17

Thermodynamic Considerations for Evolution by RNA

Douglas H. Turner and Philip C. Bevilacqua
Department of Chemistry
University of Rochester
Rochester, New York 14627-0216

The only definitive statement that thermodynamics makes about evolution is that ultimately it will result largely in CO_2, H_2O, and N_2. Fortunately, the sun allows us to live in a world that is kinetically controlled due to an input of energy. Nevertheless, since many equilibria are reached rapidly compared with the age of the earth, thermodynamic principles can be used to predict what reactions are possible in this transitory world. In this chapter, we argue that known thermodynamic properties of RNA suggest it is well adapted to play a central role in the evolution of information transfer and catalysis. This is because mononucleotides contain many functional groups and are therefore capable of strong, specific binding interactions. The evidence for the strength of the binding interactions comes largely from optical melting studies of oligonucleotides (Turner et al. 1988). Since evolution presumably started with rather short molecules, oligonucleotides are reasonable model systems for studying the interactions important in early times. In the following, we discuss the interactions thought to be important for organizing RNA, describe some implications for catalysis and information transfer, and speculate on implications for evolution.

DISCUSSION

Fundamental Interactions Determining RNA Structure

Molecular association and organization are essential for information transfer and catalysis. The thermodynamic measure of the strength of an association or the extent of organization is the change in free energy, ΔG°, since the equilibrium constant for either association or organization into a specific conformation relative to the random coil is given by $K = e^{-\Delta G^\circ/RT}$. Here R is 1.987 cal/M·K and T is temperature in kelvins. Thus, the more negative the free-energy change, the more favorable the association or the given conformation. For example, at 37°C (310.15 K),

The RNA World

every decrease of 1.4 kcal/mole in free energy, ΔG°_{37}, favors that conformation or association by a factor of 10 increase in the equilibrium constant. As discussed below, individual stacking and hydrogen-bonding interactions can produce changes of this magnitude.

Stacking Interactions

Stacking is the vertical interaction of flat aromatic ring systems, as cartooned in Figure 1. One measure of the strength of stacking interactions is the degree to which homopolymers like polyadenylic acid organize from random coil to stacked, helical conformations. This has been studied by optical, calorimetric, and other methods (Richards et al. 1963; Inners and Felsenfeld 1970; Suurkuusk et al. 1977; Filimonov and Privalov 1978; Freier et al. 1981). The ΔG°_{37} for this intrastrand stacking is both modest and sequence dependent. For polycytidylic, polyadenylic, and polyuridylic acids, the ΔG°_{37} values are −0.3, −0.02, and >+1 kcal/mole, respectively (Richards et al. 1963; Inners and Felsenfeld 1970; Freier et al. 1981). The positive ΔG°_{37} for polyuridylic acid indicates negligible stacking at 37°C.

Another measure of the strength of stacking interactions is the enhanced association constant observed when unpaired nucleotides, dangling ends, are added to the 3′ ends of RNA duplexes (Fig. 1). This can be determined by measuring optical melting curves for duplexes with and without dangling ends (Freier et al. 1983; Petersheim and Turner 1983; Sugimoto et al. 1987; Turner et al. 1988). The interstrand stacking effect can be large. For example, 3′ dangling purines, R, in the sequence $^{CR}_{G}$ make ΔG°_{37} for oligomer association more favorable by 1.7 kcal/mole on average. Thus, a 3′ dangling purine can make a substantial contribution to stability without any requirement for additional complementarity. One trace of an early use of this interaction may be the fact that a 3′ dangling purine typically follows the anticodon in tRNA. The effect is very sequence dependent, however. The ΔG°_{37} increment from a 3′ dangling pyrimidine in the sequence $^{UY}_{A}$ is only −0.1 kcal/mole.

In contrast to 3′ dangling ends, all 5′ dangling ends in RNA add little additional stability, on average making duplex formation more favorable by only 0.2 kcal/mole at 37°C. This is expected from the geometry of an A-form RNA helix, which places a 5′ dangling end away from the opposite strand (Freier et al. 1985).

Hydrogen-bonding Interactions

It is clear that hydrogen bonds are important for providing specificity in the binding and folding of macromolecules. The contributions of hydrogen bonds to the free-energy changes for binding and folding have been controversial, however (Abeles et al. 1992). This is because formation of hydrogen bonds in binding and folding must be accompanied by breakage of hydrogen bonds to water. For RNA, thermodynamic studies suggest the net contribution of a hydrogen bond to binding or folding is

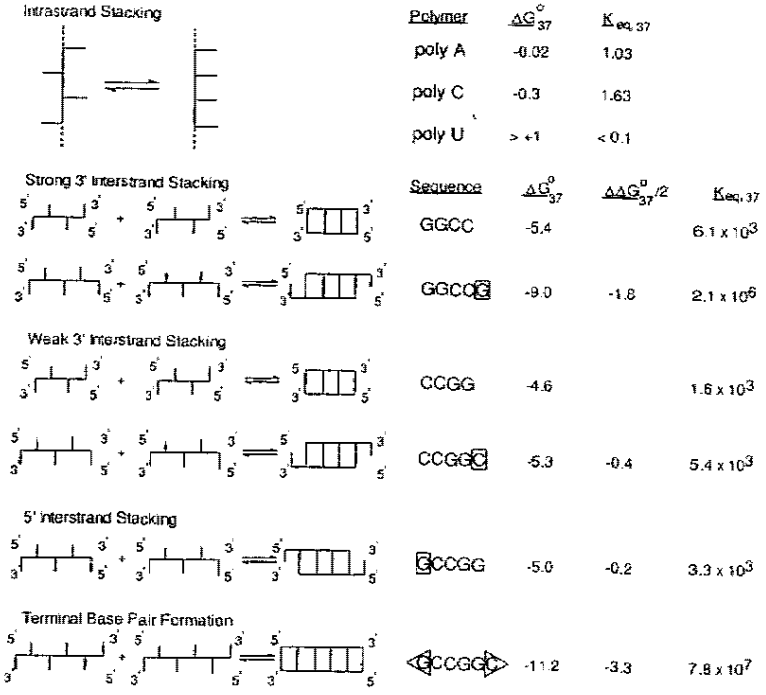


Figure 1 Some examples of free-energy changes, ΔG°_{37}, and free-energy increments, $\Delta\Delta G^{\circ}_{37}/2$ for stacking and base-pairing at 37°C. For intrastrand stacking in homopolymers, ΔG°_{37} values come from combining optical and calorimetric data (Freier et al. 1981) or from the absence of any observable stacking in polyuridylic acid (Richards et al. 1963). For interstrand stacking, $\Delta\Delta G^{\circ}$ values come from subtracting the ΔG°_{37} for duplex formation by a fully Watson-Crick-paired core from the ΔG°_{37} for duplex formation by the core sequence with an added nucleotide that does not pair. For example, $\Delta\Delta G^{\circ}_{37}$ ($2\ ^{CG}_{G}$) = ΔG°(GGCCG) − ΔG°(GGCC). The unpaired nucleotide is boxed in the sequence. Also shown is the free-energy increment for adding two terminal CG pairs to a $(CCGG)_2$ core. Note that in all cases, the $\Delta\Delta G^{\circ}_{37}$ has been divided by two to give the free-energy increment per end.

context-dependent (SantaLucia et al. 1992), but can be significant (Freier et al. 1985, 1986b; Turner et al. 1987). The evidence is discussed below.

The first indication that hydrogen bonds contribute to the ΔG° values for interactions between RNA oligomers comes from the changes in free energy of duplex formation when an oligomer with a dangling end is extended by one nucleotide to allow the previously unpaired dangling end to form a base pair (Freier et al. 1985, 1986b). One example is illustrated in Figure 1. For this case at 37°C, the duplex with two terminal CG pairs, $(GCCGGC)_2$, is 5.9 kcal/mole more stable than the duplex with two terminal 3′ dangling ends, $(CCGGC)_2$ (Freier et al. 1985). Most of this extra stability depends on the complementarity of the 5′ and 3′ terminal residues, since the duplex with two terminal AC mismatches, $(ACCGGC)_2$, is only 1.4 kcal/mole more stable than $(CCGGC)_2$ (Hickey and Turner 1985). This suggests the hydrogen bonds in the terminal CG pairs are each worth at least (5.9–1.4)/6 = 0.8 kcal/mole. A more refined estimate for the ΔG° of a hydrogen bond requires a correction for conformational entropy effects. This gives a value of 1.6 kcal/(mole hydrogen bond) for this system (Freier et al. 1986b).

Another approach to estimating the contribution of a hydrogen bond to binding or folding is to measure the free-energy consequences of replacing a single hydrogen-bonding group with a hydrogen (Turner et al. 1987). This "atomic mutation" approach is made possible by advances in the chemical synthesis of RNA (Kierzek et al. 1986; Usman et al. 1987). The results from several studies involving substitutions of single functional groups are illustrated in Figure 2. The changes in free energy upon substituting hydrogen for a single hydrogen-bonding group range from 0 to 1.6 kcal/mole. This range reflects different contexts for the hydrogen bonds. For example, substituting the 5′-terminal Gs of $(GCCGGC)_2$ with inosine replaces the 5′-terminal G amino groups with hydrogens (Fig. 2). This reduces duplex stability by 3.2 kcal/mole or 1.6 kcal/(mole hydrogen bonds), roughly as expected from the comparison of dangling ends and terminal mismatches described above (Turner et al. 1987). Substituting purine for A in GA mismatches replaces the amino group of A with a hydrogen. As shown in Figure 2, this substitution reduces duplex stability by 1.3–1.5 kcal/(mole mismatch) when two GA mismatches are in the center of a helix and by 0.4 kcal/mole for a GA mismatch in a 4-nucleotide hairpin, but it has no effect for a GA mismatch at the end of a helix (SantaLucia et al. 1991a, 1992). The reasons for this range are not clear. One possibility is that loss of a hydrogen bond at the end of a helix is compensated by enhanced conformational freedom. Clearly, however, a single hydrogen bond between bases can contribute on the order of 2 kcal/mole to binding or folding.

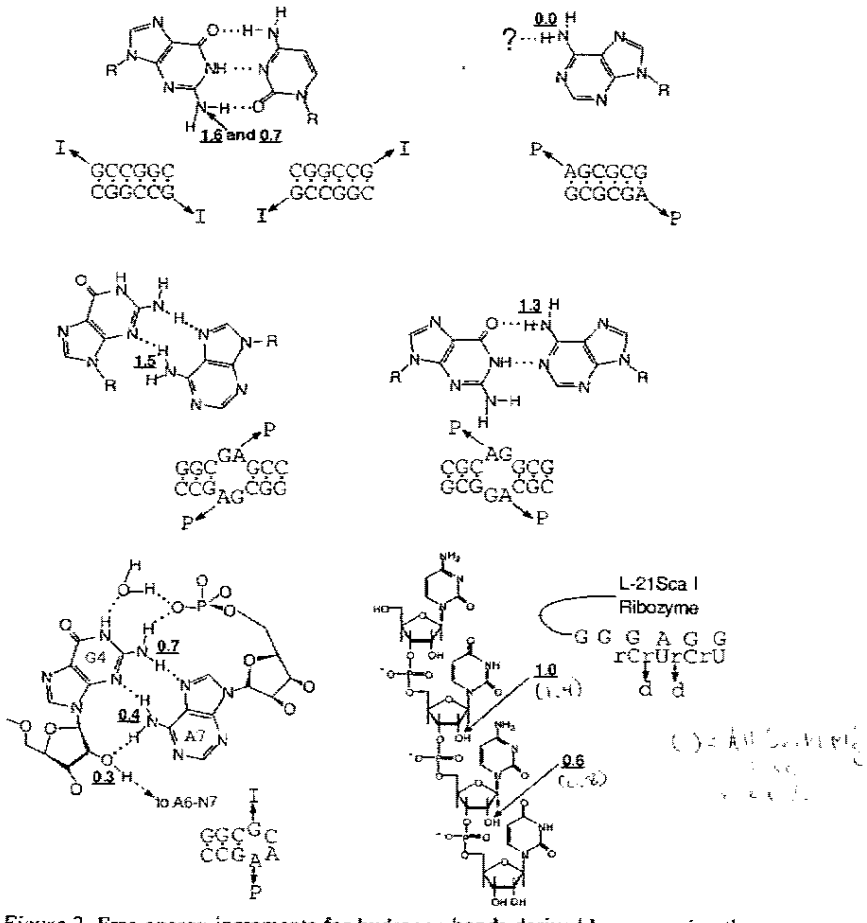


Figure 2 Free-energy increments for hydrogen bonds derived by measuring the thermodynamic effect of substituting H for a functional group. Substitutions are denoted by arrows: I = inosine, P = purine, d = deoxyribose. Changes observed in binding or folding free energy (kcal/mole) are listed in bold, underlined numbers near the functional group that was substituted. Changes in duplex formation (*top* 4 structures) are at 37°C (Turner et al. 1987; SantaLucia et al. 1991a). Changes in hairpin stability (*bottom left*) are at 70°C and are about 0.1 kcal/mole smaller at 37°C (SantaLucia et al. 1992). Changes in ribozyme binding are at 15°C (Bevilacqua and Turner 1991).

Hydrogen-bonded interactions in RNA are not restricted to the bases. It has been shown by kinetic (Sugimoto et al. 1989), gel retardation (Pyle and Cech 1991), and equilibrium dialysis (Bevilacqua and Turner 1991) methods that binding of substrates to a group I ribozyme is partially dependent on particular 2′-OH groups. This provides an advantage for folding and catalytic diversity not possessed by DNA. For example, comparisons of the effects on binding of substrate CUCU to the ribozyme or to an oligomer after replacing substrate 2′-OH groups with hydrogens suggest interactions with two of the 2′-OH groups contribute about 1 kcal/mole each to binding free energy at 15°C (see Fig. 2) (Bevilacqua and Turner 1991). Additionally, 2′-OH group interactions might also help catalyze reactions without affecting binding in the ground state. For example, substitution of H for the 2′ OH of the reactive nucleotide of substrates for two different forms of a group I intron hinders reactivity (Sugimoto et al. 1989; Bevilacqua and Turner 1991 and unpubl.) but does not affect ground-state binding. Thus, a 2′-OH can enhance binding in both ground and transition states.

Presumably, hydrogen bonds to phosphate groups are also important for RNA binding and folding (Quigley and Rich 1976). Thus, a single nucleotide has many contact points that can provide significant contributions to free-energy changes for binding and folding.

Counterion Condensation

A unique property of nucleic acids is their high density of negative charge. This leads to condensation of neutralizing counterions into a relatively small volume around the nucleic acid chain, producing a high local concentration of counterions that is independent of the counterion concentration in the bulk solvent (Manning 1978; Record et al. 1978). This concentrating of counterions around the helix is entropically unfavorable, but less so as the counterion concentration in the bulk solvent increases. This underlies the common linear dependence of duplex melting temperature, T_M, as a function of log [Na^+]. Another consequence of counterion condensation is that multivalent ions will displace essentially all univalent cations around the chain. Thus, the local volume around a nucleic acid is enriched in multivalent cations. This guarantees a ready source of the metal ions required for catalysis. The details of counterion condensation are relatively well understood for polymeric nucleic acids (Manning 1978; Record et al. 1978). For example, the local concentration of +2 cations near a polymer double helix is on the order of 0.4 M, independent of the bulk concentration of +2 cations (Manning 1978). Studies of the salt dependence of the T_M and the association rate for the hexamer dGCATGC suggest that counterion condensation is also important for oligomers (Williams et al. 1989). The theoretical framework for a quantitative description of the effect with oligomers is currently being developed (Olmsted et al. 1989).

Thermodynamic Advantages for RNA as the First Macromolecular Catalyst and Information Carrier

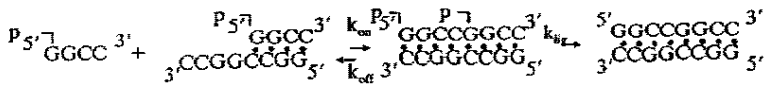
A prerequisite for catalysis and information transfer is strong binding. For catalysis, simply bringing reactants together into the correct position for reaction can enhance rates by factors of up to 10^8 (Fersht 1985; Jencks 1987; Abeles et al. 1992). We have argued above that individual stacking and hydrogen-bonding interactions in RNA can each contribute on the order of 2 kcal/mole to the ΔG^o for binding at 37°C. This value is even more favorable at low temperature, especially for the stacking component. Thus, the most favorable stacking interaction for a 3′ dangling end contributes −2.6 kcal/mole at 0°C, enhancing a binding constant by a factor of more than 100 (Turner et al. 1988). This interaction would favor chain extension in a 5′ to 3′ direction, since a nucleotide coming onto the 3′ end of an RNA chain would be held tighter than one coming onto the 5′ end where stacking is negligible. In addition to stacking surfaces, of course, a single nucleotide has many hydrogen-bonding groups that can also enhance binding.

In contrast to nucleotides, the number of strong interactions available to a single amino acid is much smaller. This is reflected in the number of monomers required to form ordered structures in oligonucleotides and oligopeptides. Many dinucleotides have equilibrium constants of about 1 for stacking in a helical manner (Powell et al. 1972; Olsthoorn et al. 1981); the equilibrium constant for hairpin formation by $_{GGC}{}^{GCAA}{}_{GCC}$ at 0°C is 1×10^5 (SantaLucia et al. 1992). Even for optimized cases, however, an oligopeptide of 14 amino acids is required to get an equilibrium constant of order 1 for α-helix formation at 3°C (Shoemaker et al. 1987). The dissociation constant for duplex formation by GGCC at 0°C is 62 nM (Freier et al. 1983). The dissociation constants for assembly of short peptide strands have not been measured yet, although assembly of covalently linked fragments has been demonstrated (Oas and Kim 1988). Note that formation of GGCC from monomers requires only three coupling reactions; from dimers it requires only one coupling. In a prebiotic world with coupling reactions difficult, nucleic acids have a clear thermodynamic advantage over proteins. Few coupling reactions are required to allow strong associations into ordered assemblies.

Another thermodynamic-based advantage for evolution of nucleic acids is the large temperature dependence for binding constants between

nucleic acids. For example, before enzymes evolved, RNA elongation may have been catalyzed by bringing two substrates close via a complementary template. The large temperature dependence of these associations could give rise to a catalytic cycle where low temperatures favored binding for efficient elongation of short oligomers, but high temperatures allowed efficient release of the longer products. Effectively, chain growth becomes a two-step cycle, perhaps regulated by day and night.

Consider the following simple mechanism for the first step of the cycle in which two identical tetramers, GGCC, bind a complementary template, GGCCGGCC, and ligate to form a template/product complex:



For short oligomers, K_m (=$(k_{lig} + k_{off})/k_{on}$) is large. Since oligomer abundance is presumably low in early stages of evolution, K_m >> $[GGCC]_o$ >> $[GGCCGGCC]_o$. The apparent rate, k_{app}, for conversion of GGCC to the template/product complex is then given by $k_{app} = \{k_{lig}k_{on}/(k_{lig} + k_{off})\}$ $[GGCC]_o$. The rate of association for oligomers making GC pairs, k_{on}, is essentially independent of temperature (Pörschke et al. 1973; Williams et al. 1989), so the large temperature dependence of the binding constant is manifested primarily in the rate of dissociation, k_{off}. For example, the half-life, $t_{1/2}$, for dissociation of the first GGCC and template is predicted to be 5 minutes at 0°C, and 9.4 milliseconds at 37°C (Freier et al. 1986a; Turner et al. 1988). Since ligation is likely to be slow, most template-binding events will not result in reaction (k_{off} >> k_{lig}). Under these "rapid equilibrium" conditions, the ratio of k_{app} at two temperatures reduces to:

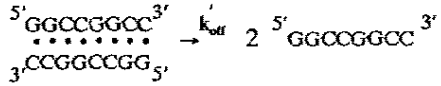
$$\frac{k_{app}^{T_{low}}}{k_{app}^{T_{high}}} = \frac{k_{lig}^{T_{low}}}{k_{lig}^{T_{high}}} \frac{k_{off}^{T_{high}}}{k_{off}^{T_{low}}}$$

The temperature dependence of each rate can be approximated by the Arrhenius equation, $k = A \exp(-E_a/RT)$. This gives:

$$\frac{k_{app}^{T_{low}}}{k_{app}^{T_{high}}} \propto \exp[\frac{1}{R}(\frac{1}{T_{low}} - \frac{1}{T_{high}})(E_{a,off} - E_{a,lig})]$$

Thus, if the temperature dependence of k_{off} is larger than k_{lig} (i.e., $E_{a,off} > E_{a,lig}$), conversion of GGCC to the template/product complex becomes favored at low temperature by this exponential term. Moreover, under such conditions, the template position of tightest substrate binding is also most efficient, adding a primitive degree of fidelity.

After one or more ligations, no sites would be left on the template for substrate binding, necessitating the second step of the cycle: release of the template/product complex.



Kinetic experiments on site-specific cleavage by a group I ribozyme clearly show that product dissociation is rate-limiting under certain conditions (Herschlag and Cech 1990). Presumably, this will be equally or more limiting for the products of ligation reactions on a template. If the length of the product is double that of the substrate as considered above, nearest-neighbor rules indicate ΔG^o will roughly double. As a consequence, $t_{1/2}$ at 0°C for dissociation of the product, GGCCGGCC, from its template becomes 250,000 years! The large temperature dependence of k_{off}, however, permits dissociation by temperature increase; $t_{1/2}$ for dissociation of GGCCGGCC from its template is only 16 hours at 37°C, one warm cycle. In a sense, high temperature could serve as a primitive helicase. Once released, the ligated product could self-assemble and/or undergo further rounds of elongation. Thus, low temperature favors chain elongation of short oligomers, but release of longer products may require high temperature. This is possible because of the large temperature dependence for RNA associations.

It should be noted that template/product formation has an optimum temperature; that is, lower temperature does not *always* favor chain elongation. Temperatures below the optimum disfavor the first step of the cycle for several reasons. Oligomers could bind so tightly to incorrect template positions or positions separated by gaps that a steady-state population of correctly aligned intermediates does not accumulate on a reasonable time scale. Such effects would decrease fidelity and prevent efficient chain elongation. Thus, the optimum temperature for chain elongation depends on both sequence and length. In general, however, release of the elongated product on a reasonable time scale requires an increase above the optimum temperature for chain elongation. It is clear that the strong dependence of k_{off} on both temperature and oligomer length must be considered in discussions of template-mediated coupling of RNA. These considerations suggest that in vitro experiments aimed at simulating evolution might benefit from incorporation of a temperature cycling step.

Thermodynamic Considerations for Evolution of the First Binding Sites and Codes

Even with the many binding contacts available on a nucleotide, initial assembly must have been difficult, given the presumably low concentrations of substrates. This requirement for tight binding may have been one element in the selection of the 3′-5′ sugar-phosphate backbone instead of the 2′-5′ alternative. Short oligomers with 2′-5′ linkages form base pairs, but the ΔG°_{37} per base pair for duplex formation is typically about 1 kcal/mole less favorable than for the corresponding 3′-5′ linked oligomer (Kierzek et al. 1992). Another factor favoring the 3′-5′ linkage is chemical stability (Usher and McHale 1976).

Given a 3′-5′ backbone, the ΔG° values for helix propagation by Watson-Crick base pairs (Freier et al. 1986a) suggest that GC pairs would be favored, since on average, adjacent GC pairs contribute 2.2 kcal/mole more than adjacent AU pairs at 0°C. The ΔG° values for GU pairs are similar to those for AU pairs (He et al. 1991). Thus, there is no thermodynamic requirement for A in early helices. Interestingly, anticodons containing A typically code for amino acids that may not have been necessary at early times: Phe, Tyr, Trp, Cys, Met, Leu, Ile, and Val. Absence of A, however, would give replication in which U could substitute for C, although C would still be favored due to stronger binding.

Formation of simple helices, of course, is probably not sufficient to produce a range of catalytic activities. Presumably, base functional groups must be available to bind substrates in particular orientations and to provide catalysis. This focuses attention on motifs that are not Watson-Crick paired. Unfortunately, most such motifs destabilize a duplex and therefore require more Watson-Crick base pairs to hold it together. The only known exceptions to this general rule involve insertions of single and double GU mismatches into helices (He et al. 1991). Some mismatches are less destabilizing than others, however. For example, recent studies of internal loops containing two mismatches show that insertion of $^{GA}_{AG}$, $^{AG}_{GA}$, and $^{UU}_{UU}$ double mismatches into a helix destabilizes the fully Watson-Crick-paired duplex much less than other non-GU double mismatches (SantaLucia et al. 1991b). Figure 3 shows the change in ΔG°_{37} measured upon insertion of these and other double mismatches into a helix. The thermodynamic advantage for insertion of the GA and UU double mismatches averages 2.1 kcal/mole at 37°C. NMR studies suggest the enhanced stabilities of loops with GA, UU, and GU mismatches are due to hydrogen bonding (He et al. 1991; SantaLucia et al. 1991b; Nikonowicz and Pardi 1992). If short helices were initially preferred due to difficulty in coupling nucleotides, then the thermodynamic results suggest internal loops with GA, UU, and GU mismatches would also be preferred. (Presumably an isolated AU or GC pair in the middle of a loop could also be favorable.)

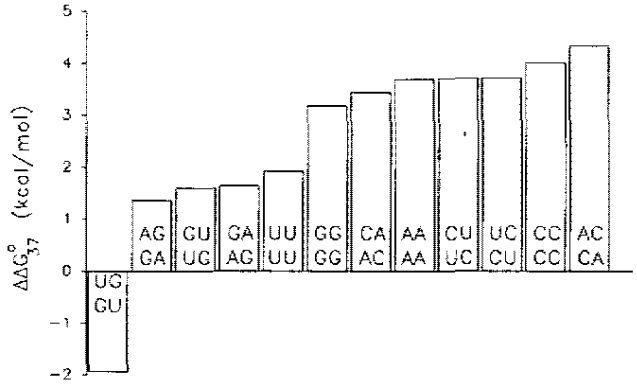


Figure 3 Change in free energy of duplex formation at 37°C upon inserting double mismatches into the middle of a helix. For example, $\Delta\Delta G^{\circ}_{37}$ ($^{AG}_{GA}$) = ΔG°_{37}(CGC<u>AG</u>GCG) − ΔG°_{37} (CGCGCG) (SantaLucia et al. 1991b). Positive values of $\Delta\Delta G^{\circ}_{37}$ are unfavorable for duplex formation.

Modeling of the $^{GA}_{AG}$ mismatch, as shown in Figure 4, suggests even a small internal loop can provide a useful recognition element (SantaLucia 1991). Each GA mismatch is hydrogen-bonded as shown in Figure 2 in a pattern initially seen in DNA by Li et al. (1991) and in RNA by Heus and Pardi (1991). This leaves normal Watson-Crick hydrogen-bonding positions available on G for binding something into the pocket that is formed. Interestingly, the $^{AG}_{GA}$ internal loop has a hydrogen-bonding pattern involving the G imino and carbonyl groups (Fig. 2). Thus, it has fewer functional groups available for further binding. In a survey of 21 16S rRNAs (Gutell et al. 1985) and 38 23S rRNAs (Gutell and Fox 1988), the $^{GA}_{AG}$ motif occurs 45 times, whereas the $^{AG}_{GA}$ motif is absent. Since $^{GA}_{AG}$ and $^{AG}_{GA}$ are thermodynamically similar, this suggests $^{GA}_{AG}$ was selected for its binding possibilities.

Thermodynamic considerations also favor placing mismatches between helices instead of at the ends in order to optimize their effectiveness for binding and catalysis. Mismatches flanked by base pairs have less conformational freedom than mismatches at the ends of helices. This results in greater preorganization of the binding site and hence tighter binding. Moreover, the local Mg^{++} concentration should be greater in the middle than at the end of a helix (Olmsted et al. 1989).

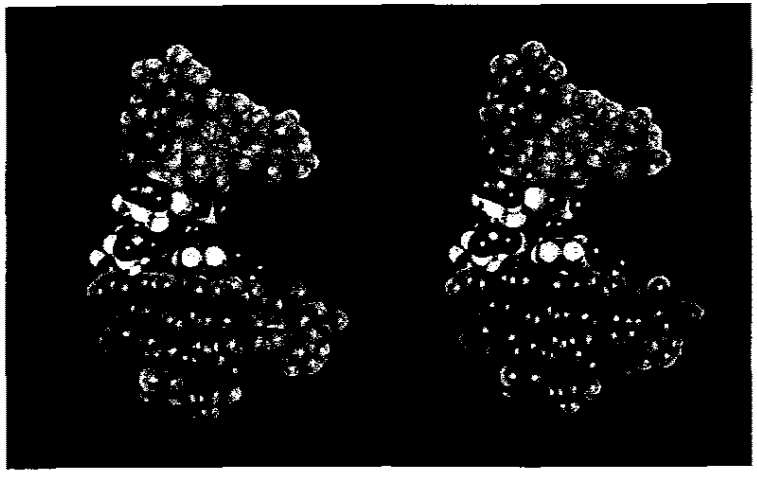


Figure 4 Model of the duplex $\begin{matrix} & & GA & & \\ GGC & & & & GCC \\ CCG & & & & CGG \\ & & AG & & \end{matrix}$ based on NMR data (SantaLucia 1991).

Base pairs are colored green and the GA mismatches have traditional colors for atoms (O = red, H = white, C = black, N = blue).

Hairpin loops provide another unpaired motif that can be added to an RNA helix while favoring folding. For example, the free-energy penalty for initiating a helix from two separate strands is 3.4 kcal/mole at 37°C (Freier et al. 1986a). If the helix is initiated intramolecularly, giving rise to a hairpin loop of 4, 5, or 7 Us, then the initiation penalty is 3.0, 3.1, or 3.8 kcal/mole, respectively (Groebe and Uhlenbeck 1988). Some loops of 4, called tetraloops, are even less destabilizing (Tuerk et al. 1988). For example, Antao and Tinoco (1992) report a penalty of only 1.0 kcal/mole for initiating a loop at 37°C with the sequence UUCG. In addition to the relatively favorable ΔG^o for initiation of a hairpin, hairpin folding is further favored by the transformation of the folding from an intermolecular to an intramolecular process. For the intramolecular folding of a hairpin, half the molecules have a base-paired stem when $\Delta G^o_{\text{hairpin}} = 0$ kcal/mole. For helix formation by two self-complementary strands, a ΔG^o of 0 kcal/mole results in half the strands forming duplex helix only when the strand concentration is 1 M. For more realistic strand concentrations, the ΔG^o must be much more favorable. For self-complementary duplexes, the ΔG^o required to drive half the strands into duplex is given by $\Delta G^o = RT \ln C_T$, where C_T is the total strand concentration. Thus, at 10^{-6} M strand concentration, a ΔG^o of −8.5 kcal/mole is required to drive half the strands into duplex at 37°C. Evidently, there was a strong thermodynamic pressure to select hairpins at an early stage in evolution.

Assembling Motifs

The above discussion focused largely on thermodynamic considerations for the coupling and interaction of short oligonucleotides. These could readily fold into well-known elements of secondary structure such as hairpins and internal loops. At some point, however, it must have been advantageous to assemble two or more motifs into more complicated structures. As with base-pairing, such assembly is facilitated when the process is intramolecular. Moreover, the same stacking and hydrogen-bonding interactions that drive base-pairing can also promote assembly of motifs. This can be seen in the crystal structure of tRNA, where such interactions fix the relative positions of the D and TΨC loops (Kim et al. 1974; Robertus et al. 1974). It is likely, however, that additional interactions are also important in this tertiary folding process.

One tertiary interaction unique to RNA is hydrogen bonding to the 2′-OH group. The role of a 2′-OH group in stabilizing a helix was first suggested by kinetic experiments using chimeric DNA-RNA substrates to open a covalently closed, circular group I intron (Sugimoto et al. 1989). Subsequent gel retardation (Pyle and Cech 1991) and equilibrium dialysis (Bevilacqua and Turner 1991) studies suggest the interaction of a 2′-OH group can contribute on the order of 1 kcal/mole to helix stability. Proposed three-dimensional structures of the group I intron suggest this 2′-OH interaction helps hold and orient the substrate helix in the active site (Kim and Cech 1987; Michel and Westhof 1990). Comparison of the effects of substituting a 2′OH with a 2′H on binding to ribozyme and to a hexamer suggest additional, as yet unidentified interactions are also involved in this helix positioning (Bevilacqua and Turner 1991). It has also been shown that the intramolecular rate for this helix positioning is about 3 s^{-1} at 15°C, much slower than the rate for formation of secondary structure (Bevilacqua et al. 1992). This might be expected if, in fact, secondary structure initially evolved in the absence of tertiary structure.

An electric dichroism study of bundling of DNA helices also suggests there are new interactions involved in helix packing (Mandelkern et al. 1981). This study showed that seven DNA helices, each of about 140 base pairs, can pack together side by side, perhaps with six helices hexagonally surrounding a center helix. Counterintuitively, this association is promoted at low concentrations of monovalent counterion, sug-

gesting it is accompanied by release of Na^+. The assembly is also promoted by +3 ions. Although no equivalent study has been done on RNA, it would not be surprising to find similar effects.

Assembly of helices could have important implications for RNA catalysis by providing a region of low dielectric constant and low water activity. Such a milieu is often important for catalysis by proteins (Abeles et al. 1992). The proposed three-dimensional structures of group I introns contain such a catalytic core (Kim and Cech 1987; Michel and Westhof 1990). Moreover, mapping experiments with hydroxyl radicals provide experimental evidence that substantial regions of the ribozyme are inaccessible to solvent after tertiary folding is induced by Mg^{++} (Celander and Cech 1991).

CONCLUSIONS AND PERSPECTIVES

The many functional groups, high charge density, and strong interactions of RNA provide thermodynamic advantages for RNA in initial stages of evolution. In particular, relatively few coupling reactions are required to form an oligomer capable of tight binding and stable folding. Many of the interactions important for determining secondary and tertiary structure are known, but additional interactions remain to be discovered. The products of evolution were restricted by the building blocks available and their environment. Recent studies of rationally designed, organic replicating systems suggest that similar interactions can result in self-assembly of different building blocks in organic solvents (Tjivikua et al. 1990; Feng et al. 1992). In vitro selection methods have been used to identify nucleic acid sequences capable of specific functions (Ellington and Szostak 1990, 1992; Tuerk and Gold 1990). This suggests similar methods may ultimately be developed for use with different building blocks, thus expanding the possibilities available. Thus, understanding intermolecular interactions may facilitate rational design of a replicating system that can be used to evolve the best structures available for a particular catalytic or other purpose.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant GM-22939.

REFERENCES

Antao, V.P. and I. Tinoco, Jr. 1992. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res.* **20:** 819–824.

Bevilacqua, P.C. and D.H. Turner. 1991. Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: Evidence for ribozyme specific interactions with 2′OH groups. *Biochemistry* **30:** 10632–10640.

Bevilacqua, P.C., R. Kierzek, K.A. Johnson, and D.H. Turner. 1992. Dynamics of ribozyme binding of substrate revealed by fluorescence detected stopped-flow. *Science* **258:** 1355–1358.

Celander, D.W. and T.R. Cech. 1991. Visualising the higher order folding of a catalytic RNA molecule. *Science* **251:** 401–407.

Ellington, A.D. and J.W. Szostak. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346:** 818–822.

———. 1992. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355:** 850–852.

Feng, Q., K. Park, and J. Rebek, Jr. 1992. Crossover reactions between synthetic replicators yield active and inactive recombinants. *Science* **256:** 1179–1180.

Fersht, A. 1985. *Enzyme structure and mechanism*, chapter 2. W.H. Freeman, New York.

Filimonov, V.V. and P.L. Privalov. 1978. Thermodynamics of base interaction in $(A)_n$ and $(A^*U)_n$. *J. Mol. Biol.* **122:** 465–470.

Freier, S.M., D. Alkema, A. Sinclair, T. Neilson, and D.H. Turner. 1985. Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helixes. *Biochemistry* **24:** 4533–4539.

Freier, S.M., B.J. Burger, D. Alkema, T. Neilson, and D.H. Turner. 1983. Effects of 3′ dangling end stacking on the stability of GGCC and CCGG double helices. *Biochemistry* **22:** 6198–6206.

Freier, S.M., K.O. Hill, T.G. Dewey, L.A. Marky, K.J. Breslauer, and D.H. Turner. 1981. Solvent effects on the kinetics and thermodynamics of stacking in poly(cytidylic acid). *Biochemistry* **20:** 1419–1426.

Freier, S.M., R. Kierzek, J.A. Jaeger, N. Sugimoto, M.H. Caruthers, T. Neilson, and D.H. Turner. 1986a. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci.* **83:** 9373–9377.

Freier, S.M., N. Sugimoto, A. Sinclair, D. Alkema, T. Neilson, R. Kierzek, M.H. Caruthers, and D.H. Turner. 1986b. Stability of XGCGCp, GCGCYp, and XGCGCYp helixes: An empirical estimate of the energetics of hydrogen bonds in nucleic acids. *Biochemistry* **25:** 3214–2319.

Groebe, D.R. and O.C. Uhlenbeck. 1988. Characterization of RNA hairpin loop stability. *Nucleic Acids Res.* **16:** 11725–11735.

Gutell, R.R. and G.E. Fox. 1988. Compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res.* (suppl.) **16:** 175–269.

Gutell, R.R., B. Weiser, C.R. Woese, and H.F. Noller. 1985. Comparative anatomy of 16-S-like ribosomal RNA. *Prog. Nucleic Acids Res. Mol. Biol.* **32:** 155–216.

He, L., R. Kierzek, J. SantaLucia, Jr., A.E. Walter, and D.H. Turner. 1991. Nearest-neighbor parameters for G*U mismatches: 5′GU3′/3′UG5′ is destabilizing in the contexts CGUG/GUGC, UGUA/AUGU, and AGUU/UUGA but stabilizing in GGUC/CUGG. *Biochemistry* **30:** 11124–11132.

Herschlag, D. and T.R. Cech. 1990. 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.

Heus, H.A. and A. Pardi. 1991. Structural features that give rise to the unusual stability of

RNA hairpins containing GNRA loops. *Science* **253:** 191–194.
Hickey, D.R. and D.H. Turner. 1985. Effects of terminal mismatches on RNA stability: Thermodynamics of duplex formation for ACCGGGp, ACCGGAp, and ACCGGCp. *Biochemistry* **24:** 3987–3991.
Inners, L.D. and G. Felsenfeld. 1970. Conformation of polyribouridylic acid in solution. *J. Mol. Biol.* **50:** 373–389.
Jencks, W.P. 1987. *Catalysis in chemistry and enzymology*. Dover, New York.
Kierzek, R., L. He, and D.H. Turner. 1992. Association of 2′-5′ oligoribonucleotides. *Nucleic Acids Res.* **20:** 1685–1690.
Kierzek, R., M.H. Caruthers, C.E. Longfellow, D. Swinton, D.H. Turner, and S.M. Freier. 1986. Polymer-supported RNA synthesis and its application to test the nearest-neighbor model for duplex stability. *Biochemistry* **25:** 7840–7846.
Kim, S.-H. and T.R. Cech. 1987. Three-dimensional model of the active site of the self-splicing rRNA precursor of *Tetrahymena*. *Proc. Natl. Acad. Sci.* **84:** 8788-8792.
Kim, S.H., F.L. Suddath, G.J. Quigley, A. McPherson, J.L. Sussman, A.H.J. Wang, N.C. Seeman, and A. Rich. 1974. Three dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* **185:** 435–440.
Li, Y., G. Zon, and W.D. Wilson. 1991. NMR and molecular modeling evidence for a G•A mismatch base pair in a purine-rich DNA duplex. *Proc. Natl. Acad. Sci.* **88:** 26–30.
Mandelkern, M., N. Dattagupta, and D.M. Crothers. 1981. Conversion of B DNA between solution and fiber conformations. *Proc. Natl. Acad. Sci.* **78:** 4294–4298.
Manning, G.S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11:** 179–246.
Michel, F. and E. Westhof. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216:** 585–610.
Nikonowicz, E.P. and A. Pardi. 1992. Three-dimensional heteronuclear NMR studies of RNA. *Nature* **355:** 184–186.
Oas, T.G. and P.S. Kim. 1988. A peptide model of a protein folding intermediate. *Nature* **336:** 42–48.
Olmsted, M.C., C.F. Anderson, and M.T. Record, Jr. 1989. Monte Carlo description of oligoelectrolyte properties of DNA oligomers: Range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits. *Proc. Natl. Acad. Sci.* **86:** 7766–7770.
Olsthoorn, C.S.M., L.J. Bosteläar, J.F.M. De Rooij, J.H. van Boom, and C. Altona. 1981. Circular dichroism study of stacking properties of oligodeoxyadenylates and polydeoxyadenylate. *Eur. J. Biochem.* **115:** 309–321.
Petersheim, M. and D.H. Turner. 1983. Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* **22:** 256–263.
Pörschke, D., O.C. Uhlenbeck, and F.H. Martin. 1973. Thermodynamics and kinetics of the helix-coil transition of oligomers containing GC base pairs. *Biopolymers* **12:** 1313–1335.
Powell, J.T., E.G. Richards, and W.B. Gratzer. 1972. The nature of stacking equilibria in polynucleotides. *Biopolymers* **11:** 235–250.
Pyle, A.M. and T.R. Cech. 1991. Ribozyme recognition of RNA by tertiary interactions with specific ribose 2′-OH groups. *Nature* **350:** 628–631.
Quigley, G.J. and A. Rich. 1976. Structural domains of transfer RNA molecules. *Science* **194:** 796–806.

Record, M.T., Jr., C.F. Anderson, and T.M. Lohman. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **11:** 103–178.
Richards, E.G., C.P. Flessel, and J.R. Fresco. 1963. Polynucleotides. VI. Molecular properties and conformation of polyribouridylic acid. *Biopolymers* **1:** 431–446.
Robertson, D.L. and G.F. Joyce. 1990. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344:** 467–468.
Robertus, J.D., J.E. Ladner, J.T. Finch, D. Rhodes, R.S. Brown, B.F.C. Clark, and A. Klug. 1974. Structure of yeast phenylalanine tRNA at 3Å resolution. *Nature* **250:** 546–551.
SantaLucia, J., Jr. 1991. "The role of hydrogen bonding in the thermodynamics and structure of mismatches in RNA oligonucleotides." Ph.D. thesis. University of Rochester, New York.
SantaLucia, J., Jr., R. Kierzek, and D.H. Turner. 1990. Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry* **29:** 8813–8819.
———. 1991a. Functional group substitutions as probes of hydrogen bonding between GA mismatches in RNA internal loops. *J. Am. Chem. Soc.* **113:** 4313–4322.
———. 1991b. Stabilities of consecutive A•C, C•C, G•G, U•C, and U•U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U•U and C•C$^+$ pairs. *Biochemistry* **30:** 8242–8251.
———. 1992. Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science* **256:** 217–219.
Shoemaker, K.R., P.S. Kim, E.J. York, J.M. Stewart, and R.L. Baldwin. 1987. Tests of the helix dipole model for stabilization of α-helices. *Nature* **326:** 563–567.
Sugimoto, N., R. Kierzek, and D.H. Turner. 1987. Sequence dependence for the energetics of dangling ends and terminal base pairs in ribonucleic acid. *Biochemistry* **26:** 4554–4558.
Sugimoto, N., M. Tomka, R. Kierzek, P.C. Bevilacqua, and D.H. Turner. 1989. Effects of substrate structure on the kinetics of circle opening reactions of the self-splicing intervening sequence from *Tetrahymena thermophila*: Evidence for substrate and Mg^{2+} binding interactions. *Nucleic Acids Res.* **17:** 355–371.
Suurkuusk, J., J. Alvarez, E. Freire, and R. Biltonen. 1977. Calorimetric determination of the heat capacity changes associated with the conformational transitions of polyriboadenylic acid and polyribouridylic acid. *Biopolymers* **16:** 2641–2652.
Tjivikua, T., P. Ballester, and J. Rebek, Jr. 1990. Self-replicating system. *J. Am. Chem. Soc.* **112:** 1249–1250.
Tuerk, C. and L. Gold. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249:** 505–510.
Tuerk, C., P. Gauss, C. Thermes, D.R. Groebe, M. Gaule, N. Guild, G. Stormo, Y. d'Aubenton-Carafa, O.C. Uhlenbeck, I. Tinoco, Jr., E.N. Brody, and L. Gold. 1988. CUUCGG hairpins: Extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proc. Natl. Acad. Sci.* **85:** 1364–1368.
Turner, D.H., N. Sugimoto, and S.M. Freier. 1988. RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.* **17:** 167–192.
Turner, D.H., N. Sugimoto, R. Kierzek, and S.D. Dreiker. 1987. Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.* **109:** 3783–3785.
Usher, D.A. and A.H. McHale. 1976. Hydrolytic stability of helical RNA: A selection advantage for the natural 3′, 5′-bond. *Proc. Natl. Acad. Sci.* **73:** 1149–1153.
Usman, N., K.K. Ogilvie, M.Y. Jiang, and R.L. Cedergren. 1987. Automated chemical

synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramides on a controlled-pore glass support: Synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an *Escherichia coli* formylmethionine tRNA. *J. Am. Chem. Soc.* **109:** 7845–7854.

Williams, A.P., C.E. Longfellow, S.M. Freier, R. Kierzek, and D.H. Turner. 1989. Laser temperature-jump, spectroscopic, and thermodynamic study of salt effects on duplex formation by dGCATGC. *Biochemistry* **28:** 4283–4291.