveld K, van Poelgeest R, Pleij CWA, van Boom JH, Bosch L (1982) The tRNA-like structure at the 3′ terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. Nucleic Acids Res 10:1929–1946
Romby P, Westhof E, Toukifimpa R, Mache R, Ebel J-P, Ehresmann C, Ehresmann B (1988) Higher order structure of chloroplastic 5S ribosomal RNA from spinach. Biochemistry 27:4721–4730
Saenger W (1984) Principles of nucleic acid structure. Springer, Berlin Heidelberg New York Tokyo
Sampson JR, DiRenzo AB, Behlen LS, Uhlenbeck OC (1989) Nucleotides in yeast $tRNA^{Phe}$ required for the specific recognition by its cognate synthetase. Science 243:1363–1366
Schimmel P (1989) RNA pseudoknots that interact with components of the translation apparatus. Cell 58:9–12
Stahl DA, Walker TA, Meyhack B, Pace NR (1979) Precursor-specific nucleotide sequences can govern RNA folding. Cell 18:1133–1143
Stern S, Weiser B, Noller HF (1988) Model for the three-dimensional folding of 16S ribosomal RNA. J Mol Biol 204:447–481
Studnicka GM, Rahn GM, Cummings IW, Salzer WA (1978) Computer method for predicting the secondary structure of single-stranded RNA. Nucleic Acids Res 5:3365–3387
Sugimoto N, Kierzek R, Freier SM, Turner DH (1986) Energetics of internal GU mismatches in ribooligonucleotide helixes. Biochemistry 25:5755–5759
Tang CK, Draper DE (1989) Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. Cell 57:531–536
Thiele D, Guschlbauer W (1971) Protonated polynucleotide structures. IX. Disproportionation of poly(G)·poly(C) in acid medium. Biopolymers 10:143–157
Tuerk C, Gauss P, Thermes C, Groebe DR, Guild N, Stormo G, Gayle M, d'Auberton-Carafa Y, Uhlenbeck OC, Tinoco I Jr, Brody EN, Gold L (1988) CUUCGG hairpins: extraor-

dinarily stable RNA secondary structure associated with various biochemical processes. Proc Natl Acad Sci USA 85:1364–1368
Turner DH, Sugimoto N, Freier SM (1988) RNA structure prediction. Annu Rev Biophys Biophys Chem 17:167–192
Uhlenbeck OC (1972) Complementary oligonucleotide binding to transfer RNA. J Mol Biol 65:25–41
Uhlenbeck OC (1987) A small catalytic oligoribonucleotide. Nature 328:596–600
van den Hoogen YT, van Beuzekom AA, de Vroom E, van der Marel G, van Boom JH, Altona C (1988) Bulge-out structures in the single-stranded trimer AUA and in the duplex (CUGGUGCGG)·(CCGCCCAG). A model-building and NMR study. Nucleic Acids Res 16:5013–5030
van de Sande JH, Ramsing NB, Germann MW, Elhorst W, Kalisch BW, v Kitzing E, Pon RT, Clegg RC, Jovin TM (1988) Parallel stranded DNA. Science 241:551–557
Varani G, Wimberly B, Tinoco I Jr (1989) Conformation and dynamics of an RNA internal loop. Biochemistry 28:7760–7772
Waring RB, Davies RW (1984) Assessment of a model for intron secondary structure relevant to RNA self-splicing – a review. Gene 28:277–291
Wells RD, Collier DA, Hanvey JC, Shimuzu M, Wohlrab F (1988) The chemistry and biology of unusual DNA structures adopted by oligopurine·oligopyrimidine sequences. FASEB J 2:2939–2949
White SA, Draper DE (1987) Effects of single-base bulges on intercalator binding to small RNA and DNA hairpins and a ribosomal RNA fragment. Biochemistry 28:1892–1897
Williams AL Jr, Tinoco I Jr (1986) A dynamic programming algorithm for finding alternative RNA secondary structures. Nucleic Acids Res 14:299–315
Woodson SA, Crothers DM (1989) Conformation of a bulge-containing oligomer from a hot-spot sequence by NMR and energy minimization. Biopolymers 28:1149–1177
Wu H-N, Uhlenbeck OC (1987) Role of a bulged A residue in a specific RNA protein interaction. Biochemistry 26:8221–8227
Wyatt JR (1990) RNA pseudoknots: effect of loop sequence and size on conformational equilibria. Ph.D. Thesis, University of California, Berkeley
Wyatt JR, Puglisi JD, Tinoco I Jr (1989) Pseudoknotted RNA oligonucleotides. In: Cech T (ed) Molecular biology of RNA. Alan R Liss, New York, pp 25–32
Zarling DA, Calhoun CJ, Hardin CC, Zarling AH (1987) Cytoplasmic Z-RNA. Proc Natl Acad Sci 84:6117–6121
Zimmerman SB, Cohen GS, Davies DR (1975) X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid. J Mol Biol 92:181–192
Zuker M (1989) On finding all suboptimal foldings of an RNA molecule. Science 244:48–52
Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 9:133–148