BIOCHEMISTRY

 Volume 44, Number 40 October 11, 2005

Current Topics

RNA Challenges for Computational Chemists†

Ilyas Yildirim‡ and Douglas H. Turner*,§

Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, and Departments of Chemistry and of Pediatrics, University of Rochester, Rochester, New York 14627

Received June 27, 2005; Revised Manuscript Received August 9, 2005

ABSTRACT: Some experimental results for the thermodynamics of RNA folding cannot be explained by simple pairwise hydrogen-bonding models. Such effects include the stabilities of isoguanosine−isocytidine (iG−iC) base pairs and of various 2 × 2 nucleotide internal loops. Presumably, these results can be explained by base stacking effects, which can be partitioned into Coulombic and overlap effects. We review experimental measurements that provide benchmarks for testing the approximations and theories used for modeling nucleic acids. Quantitative agreement between experiment and theory will indicate understanding of the interactions determining RNA stability and structure.

An understanding of the physical−chemical interactions underlying RNA folding would allow predictions of structure and perhaps function from sequence. The large number of atoms in an RNA molecule necessitates the use of approximate methods, rather than the rigorous equations of quantum mechanics. Many approaches are possible, including coarse-grained potentials (*1*), which use residue-centered force fields; molecular mechanics, which uses atom-centered force fields such as AMBER (*2*), CHARMM (*3*, *4*), and GROMOS (*5*, *6*); approximate quantum mechanics, which considers interactions of electrons and nuclei (*7*, *8*); and QM/MM, which combines quantum mechanics with molecular mechanics (*9*−*16*). Depending on the property to be predicted, the size of the RNA, and the domain of interest, different approaches will provide acceptable approximations.

The secondary structure of RNA can sometimes be deduced by sequence comparison, which relies on the Watson−Crick rules for base pairing and the assumption that secondary structures are more conserved than sequences for function (*17*). Often, however, there are not enough sequences to determine a definitive secondary structure. For these cases, free-energy minimization with a nearest neighbor model is the most popular method for predicting secondary structures (*18*−*38*). In this method, possible secondary structure motifs are assigned free-energy parameters, and these values are added to predict the total free energy of forming a secondary structure. The structure with the lowest free energy is assumed to dominate in solution. Rigorously, however, the concentrations of the various possible structures are predicted to be weighted by a Boltzmann factor, $\exp(-\Delta G^\circ/RT)$, where ΔG° is the free energy change for folding, R is the gas constant, 1.987 cal $K^{-1}mol^{-1}$, and T is the temperature in kelvins. Understanding intermolecular interactions such as hydrogen bonding, stacking, and so forth would allow accurate prediction of ΔG° and therefore secondary structure for an RNA.

Watson−Crick base pairs are the most common and most extensively studied motif in RNA structures. Configurations with the same base compositions but different permutations of base pairs generally have different free energies. For example, the duplexes $(5'CGCG3')_2$ and $(5'GGCC3')_2$ both have four GC base pairs, but have free-energy changes of duplex formation in 1 M NaCl at 37 °C of −3.7 and −5.4 kcal/mol, respectively (*33*). Presently, nearest-neighbor models are the best approximate methods to predict ther-

† This work was supported by NIH Grant GM22939 (D.H.T.).

* To whom correspondence should be addressed. Phone: (585) 275-3207. Fax: (585) 506-0205. E-mail: turner@chem.rochester.edu.

‡ Department of Physics and Astronomy, University of Rochester.

§ Department of Chemistry, University of Rochester.

10.1021/bi051236o CCC: $30.25
Published on Web 09/17/2005

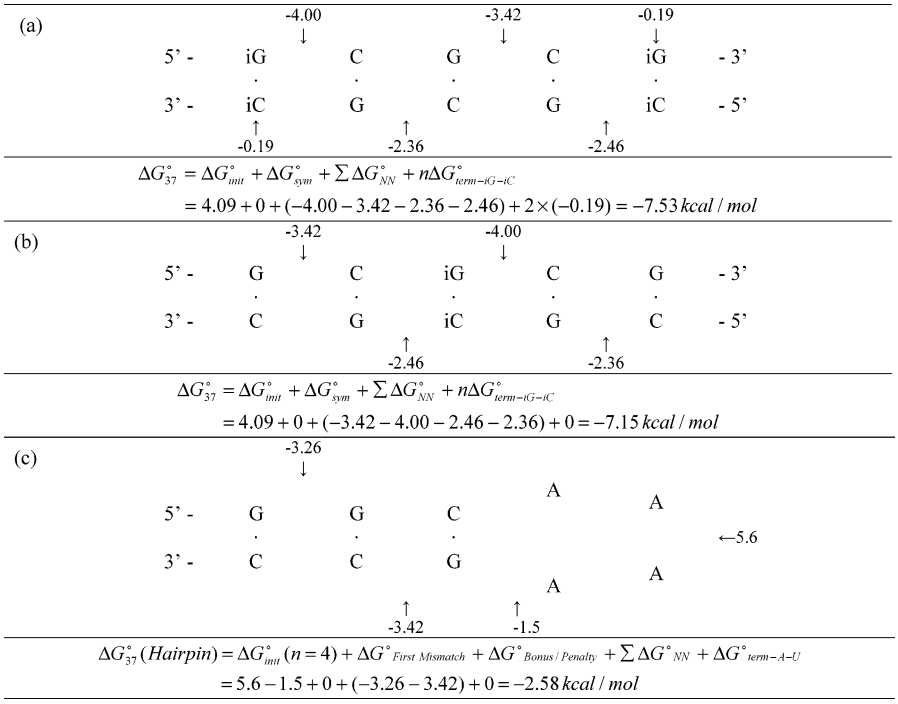

FIGURE 1: Application of INN-HB model (*33*) to duplex formation by 5′iGCGCiG/3′iCGCGiC and 5′GCiGCG/3′CGiCGC (a and b) and to intramolecular folding by 5′GGCAAAAGCC3′ (*26*) (c) at 37 °C. Additional examples are given elsewhere (*31*, *33*, *93*). In this model, ΔG°_{init} is the free-energy change for forming the first base pair and is assumed to depend on the free energy for hydrogen bonding within a GC pair and the free-energy penalty for either bringing two strands together (a and b) or forming a loop (c). ΔG°_{sym} is a symmetry correction that is only applied when two strands of identical sequence form a duplex, $\sum \Delta G^\circ_{NN}$ is the sum of nearest neighbor free-energy parameters (*33*, *76*), each of which contains a contribution from half of the hydrogen bonds in the nearest neighbor base pairs and from the stacking of the nearest neighbor base pairs; $\Delta G^\circ_{term-iG-iC}$ is half the difference in hydrogen bonding free energy between an iGiC and GC pair (*76*). This term accounts for the possibility that two helices may contain the same nearest neighbors but different base pair compositions. For example, the duplexes (a) and (b) have the same nearest neighbors but (a) has two iG-iC pairs while (b) has only one. For natural sequences, $\Delta G^\circ_{term-iG-iC}$ is replaced by $\Delta G^\circ_{term-A-U}$. For the hairpin, $\Delta G^\circ_{First\ Mismatch}$ is the free energy for stacking of the first mismatch on the stem helix and $\Delta G^\circ_{Bonus/Penalty}$ is applied to loop sequences with unusual stability.

modynamic properties of RNA and DNA duplexes containing only Watson−Crick pairs (*18*−*22*, *28*−*31*, *33*). Figure 1 illustrates how the INN-HB[1] (individual nearest neighbor-hydrogen bonding) model (*33*) is used to calculate the free-energy change for folding the duplexes 5′iGCGCiG/3′iCGCGiC and 5′GCiGCG/3′CGiCGC and the hairpin 5′GGCAAAAGCC3′.

Free-energy minimization with a nearest neighbor model is a useful approximation for predicting RNA secondary structure, but on average, predicts only 73% of known base pairs (*25*). One major limitation is that most algorithms do not include pseudoknots. In a pseudoknot, a nucleotide between paired nucleotides i and j forms a base pair with a nucleotide not between i and j. Complete inclusion of pseudoknots is an NP-complete problem, that is, one for which there is no known polynomial time solution (*35*). Algorithms that allow many types of pseudoknots are available, however, and able to run in time that scales as N^5 (*36*, *37*) or N^6 (*38*), where N is the number of nucleotides. The algorithm of Rivas and Eddy (*38*) is able to handle 90% of the pseudoknots found in a database of 486 structures (*39*). Good parameters for predicting the stabilities of pseudoknots are not available, however, due to lack of experimental data on model systems. This lack of experimentally measured parameters is also a problem for other types of loops, especially multibranch loops (*40*, *41*). The parameters used in most algorithms (*25*, *33*) have been measured in 1 M NaCl, which usually mimics well typical intracellular ionic conditions, for example, 0.15 M KCl and several millimolar Mg^{2+} (*42*−*44*). Motifs have been discovered, however, where Mg^{2+} or K^+ increase stability by several kilocalories per mole (*45*−*47*). The limited knowledge of such effects also limits the accuracy of structure prediction.

Secondary structures may also be perturbed by other interactions not included in prediction algorithms. For example, tertiary interactions and/or interactions with proteins may be important. Both types of interactions are expected to be relatively weak compared to base pairing, however. A single GC pair can stabilize a helix by 3 kcal/mol at 37 °C. In contrast, the entire folding free energy for proteins having molecular weights of about 10 000 averages about 9 kcal/mol (*48*).

Of course, free-energy minimization methods assume that the RNA is in equilibrium. This may not be true for all RNAs. The very favorable free energy associated with a canonically paired helix can kinetically trap an RNA secondary structure. Such kinetic trapping is often observed when RNA is renatured (*49*). The reasonable success of free-energy minimization despite the limitations discussed above, however, suggests that kinetics does not usually play a dominant role in determining secondary structure.

[1] Abbreviations: N and M, any nucleotide including A or G; iC, isocytidine; iG, isoguanosine; I, inosine; INN-HB, individual nearest neighbor-hydrogen bonding; R.E.D., RESP ESP charge derive; RESP, restrained electrostatic potenital; eu, cal K^{-1} mol^{-1}; T_M, melting temperature in kelvins; T_m, melting temperature in degrees Celsius; C_T, total RNA strand concentration.

Some of the above limitations can be overcome by combining free-energy minimization with sequence comparison and/or constraints derived from experiments (*23*−*27*, *32*, *34*). These enhancements, however, do not compensate well for pseudoknots. For example, when predictions are constrained by chemical modification data in an algorithm that does not allow pseudoknots, the average accuracy of secondary structure predictions is only 60% for five sequences having 7−10% of base pairs in pseudoknots. In contrast, the average accuracy is 84% for 11 sequences with ≤5% of base pairs in pseudoknots (*25*).

In principle, if the intermolecular interactions in RNA were understood, then it should be possible to predict secondary structure and subsequently local and global three-dimensional structure. This review discusses some considerations for theoretical approaches and then presents some experimental benchmarks for testing these approaches.

THEORY

Theories for important interactions are required to rationalize experimental results and predict the secondary and three-dimensional structures of RNA molecules. Quantum mechanics is the best theory for describing the interactions in atomic systems. Unfortunately, even the simplest RNA systems have hundreds of atoms. As a result, direct application of ab initio quantum mechanics is not yet possible. Improvement in computer power and algorithms may allow the direct use of quantum mechanics on these complicated systems in the future, however.

Quantum mechanics cannot be applied directly to even small RNAs, but it can be applied to individual components such as guanine, adenine, uracil, and cytosine, and those results can be used to build models to predict experimental observables. These models are called force fields. Most of the popular force fields (*2*−*6*) include an atom-centered point-charge model. Therefore, these simulation methods need a set of point charges for all of the atoms in the system. Methods to determine the charges include CHELP (*50*), CHELPG (*51*), RESP (*52*), and the Merz−Kollman scheme (*53*, *54*). The best method for finding charges depends on the system of interest (*55*). Starting with an initial structure and using a force field, properties such as stability and dynamics can be predicted for complicated RNA configurations. The initial structure can be obtained from X-ray crystallography or NMR spectroscopy or can be homology-modeled from X-ray and NMR structures.

Theoretical simulations are done using molecular mechanics (*56*). In these simulations, the initial structure is varied in order to minimize the energy. Molecular mechanics is combined with molecular dynamics and Monte Carlo methods to predict properties. Molecular dynamics uses Newtonian mechanics (*57*−*61*). A molecular simulation is created by changing the positions and velocities of the initial configuration as a function of time. A configuration will be reached, which corresponds to a local or global energy minimum. In the Monte Carlo method (*62*−*64*), the initial configuration is changed randomly, and configurations are retained if they lower energy or pass a Boltzmann-weighted probability test (Metropolis Criterion).

The crucial point in molecular mechanics is to find a force field that provides a good description of the system. The potential functions used in force fields can include electrostatic, exchange repulsion, and dispersion (induced dipole−induced dipole) terms, among others. Adding more terms and better parameters to the potential functions will improve predictions. For example, there is a correlation between stacking free energies, melting temperatures, and polarizabilities of DNA bases (*65*, *66*), which suggests that a polarizable force field might improve predictions of RNA thermodynamics and conformations.

RNA and DNA molecules have polyelectrolyte character because of the phosphate group at each residue. This highly negative charge character of RNA and DNA opposes folding. Positive counterions such as Mg^{2+}, K^+, and Na^+ stabilize nucleic acid structures by reducing the Coulombic interactions between the phosphate groups. The positive ions accumulate around the RNA molecule, as described by polyelectrolyte theory (*67*−*69*). This "counterion condensation" reduces the effective charge of RNA and DNA molecules.

Ion−RNA interactions are not restricted to counterion condensation along the RNA backbone. Some ions, such as Mg^{2+} and K^+, stabilize tertiary structures better than other ions (*45*, *47*, *70*, *71*). Such stabilization is not well-understood. Because the cell environment of living organisms is ionic, ionic effects must be included in models for predicting stability and structure. Approaches to this include nonlinear Poisson−Boltzmann theory (*72*) and Ewald summation of Coulombic interactions (*73*, *74*).

EXPERIMENT

A variety of experimental results are required to test the effectiveness of the various approximations and parameters used to predict the properties of RNAs. Some possible benchmarks are described below.

iG−iC versus G−C. A relatively straightforward challenge for computational methods is the description of isoguanosine (iG) and isocytidine (iC). As shown in Figure 2, they form an unnatural base pair, iG−iC, in which the amino and carbonyl groups of G and C are transposed (*75*). Table 1 lists thermodynamic parameters measured for several RNA duplexes containing G−C or iG−iC base pairs. Table 2 lists nearest neighbor thermodynamic parameters for G−C and iG−iC base pairs, as derived from experimental data such as that listed in Table 1 (*76*). Within experimental error, the stability of only the 5′GC/3′CG nearest neighbor is affected by iG-iC substitution. In particular, the duplex stability of 5′GC/3′CG is increased by 0.6 kcal/mol for each iG−iC substitution.

Because only the carbonyl and amino groups of the G and C are transposed to generate iG and iC, iG−iC and G−C pairs are isosteric. The electron density, however, is different (Figures 2 and 3). Therefore, the iG−iC pair is a good system to test predictions of the effects of different electron distributions.

One likely factor determining the stabilities of iG−iC base pairs is increased strength of the hydrogen bonds relative to G−C pairs. This is expected from the increased and decreased electron densities on hydrogen bonding acceptors and donors, respectively (Figure 2). An experimental estimate of the free-energy contribution from these stronger hydrogen bonds can be deduced from the terminal iG−iC term of −0.19 kcal/mol in Table 2 (*76*). This term effectively ac-

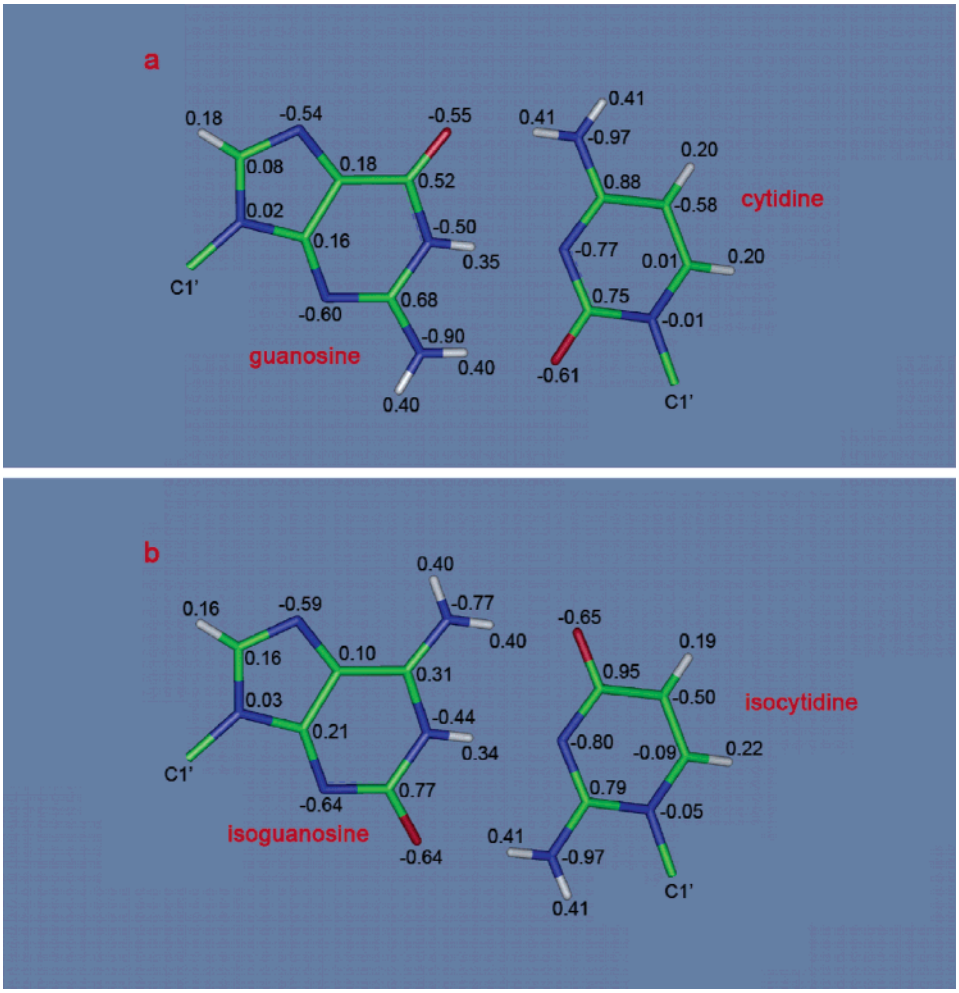

FIGURE 2: GC (a) and iGiC (b) base pairs with their RESP charges. Guanosine and cytidine are natural nucleosides, while isoguanosine and isocytidine are unnatural nucleosides. The GC pair was extracted from the crystal structure of 1QC0 (*94*). Transposing the carbonyl and amino groups of G and C creates iG and iC, respectively. The RESP charges (*52*, *95*) were calculated by following the R.E.D. procedure (*96*). Calculations were done on the nucleosides, but only the C1′ of the ribose is shown.

Table 1: Thermodynamics of Formation of RNA Duplexes in 1 M NaCl

	$1/T_M$ vs $\ln(C_T)$ parameters			
sequence (5′→3′)	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	T_m (°C)
GCGC[a]	−4.6 ± 0.1	−30.5 ± 1.5	−83.4 ± 4.8	26.6
iGiCiGiC[b]	−7.0 ± 0.1	−50.6 ± 5.4	−140.6 ± 17.1	45.2
CCGG[c]	−4.6 ± 0.2	−34.2 ± 0.2	−95.6 ± 0.8	27.2
iCiCiGiG[b]	−4.8 ± 0.1	−39.8 ± 1.1	−112.9 ± 3.6	30.5

[a] Ref *86*. [b] Ref *76*. [c] Ref *87*.

Table 2: Nearest-Neighbor Parameters for RNA Duplexes Containing iG−iC and G−C Pairs in 1 M NaCl (*33*, *76*)[a]

	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)
5′CG3′ 3′GC5′	−2.36 ± 0.09	−10.64 ± 1.65	−26.70 ± 5.0
5′iCG3′ 3′iGC5′	−2.46 ± 0.08	−10.80 ± 1.12	−27.01 ± 3.13
5′iCiG3′ 3′iGiC5′	−2.45 ± 0.17	−12.69 ± 2.23	−33.28 ± 6.29
5′GG3′ 3′CC5′	−3.26 ± 0.07	−13.39 ± 1.24	−32.70 ± 3.8
5′iGG3′ 3′iCC5′	−3.46 ± 0.11	−14.94 ± 1.44	−36.80 ± 4.04
5′GiG3′ 3′CiC5′	−3.07 ± 0.11	−14.67 ± 1.48	−37.34 ± 4.20
5′iGiG3′ 3′iCiC5′	−3.30 ± 0.17	−14.01 ± 2.26	−34.58 ± 6.35
5′GC3′ 3′CG5′	−3.42 ± 0.08	−14.88 ± 1.58	−36.90 ± 4.9
5′iGC3′ 3′iCG5′	−4.00 ± 0.09	−16.90 ± 1.19	−41.32 ± 3.33
5′iGiC3′ 3′iCiG5′	−4.61 ± 0.17	−19.98 ± 2.31	−49.36 ± 6.53
per terminal iG−iC[a]	−0.19 ± 0.07	−1.00 ± 0.98	−2.54 ± 2.77

[a] See caption to Figure 1 for molecular interpretation.

counts for the different base compositions of sequences with the same nearest neighbors, but different terminal base pairs. For example, as illustrated in Figure 1, if two duplexes have the same nearest neighbors, but one has two terminal G−C pairs, while the other has two terminal iG−iC pairs, then the one with the iG−iC terminal pairs will have one more iG−iC pair than the duplex with two terminal G−C pairs. Because the nearest neighbors and therefore presumably stacking interactions in the two duplexes are the same, the difference in hydrogen bonding within an iG−iC and G−C pair is estimated as $2 \times (-0.19) = -0.38$ kcal/mol. Because a single nearest neighbor interaction accounts for half of the hydrogen bonds in the two neighboring base pairs, the difference in hydrogen bonding strength between iG−iC and G−C accounts for about 0.4 kcal/mol of the 1.2 kcal/mol enhanced stability of 5′iGiC/3′iCiG relative to 5′GC/3′CG.

The remaining 0.8 kcal/mol of increased stability for 5′iGiC/3′iCiG compared to 5′GC/3′CG is likely due to the stacking interactions for the iG−iC pairs. The stacking interaction can be partitioned into Coulombic and overlap effects. Coulombic effects are due to interactions between

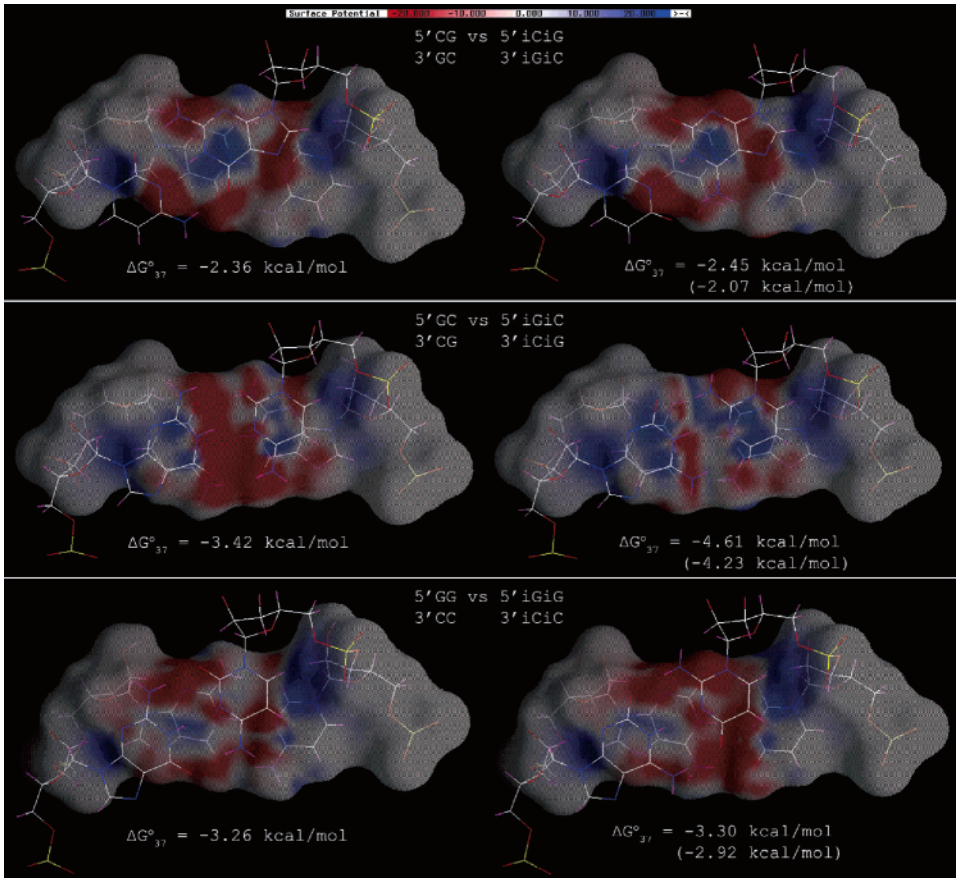

FIGURE 3: The calculated electrostatic potential surfaces of 5′CG/3′GC vs 5′iCiG/3′iGiC (top), 5′GC/3′CG vs 5′iGiC/3′iCiG (middle), and 5′GG/3′CC vs 5′iGiG/3′iCiC (bottom). Structures for 5′CG/3′GC, 5′GC/3′CG, and 5′GG/3′CC were extracted from the crystal structure of 1QC0 (*94*), and the iso structures were created using the program Insight II (Biosym Technologies, San Diego, CA) by transposing the amino and carbonyl groups. The potential surfaces were created with the program GRASP (*77*). To create these surfaces, the RESP charges (*52*, *95*) for G, C, iG, and iC nucleosides were calculated using R.E.D. (*96*), and only the charges on the base rings were used in order to focus on the Coulombic effects. These surfaces were created around the bottom base pairs with a probe radius of 1.4 Å. The red regions represent negative potential, while the blue regions represent positive potential. The red-blue color spectrum runs from −20 to +20 kT/e, respectively. Namely, it shows the potential of a positive unit charge on these surfaces when the charges on the base rings are present. Below each structure, the experimental free energies are presented. To compare the stacking interactions between the G−C and iG−iC pairs, the hydrogen bond strength enhancement of the iG−iC pairs is excluded, and these hydrogen-bond-corrected free energies are shown in parentheses. The top parts of the structures are in the major groove, while the bottom parts are in the minor groove.

effective permanent partial charges on each atom. Overlap effects will include other electrostatic interactions between bases, for example, dispersion interactions, and may also include effects due to burial of bases away from water.

Figure 3 shows electrostatic potential surfaces calculated by the program GRASP (*77*) for the nearest neighbors 5′CG/3′GC versus 5′iCiG/3′iGiC (top), 5′GC/3′CG versus 5′iGiC/3′iCiG (middle), and 5′GG/3′CC versus 5′iGiG/3′iCiC (bottom) in A-form RNA. Only the charges on the base part of the structures were used in the calculations in order to focus on the stacking interactions. Thus, Figure 3 provides a qualitative picture of the Coulombic effect in these nearest neighbor base pairs. For example, there is a big difference between 5′GC/3′CG and 5′iGiC/3′iCiG potential surfaces. In contrast, the differences between potential surfaces are modest for 5′CG/3′GC versus 5′iCiG/3′iGiC and 5′GG/3′CC versus 5′iGiG/3′iCiC. This suggests that the Coulombic interaction is the dominant factor when 5′GC/3′CG is compared to 5′iGiC/3′iCiG.

A qualitative comparison can be made between the pictures in Figure 3 and the experimental results in Table 2. The red-blue color spectrum reflects the value of the electrostatic potential on the surface created around the bottom base pair. The most favorable stacking would place opposite charges on top of each other, producing a neutral (white) potential halfway between the charges. The least favorable stacking would place identical charges on top of each other, producing a large positive (blue) or negative (red) potential halfway between the charges. Qualitatively, the potential surface of 5′iGiC/3′iCiG is smaller in magnitude than 5′GC/3′CG, suggesting more favorable Coulombic interactions. This is consistent with 5′iGiC/3′iCiG being more stable than 5′GC/3′CG even after correction for differences in hydrogen bond strength. Conversely, the electrostatic potential of 5′GG/3′CC is smaller in magnitude than that of 5′iGiG/3′iCiC, suggesting more favorable Coulombic interactions. After correction for differences in hydrogen bonding, 5′GG/3′CC is more stable than 5′iGiG/3′iCiC, although the difference is within experimental error. For 5′iCiG/3′iGiC and 5′CG/3′GC, there is little difference in the electrostatic potential or stability. It is likely that Figure 3 does not capture all the important physics and chemistry. For example, the minor groove is more accessible to water than the major groove, so that the local dielectric constants may not be identical. Consideration of only duplex geometries also neglects any differential stacking in the single strands.

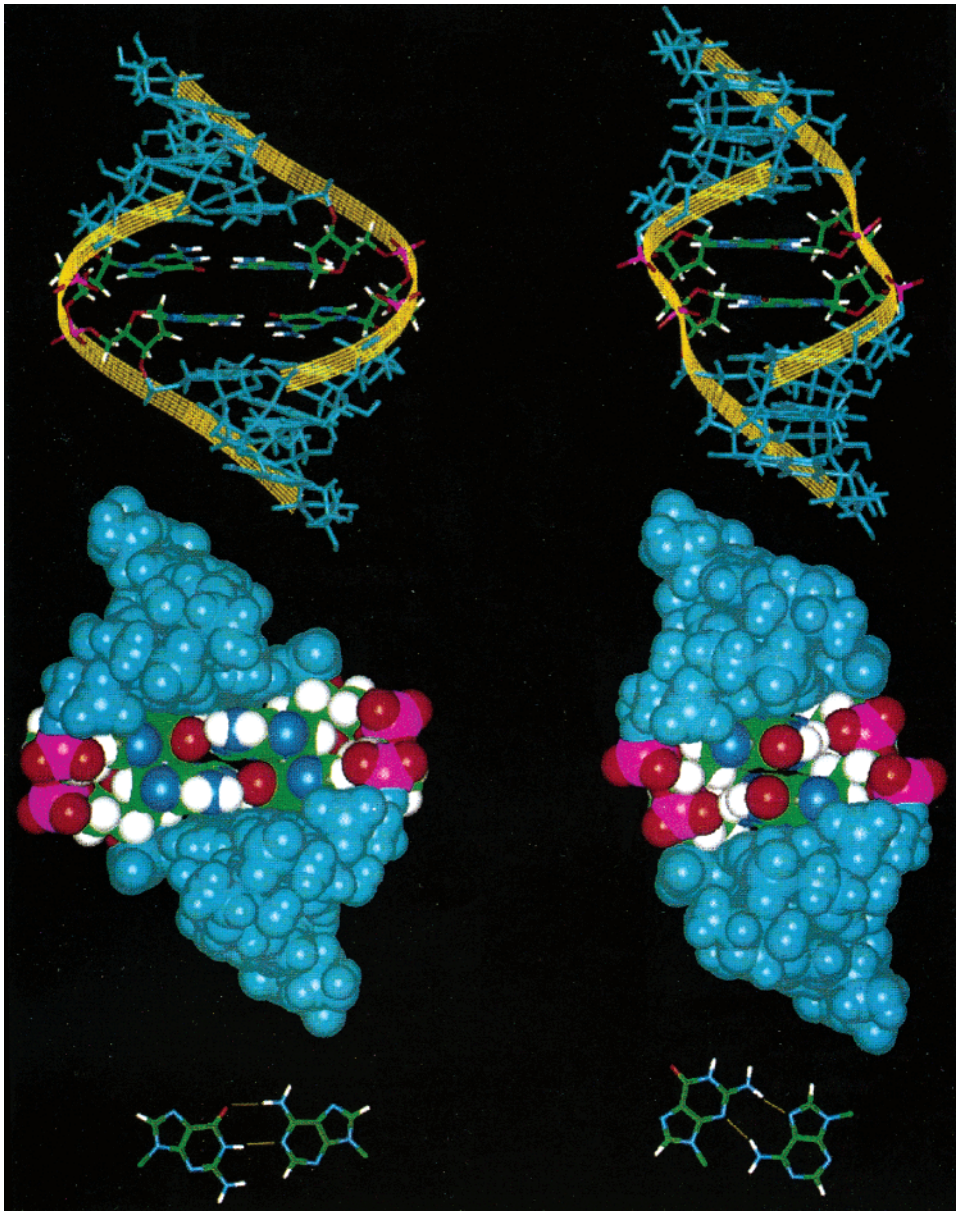

FIGURE 4: Stacking interactions (Coulombic and overlap effects) can change local 3D structures (*81*). The NMR structures of 1MIS, $^{5'GCG\mathbf{GA}CGC3'}_{3'CGC\mathbf{AG}GCG5'}$, (left (*81*)) and 1YFV, $^{5'GGC\mathbf{GA}GCC3'}_{3'CCG\mathbf{AG}CGG5'}$, (right (*97*)) are shown in stick (top two) and space-filling (center two) representations, viewed from major grooves. Imino hydrogen-bonded (bottom left) and sheared (bottom right) GA mismatches are shown in stick representations. Yellow ribbons follow the sugar−phosphate backbone. Because of the electronic structure differences of these two motifs, the local 3D structures are different.

Another determinant of stability that can be seen in Figure 3 is the base overlap. The base rings of 5′GC/3′CG and 5′iGiC/3′iCiG helices have more overlap than the other nearest neighbors. This overlap will enhance dispersion interactions and thus the stability of the structure. Burial of surface area away from water may also be important (*78*). Interestingly, at 37 °C, 5′GC/3′CG is −0.16 kcal/mol more favorable than 5′GG/3′CC, even though the electrostatic potential surface of 5′GG/3′CC appears more favorable than that of 5′GC/3′CG. This suggests that 5′GC/3′CG gains stability because of overlap effects and that the free-energy gain from overlap effects in 5′GC/3′CG compensates for the less favorable Coulombic interactions relative to 5′GG/3′CC. Consistent with this speculation is that 5′iGiG/3′iCiC appears to have less favorable Coulombic interactions than 5′GG/3′CC and is thermodynamically less stable than 5′GC/3′CG. As another example, comparison of 5′CG/3′GC with 5′GC/3′CG reveals that the base overlap in 5′CG/3′GC is less than in 5′GC/3′CG, and the Coulombic interactions in 5′CG/3′GC are similar to those in 5′GC/3′CG. Thus, qualitatively, 5′GC/3′CG is expected to be more stable than 5′CG/3′GC, and this is the case (Table 2).

The qualitative comparisons above suggest that much of the sequence dependence of base pair stability in RNA can be rationalized by current force fields. Quantitative calculations are required, however, to test this hypothesis.

G·U Closure of 2 × 2 Internal Loops: $^{5'G^{NM}C3'}_{3'C_{MN}G5'}$ *versus* $^{5'G^{NM}U3'}_{3'U_{MN}G5'}$. Internal loops with G·U closing pairs are often involved in tertiary and quaternary interactions that are important for structural and functional roles (*79*, *80*). As a result, a good understanding of the thermodynamics of internal loops will help predict both secondary and tertiary structures of RNA. Table 3 lists experimental free-energy

Table 3: Free-Energy Increments (kcal/mol at 37 °C) for Symmetric 2 × 2 Purine Loops (*88*)[a]

closing base pairs	loops GA/AG	AG/GA	AA/AA
5′G/3′C	−2.6	−1.3	1.5
5′C/3′G	−0.7	−0.7	1.3
5′A/3′U	0.3	1.7	2.8
5′U/3′A	0.7	0.9	2.8
5′G/3′U	1.8	2.6	4.1
5′U/3′G	0.1	3.4	2.1

[a] All motifs are symmetric. Thus, a closing base pair of 5′G/3′C and a loop of GA/AG represent the motif 5′G**GA**C3′/3′C**AG**G5′. Values are from refs *26* and *88*, and based on the experimental data from refs *88*−*92*.

parameters for symmetric 2 × 2 purine internal loops with different closing pairs.

Internal loops closed by GC rather than AU pairs are more stable by 1.3 kcal/mol on average, and this difference has been attributed to the different number of hydrogen bonds in GC and AU pairs (*44*). AU and GU pairs both have two hydrogen bonds, but the predicted increment of 1.3 kcal/mol relative to loop closure by GC pairs cannot rationalize some of the results listed in Table 3, especially for closure by GU pairs. In particular, when the internal loops 5′GA/3′AG, 5′AG/3′GA, and 5′AA/3′AA are closed by GC rather than GU pairs, the stabilities are enhanced by 4.4, 3.9, and 2.6 kcal/mol, respectively. The same trend is seen when 5′GGAC/3′CAGG and 5′GAGC/3′CGAG are compared with 5′AGAU/3′UAGA and 5′AAGU/3′UGAA, respectively. When 5′GA/3′AG and 5′AG/3′GA loops are closed with GC base pairs, the stabilities are enhanced by 2.9 and 3.0 kcal/mol, respectively, compared to the same loops closed with AU base pairs. Because the shapes and electronic structures of GU, GC, and AU pairs differ, the Coulombic and overlap interactions between closing base pairs and loop must be different. Thus, calculations of these effects should provide another test of the ability of theory to rationalize experimental results.

GC versus CG Closure of Tandem G·A Pairs: $\begin{smallmatrix}5'G^{GA}C3'\\3'C_{AG}G5'\end{smallmatrix}$ *versus* $\begin{smallmatrix}5'C^{GA}G3'\\3'G_{AG}C5'\end{smallmatrix}$. According to Table 3, the free energy of the GC closure of the 5′GA/3′AG internal loop is enhanced by 1.9 kcal/mol compared to CG closure of the 5′GA/3′AG internal loop. Moreover, NMR structures of these two motifs are different, as shown in Figure 4 (*81*). This, again, suggests that Coulombic and overlap interactions between bases are important. Moreover, nonplanar guanine amino groups might be important, too (*82*). Thus, this system presents a challenge to theoretical methods to predict both the stabilities and structures as a function of sequence in these motifs.

Tandem G·U Pairs. Table 4 lists the free-energy increments for symmetric tandem G·U, I·U, and A−U motifs closed by GC and CG base pairs. One interesting result is that the free energy of 5′CAUG/3′GUAC is 4.2 kcal/mol more favorable than that of 5′CGUG/3′GUGC. The only difference between these two structures is that 5′CGUG/3′GUGC has two adjacent G·U base pairs, while 5′CAUG/3′GUAC has two adjacent A−U base pairs. It is likely that this difference is also due to differences in Coulombic and overlap interactions because the simple hydrogen-bond model

Table 4: Comparison of Free-Energy Increments (kcal/mol at 37 °C) for Symmetric Tandem G·U, I·U, and A−U Motifs

adjacent base pairs	UG/GU	UI/IU	UA/AU	(UG−UI)	(UG−UA)
5′G/3′C	−4.9[a]	-	−5.8[a]	-	0.9
5′C/3′G	−4.2[a]	+0.5[b]	−5.5[a]	−4.7	1.3
adjacent base pairs	**GU/UG**	**IU/UI**	**AU/UA**	**(GU−IU)**	**(GU−AU)**
5′G/3′C	−4.1[a]	-	−5.8[a]	-	1.7
5′C/3′G	−1.1[a]	+2.1[b]	−5.3[a]	−3.2	4.2

[a] Ref *83*. [b] Ref *85*.

cannot rationalize this difference. Interestingly, NMR and chemical substitution effects suggest that not all adjacent G·U pairs form two hydrogen bonds (*83*), and molecular dynamics calculations reproduce this interpretation (*84*).

Replacement of guanine by inosine produces other interesting results. The only structural difference between guanine and inosine is replacement by a hydrogen atom (−H) of an amino group (−NH_2) that is not involved in base−base hydrogen bonding. The free energies of 5′CUGG/3′GGUC and 5′CGUG/3′GUGC are 4.7 and 3.2 kcal/mol more favorable than 5′CUIG/3′GIUC and 5′CIUG/3′GUIC, respectively (*85*). Presumably, this is due to the electronic structure differences between guanine and inosine. A quantitative analysis of these two structures would provide another test of our understanding of Coulombic and overlap effects in RNA molecules.

CONCLUSION

A sufficient understanding of intermolecular interactions in RNA would permit prediction of RNA structure. In this review, some experimental results for RNA thermodynamics and structures are presented, which can serve as benchmarks for theoretical calculations. Generally, these structures have non-Watson−Crick base pairs, such as G·U, I·U, G·A, A·A, and iG−iC. Qualitative comparison between G−C and iG−iC base pairs suggests that stability differences may be rationalized with current force fields by decomposing the effects into Coulombic and overlap effects. Quantitative calculations on a variety of systems are required to test this and other approaches.

ACKNOWLEDGMENT

We thank Profs. Harry Stern, Kara Bren, David H. Mathews, and Mr. Gang Chen for helpful discussions.

REFERENCES

1. Malhotra, A., and Harvey, S. C. (1994) A quantitative model of the *Escherichia coli* 16-S RNA in the 30-S ribosomal-subunit, *J. Mol. Biol. 240*, 308−340.
2. Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Caldwell, J. W., Ross, W. S., and Kollman, P. A. (2004) AMBER 8, University of California, San Francisco, San Francisco, CA.
3. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM−a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187−217.

4. MacKerell, A. D., Jr., Brooks, B., Brooks, C. L., III, Nilsson, L., Roux, B., Won, Y., and Karplus, M. (1998) CHARMM: the energy function and its parametrization with an overview of the program, in *The Encyclopedia of Computational Chemistry* (Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, I., H. F., and Schreiner, P. R., Eds.) pp 271−277, John Wiley & Sons, Chichester, U.K.
5. Berendsen, H. J. C., and van Gunsteren, W. F. (1987) *GROMOS Reference Manual*, University of Groningen, Groningen, The Netherlands.
6. van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P., Mark, A. E., Scott, W. R. P., and Tironi, I. G. (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Verlag der Fachvereine, Zürich, Switzerland.
7. Hobza, P., and Sponer, J. (2002) Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations, *J. Am. Chem. Soc. 124*, 11802−11808.
8. Kratochvil, M., Sponer, J., and Hobza, P. (2000) Global minimum of the adenine···thymine base pair corresponds neither to Watson−Crick nor to Hoogsteen structures. Molecular dynamic/quenching/AMBER and ab initio beyond Hartree−Fock studies, *J. Am. Chem. Soc. 122*, 3495−3499.
9. Freindorf, M., and Gao, J. L. (1996) Optimization of the Lennard-Jones parameters for a combined ab initio quantum mechanical and molecular mechanical potential using the 3-21G basis set, *J. Comput. Chem. 17*, 386−395.
10. Eichinger, M., Tavan, P., Hutter, J., and Parrinello, M. (1999) A hybrid method for solutes in complex solvents: density functional theory combined with empirical force fields, *J. Chem. Phys. 110*, 10452−10467.
11. Gao, J. L., and Xia, X. F. (1992) A priori evaluation of aqueous polarization effects through Monte-Carlo QMMM simulations, *Science 258*, 631−635.
12. Murphy, R. B., Philipp, D. M., and Friesner, R. A. (2000) A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments, *J. Comput. Chem. 21*, 1442−1457.
13. Stanton, R. V., Hartsough, D. S., and Merz, K. M. (1993) Calculation of solvation free-energies using a density functional/molecular dynamics coupled potential, *J. Phys. Chem. 97*, 11868−11870.
14. Tunon, I., Martins-Costa, M. T. C., Millot, C., Ruiz-Lopez, M. F., and Rivail, J. L. (1996) A coupled density functional-molecular mechanics Monte Carlo simulation method: the water molecule in liquid water, *J. Comput. Chem. 17*, 19−29.
15. Warshel, A., and Levitt, M. (1976) Theoretical studies of enzymic reactions−dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme, *J. Mol. Biol. 103*, 227−249.
16. Bash, P. A., Field, M. J., and Karplus, M. (1987) Free-energy perturbation method for chemical-reactions in the condensed phase: a dynamical-approach based on a combined quantum and molecular mechanics potential, *J. Am. Chem. Soc. 109*, 8092−8094.
17. Pace, N. R., Thomas, B. C., and Woese, C. R. (1999) Probing RNA structure, function and history by comparative analysis, in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 113−141, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Borer, P. N., Dengler, B., Tinoco, I., and Uhlenbeck, O. C. (1974) Stability of ribonucleic-acid double-stranded helices, *J. Mol. Biol. 86*, 843−853.
19. Goldstein, R. F., and Benight, A. S. (1992) How many numbers are required to specify sequence-dependent properties of polynucleotides, *Biopolymers 32*, 1679−1693.
20. Gralla, J., and Crothers, D. M. (1973) Free-energy of imperfect nucleic-acid helices. 2. Small hairpin loops, *J. Mol. Biol. 73*, 497−511.
21. Gray, D. M. (1997) Derivation of nearest-neighbor properties from data on nucleic acid oligomers. 1. Simple sets of independent sequences and the influence of absent nearest neighbors, *Biopolymers 42*, 783−793.
22. Gray, D. M. (1997) Derivation of nearest-neighbor properties from data on nucleic acid oligomers. 2. Thermodynamic parameters of DNA−RNA hybrids and DNA duplexes, *Biopolymers 42*, 795−810.
23. Knight, R., Birmingham, A., and Yarus, M. (2004) BayesFold: rational 20 folds that combine thermodynamic, covariation, and chemical data for aligned RNA sequences, *RNA 10*, 1323−1336.
24. Luck, R., Graf, S., and Steger, G. (1999) ConStruct: a tool for thermodynamic controlled prediction of conserved secondary structure, *Nucleic Acids Res. 27*, 4208−4217.
25. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.
26. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
27. Mathews, D. H., and Turner, D. H. (2002) Dynalign: an algorithm for finding the secondary structure common to two RNA sequences, *J. Mol. Biol. 317*, 191−203.
28. SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc. Natl. Acad. Sci. U.S.A. 95*, 1460−1465.
29. SantaLucia, J., and Hicks, D. (2004) The thermodynamics of DNA structural motifs, *Annu. Rev. Biophys. Biomol. Struct. 33*, 415−440.
30. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973) Improved estimation of secondary structure in ribonucleic-acids, *Nat. New Biol. 246*, 40−41.
31. Turner, D. H. (2000) Conformational changes, in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 259−334, University Science Books, Sausalito, CA.
32. Wuchty, S., Fontana, W., Hofacker, I. L., and Schuster, P. (1999) Complete suboptimal folding of RNA and the stability of secondary structures, *Biopolymers 49*, 145−165.
33. Xia, T. B., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X. Q., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson−Crick base pairs, *Biochemistry 37*, 14719−14735.
34. Zuker, M., and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res. 9*, 133−148.
35. Lyngso, R. B., and Pedersen, C. N. S. (2000) RNA pseudoknot prediction in energy-based models, *J. Comput. Biol. 7*, 409−427.
36. Dirks, R. M., and Pierce, N. A. (2003) A partition function algorithm for nucleic acid secondary structure including pseudoknots, *J. Comput. Chem. 24*, 1664−1677.
37. Dirks, R. M., and Pierce, N. A. (2004) An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots, *J. Comput. Chem. 25*, 1295−1304.
38. Rivas, E., and Eddy, S. R. (1999) A dynamic programming algorithm for RNA structure prediction including pseudoknots, *J. Mol. Biol. 285*, 2053−2068.
39. Condon, A., Davy, B., Rastegari, B., Zhao, S., and Tarrant, F. (2004) Classifying RNA pseudoknotted structures, *Theor. Comput. Sci. 320*, 35−50.
40. Diamond, J. M., Turner, D. H., and Mathews, D. H. (2001) Thermodynamics of three-way multibranch loops in RNA, *Biochemistry 40*, 6971−6981.
41. Mathews, D. H., and Turner, D. H. (2002) Experimentally derived nearest-neighbor parameters for the stability of RNA three- and four-way multibranch loops, *Biochemistry 41*, 869−880.
42. McDowell, J. A., He, L. Y., Chen, X. Y., and Turner, D. H. (1997) Investigation of the structural basis for thermodynamic stabilities of tandem GU wobble pairs: NMR structures of $(rGGAGUUCC)_2$ and $(rGGAUGUCC)_2$, *Biochemistry 36*, 8030−8038.
43. Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., and Turner, D. H. (1989) Laser temperature-jump, spectroscopic, and thermodynamic study of salt effects on duplex formation by dGCATGC, *Biochemistry 28*, 4283−4291.
44. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA, *Biochemistry 39*, 9257−9274.
45. Conn, G. L., Gittis, A. G., Lattman, E. E., Misra, V. K., and Draper, D. E. (2002) A compact RNA tertiary structure contains a buried backbone$-K^+$ complex, *J. Mol. Biol. 318*, 963−973.
46. Draper, D. E., Grilley, D., and Soto, A. M. (2005) Ions and RNA folding, *Annu. Rev. Biophys. Biomol. Struct. 34*, 221−243.
47. Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures, *RNA 8*, 307−323.

48. Makhatadze, G. I., and Privalov, P. L. (1995) Energetics of protein structure, in *Advances in Protein Chemistry*, Vol. 47, pp 307−425, Academic Press, Inc., San Diego, CA.
49. Uhlenbeck, O. C. (1995) Keeping RNA happy, *RNA 1*, 4−6.
50. Chirlian, L. E., and Francl, M. M. (1987) Atomic charges derived from electrostatic potentials−a detailed study, *J. Comput. Chem. 8*, 894−905.
51. Breneman, C. M., and Wiberg, K. B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials−the need for high sampling density in formamide conformational-analysis, *J. Comput. Chem. 11*, 361−373.
52. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges−the RESP model, *J. Phys. Chem. 97*, 10269−10280.
53. Besler, B. H., Merz, K. M., and Kollman, P. A. (1990) Atomic charges derived from semiempirical methods, *J. Comput. Chem. 11*, 431−439.
54. Singh, U. C., and Kollman, P. A. (1984) An approach to computing electrostatic charges for molecules, *J. Comput. Chem. 5*, 129−145.
55. Sigfridsson, E., and Ryde, U. (1998) Comparison of methods for deriving atomic charges from the electrostatic potential and moments, *J. Comput. Chem. 19*, 377−395.
56. Kollman, P. A. (2000) Theoretical methods, in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 223−258, University Science Books, Sausalito, CA.
57. Allen, M. P., and Tildesley, D. J. (1987) *Computer Simulation of Liquids*, Oxford University Press, New York.
58. Burkert, U., and Allinger, N. L. (1982) *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington, DC.
59. Flores, T. P., and Moss, D. S. (1990) Simulating the dynamics of macromolecules, in *Molecular Dynamics: Applications in Molecular Biology* (Goodfellow, J. M., Ed.) pp 1−26, CRC Press, Inc., Boca Raton, FL.
60. Frenkel, D., and Smit, B. (2002) *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed., Academic Press, San Diego, CA.
61. Schlick, T. (2002) *Molecular Modeling and Simulation: An Interdisciplinary Guide*, Springer-Verlag, New York.
62. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953) Equation of state calculations by fast computing machines, *J. Chem. Phys. 21*, 1087−1092.
63. Metropolis, N., and Ulam, S. (1949) The Monte Carlo method, *J. Am. Stat. Assoc. 44*, 335−341.
64. Niederreiter, H., Hellekalek, P., Larcher, G., and Zinterhof, P., Eds. (1998) *Monte Carlo and Quasi-Monte Carlo Methods 1996*, Proceedings of a conference at the University of Salzburg, Austria, July 9−12, 1996 (lecture notes in statistics 127), Springer-Verlag, New York.
65. Guckian, K. M., Schweitzer, B. A., Ren, R. X. F., Sheils, C. J., Tahmassebi, D. C., and Kool, E. T. (2000) Factors contributing to aromatic stacking in water: evaluation in the context of DNA, *J. Am. Chem. Soc. 122*, 2213−2222.
66. Rosemeyer, H., and Seela, F. (2002) Modified purine nucleosides as dangling ends of DNA duplexes: the effect of the nucleobase polarizability on stacking interactions, *J. Chem. Soc., Perkin Trans. 2*, 746−750.
67. Manning, G. S. (1969) Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties, *J. Chem. Phys. 51*, 924−933.
68. Manning, G. S. (1978) Molecular theory of polyelectrolyte solutions with applications to electrostatic properties of polynucleotides, *Q. Rev. Biophys. 11*, 179−246.
69. Record, M. T., Anderson, C. F., and Lohman, T. M. (1978) Thermodynamic analysis of ion effects on binding and conformational equilibria of proteins and nucleic-acids−roles of ion association or release, screening, and ion effects on water activity, *Q. Rev. Biophys. 11*, 103−178.
70. Auffinger, P., Bielecki, L., and Westhof, E. (2004) Symmetric K^+ and Mg^{2+} ion-binding sites in the 5 S rRNA loop E inferred from molecular dynamics simulations, *J. Mol. Biol. 335*, 555−571.
71. Draper, D. E. (2004) A guide to ions and RNA structure, *RNA 10*, 335−343.
72. Sharp, K. A., and Honig, B. (1990) Electrostatic interactions in macromolecules−theory and applications, *Annu. Rev. Biophys. Biophys. Chem. 19*, 301−332.
73. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald−an N·Log(N) method for Ewald sums in large systems, *J. Chem. Phys. 98*, 10089−10092.
74. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth Particle Mesh Ewald method, *J. Chem. Phys. 103*, 8577−8593.
75. Switzer, C. Y., Moroney, S. E., and Benner, S. A. (1993) Enzymatic recognition of the base-pair between isocytidine and isoguanosine, *Biochemistry 32*, 10489−10496.
76. Chen, X. Y., Kierzek, R., and Turner, D. H. (2001) Stability and structure of RNA duplexes containing isoguanosine and isocytidine, *J. Am. Chem. Soc. 123*, 1267−1274.
77. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association−insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.
78. Holbrook, J. A., Capp, M. W., Saecker, R. M., and Record, M. T. (1999) Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices, *Biochemistry 38*, 8409−8422.
79. Gautheret, D., Konings, D., and Gutell, R. R. (1995) GU base-pairing motifs in ribosomal-RNA, *RNA 1*, 807−814.
80. van Knippenberg, P. H., Formenoy, L. J., and Heus, H. A. (1990) Is there a special function for UG base-pairs in ribosomal-RNA, *Biochim. Biophys. Acta 1050*, 14−17.
81. Wu, M., and Turner, D. H. (1996) Solution structure of $(rGCG-GACGC)_2$ by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 35*, 9677−9689.
82. Sponer, J., Mokdad, A., Sponer, J. E., Spackova, N., Leszczynski, J., and Leontis, N. B. (2003) Unique tertiary and neighbor interactions determine conservation patterns of Cis Watson−Crick A/G base-pairs, *J. Mol. Biol. 330*, 967−978.
83. Chen, X. Y., McDowell, J. A., Kierzek, R., Krugh, T. R., and Turner, D. H. (2000) Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for GU pairs: One hydrogen bond for each GU pair in $(rGGCGUGCC)_2$ and two for each GU pair in $(rGAGUGCUC)_2$, *Biochemistry 39*, 8970−8982.
84. Pan, Y. P., Priyakumar, U. D., and MacKerell, A. D. (2005) Conformational determinants of tandem GU mismatches in RNA: insights from molecular dynamics simulations and quantum mechanical calculations, *Biochemistry 44*, 1433−1443.
85. Serra, M. J., Smolter, P. E., and Westhof, E. (2004) Pronounced instability of tandem IU base pairs in RNA, *Nucleic Acids Res. 32*, 1824−1828.
86. Freier, S. M., Sinclair, A., Neilson, T., and Turner, D. H. (1985) Improved free-energies for GC base-pairs, *J. Mol. Biol. 185*, 645−647.
87. Petersheim, M., and Turner, D. H. (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.
88. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA, *Biochemistry 40*, 11509−11517.
89. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: symmetrical loops are more stable than asymmetric loops, *Biochemistry 30*, 6428−6436.
90. SantaLucia, J., Kierzek, R., and Turner, D. H. (1990) Effects of GA mismatches on the structure and thermodynamics of RNA internal loops, *Biochemistry 29*, 8813−8819.
91. Walter, A. E., Wu, M., and Turner, D. H. (1994) The stability and structure of tandem GA mismatches in RNA depend on closing base-pairs, *Biochemistry 33*, 11349−11354.
92. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic-table of symmetrical tandem mismatches in RNA, *Biochemistry 34*, 3204−3211.
93. Xia, T., Mathews, D. H., and Turner, D. H. (2001) Thermodynamics of RNA secondary structure formation, in *RNA* (Söll, D. G., Nishimura, S., and Moore, P. B., Eds.) pp 21−48, Elsevier Science Ltd., Kidlington, Oxford, U.K.
94. Klosterman, P. S., Shah, S. A., and Steitz, T. A. (1999) Crystal structures of two plasmid copy control related RNA duplexes: an 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution, *Biochemistry 38*, 14784−14792.
95. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollman, P. A. (1993) Application of RESP charges to calculate conformational energies,

hydrogen-bond energies, and free-energies of solvation, *J. Am. Chem. Soc. 115*, 9620−9631.
96. Pigache, A., Cieplak, P., and Dupradeau, F. Y. (2004) Automatic and highly reproducible RESP and ESP charge derivation: application to the development of programs RED and X RED, *Abstr. Pap. Am. Chem. Soc. 227*, U1011.
97. SantaLucia, J., and Turner, D. H. (1993) Structure of (rGGC-GAGCC)$_2$ in solution from NMR and restrained molecular-dynamics, *Biochemistry 32*, 12612−12623.

BI051236O