Thermodynamic Analysis of an RNA Combinatorial Library Contained in a Short Hairpin

Joanne M. Bevilacqua and Philip C. Bevilacqua*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

Received July 17, 1998; Revised Manuscript Received September 14, 1998

ABSTRACT: Prediction of nucleic acid structure from sequence requires thermodynamic parameters for a variety of motifs, many of which are complex and consist of a large number of possible sequence combinations. Here we report an experimental approach for identifying the stable and unstable members of an RNA combinatorial library. Short model RNA hairpins consisting of 13 base pairs (bp) flanked by primer binding sites are constructed and separated according to their relative thermodynamic stabilities using temperature gradient gel electrophoresis (TGGE). Partially denaturing TGGE is carried out with potassium chloride, sodium chloride, or magnesium chloride salts in the gel. The T_Ms of model hairpins can be tuned by adjusting the concentration of urea in the gel while maintaining the correct order of stabilities for the hairpins. Mixtures of RNAs differing by a single Watson−Crick base pair are resolved according to their relative thermodynamic stabilities, as are mixtures of GC or AU base pair transversions differing in $\Delta G°_{37}$ by only 0.3−0.5 kcal/mol. In addition, a simple combinatorial library with one position of randomization opposite a guanosine is prepared and separated into its four members by parallel and perpendicular TGGE. The order of thermodynamic stabilities for the library determined by TGGE is shown to be the same when assayed by UV-melting experiments. Analysis of the thermodynamics of folding of combinatorial libraries is general and may be applied to a wide variety of complex nucleic acid secondary and tertiary motifs in order to identify the stable and unstable members.

Ribonucleic acids fold into a wide variety of very stable structures. In some cases, the global minimum free energy folding is achieved, while in other cases local minimum free energy foldings are stably occupied. In either case, the three-dimensional structure of RNA is determined by the thermodynamic stability of a particular fold. There are several methods for predicting RNA secondary structure from sequence. One approach involves determining the minimum free energy for folding (*1*−*3*), and another approach simulates the kinetic pathway of RNA folding and metastable structures that can be encountered during folding (*4, 5*). Both approaches require knowledge of the thermodynamic parameters for folding of various RNA motifs.

Thermodynamic parameters for a variety of RNA structural features have been determined. These include Watson−Crick and non-Watson−Crick base pairs, 3′- and 5′-terminal nucleotides, terminal mismatches, internal loops of 2 by 1 and 2 by 2 nucleotides, and a partial set of rules for coaxial stacking interactions and hairpin loops (*3, 6*). These parameters have largely been determined by synthesizing individual RNA oligomers of interest and obtaining their thermodynamic parameters by UV-melting experiments. Free energy minimization of a set of 12 small subunit rRNAs, group I introns, and group II introns predicted 82.5% of phylogenetic base pairs correctly in the optimal secondary structure (*3*). Comparison of a wide range of computer-predicted and phylogenetically-determined base pairs from 16S and 16S-like rRNA secondary structures performed with a limited set of the above parameters indicated that Archaeal and Eubacterial 16S rRNAs are predicted as high as 80%, but that many Eucarya and mitochondrial 16S rRNAs are predicted with an accuracy of only about 30% (*7*). Overall, the mean scores for prediction of 56 16S rRNA and 72 23S rRNA secondary structures taken from a mixture of Archaea, Eubacteria, and Eucarya were 46 and 44%, respectively (*7, 8*). These comparisons indicate room for substantial improvement.

Current data suggest that thermodynamic parameters for other RNA motifs will improve structure prediction. For example, incremental improvements in RNA secondary structure prediction have resulted from the inclusion of thermodynamic parameters for tetraloops (*9*); coaxial stacking (*10*); and a combination of internal loops, hairpin loops, and altered coaxial stacking rules (*3*). Nevertheless, a large number of RNA secondary and tertiary structures have not had their thermodynamic parameters analyzed. The primary limitation to such investigations has been that many RNA secondary and tertiary structural motifs are of sufficient size to rule out direct synthesis and UV melting of each possible sequence. Within a given motif, however, it appears likely that a small population of stable and unstable sequences exists. For example, RNA is known to fold into several families of stable tetraloops, including GNRA sequences, and ultrastable tetraloops, including UNCG sequences, where N is any of the four bases and R is a purine base (*11*−*13*). Thus, a reasonable hypothesis is that for a given motif most of the sequences will fold with similar thermodynamics, but

* To whom correspondence should be addressed. Phone: (814) 863-3812. Fax: (814) 863-8403. E-mail: pcb@chem.psu.edu.

10.1021/bi981732v CCC: $15.00
Published on Web 10/22/1998

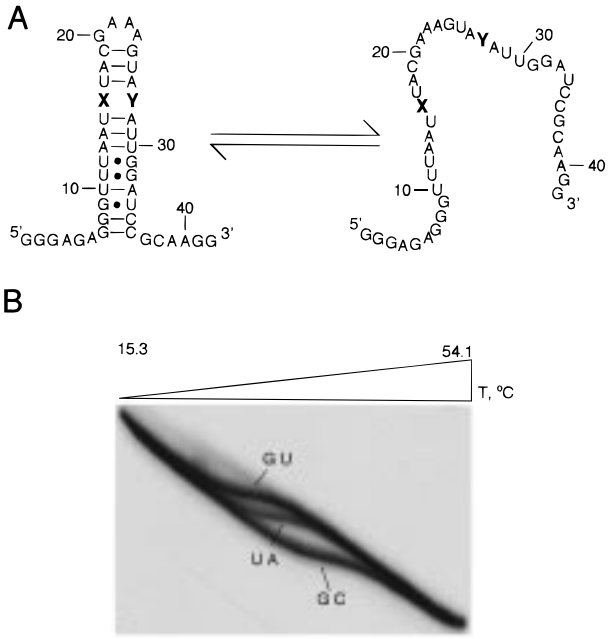

FIGURE 1: Separation of model RNA hairpins. (A) Equilibrium between folded (left) and unfolded (right) forms of model hairpins. Watson−Crick base pairs are denoted with a dash, and wobble base pairs are denoted with a dot. The letters 'X' and 'Y' denote positions of substitution with various base pairs and mismatches. (B) Perpendicular TGGE of 5′-^{32}P-labeled RNA oligomers with XY substituted to GU, UA, and GC pairs as indicated. Equal amounts of the three oligomers were loaded on the gel. The temperature extremes measured in the gel were 15.3−54.1 °C. The gel was run under standard TGGE conditions.

a small number will fold with ultrastable or unstable free energies. Identification and thermodynamic analysis of rare and unusually stable sequences may aid in prediction of RNA structure, and may lead to new principles for RNA folding.

To investigate complex libraries, we have developed a method for analyzing the thermodynamics of combinatorial libraries. The melting behavior of short RNA hairpins containing a combinatorial library is analyzed by temperature gradient gel electrophoresis (TGGE).[1] This approach is shown to be sensitive to small differences in free energy and capable of resolving a simple RNA library into stable and unstable members.

MATERIALS AND METHODS

Preparation of RNA. The DNA templates for synthesis of the RNA hairpins were prepared by solid-phase synthesis, deblocked, and desalted (Integrated DNA Technologies). All DNA templates were purified by denaturing polyacrylamide gel electrophoresis (PAGE). As an example, the DNA template for the RNA sequence XY = UA (Figure 1A) was as follows: 5′-CCTTGCGGATCCAATTTACTTTCGTAAA-TTAAACCCTCTCCCTATAGTGAGTCGTATTAATTTC-3′. Positions of randomization were prepared at the DNA level by mixing the four phosphoramidites for a single coupling. The RNA hairpins were synthesized by T7 transcription reactions (1 mL) using phage T7 RNA polymerase (*14*), purified by denaturing PAGE as previously described (*15*), and stored in TE (10 mM Tris, 1 mM Na_2-EDTA, pH 7.5). The RNAs used in the TGGE studies were dephosphorylated and 5′-^{32}P-end-labeled with polynucleotide kinase and [γ-^{32}P]ATP. Following the labeling reactions, RNAs were purified by denaturing PAGE. Labeled RNA concentrations were determined by scintillation counting. The sequences of the RNA molecules where XY = GA and GG (Figure 1A) were confirmed by limited digestions with G-specific RNase T1 and A-specific RNase U2.

TGGE. Experiments were carried out on an apparatus similar to one previously described (*16*). A vertical apparatus was used, and the glass plates containing the acrylamide gel were placed between two aluminum blocks. Channels in the blocks allowed externally circulating water to establish a temperature gradient from left to right (perpendicular to the electric field), or, using a different set of aluminum blocks, from top to bottom (parallel to the electric field). Water baths were used to control the two temperatures. The system was insulated from the environment by placing a sheet of closed cell foam outside of each aluminum plate, and securing with c-clamps. The perpendicular temperature gradient was measured by an Omega HH21 microprocessor thermometer after inserting a 0.020 in. diameter, 6 in. long thermocouple probe (Omega) into the gel at different horizontal positions, and was found to be linear with position. Linear diagonal tracking of the dyes in the gel at the end of an electrophoresis run was also consistent with a linear gradient.

Melting experiments by TGGE are confined to a limited temperature regime. Convenient lower and upper temperatures for the externally circulating water baths were found to be 4 and 65 °C, respectively. Temperatures lower than 4 °C led to excessive condensation on the apparatus and difficulty in holding a steady temperature, while temperatures higher than 65 °C stressed the glass plates. For these bath settings, temperatures measured with the thermocouple inserted directly in the gel at the edges of the single wide sample lane were found to be 15.3 and 54.1 °C. For a 2 h electrophoresis run, there was no fluctuation in the temperatures measured at various points inside of the gel (Z. Shu, unpublished data).

Standard gel conditions consisted of an 8% polyacrylamide gel with an acrylamide to bis-acrylamide ratio of 29:1. The standard buffer in the gel and the reservoirs was 1×$TBEK_{50}$ (100 mM Tris, 83 mM boric acid, 1 mM Na_2EDTA, 50 mM KCl), pH 8.77 at 20 °C, and 4 M urea. The pH dependence of 1×$TBEK_{50}$ was found to be −0.014 ΔpH/°C, providing a pH of 8.85 at the gel temperature of 15.3 °C, and a pH of 8.30 at the gel temperature of 54.1 °C. Either 0.5 mm or 1.5 mm thick spacers were used between the glass plates. Gels were preelectrophoresed for 30 min at 350 V with water baths running prior to loading of samples. To permit the RNAs to fold into the proper hairpin conformation, at the start of each experiment the appropriate mixture of RNAs was renatured in TE in the same tube by heating for 3 min at 90 °C followed by cooling on the bench for at least 10 min. Samples were then mixed with TGGE loading buffer (final concentration: 1×$TBEK_{50}$, 10% glycerol, 0.05% xylene cyanol, and 0.05% bromophenol blue) prior to

[1] Abbreviations: TGGE, temperature gradient gel electrophoresis; PAGE, polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; TE, Tris−EDTA; $TBEK_{50}$, Tris−boric acid−EDTA−50 mM KCl buffer; $P_{10}N_{100}E_{0.1}$, 10 mM sodium phosphate−100 mM NaCl−0.1 mM EDTA buffer; RT, reverse transcription; PCR, polymerase chain reaction; T_M, melting temperature.

Table 1: Effects of Salt and Urea on Perpendicular TGGE[a]

		R_f			T_M (°C)		
salt[a]	[urea],M	lower base line	upper base line	% hypomobility[b]	UA	GC	GU
K_{50}	0	1.06	0.85	19.8	48.6	53.8	45.4
K_{50}	1	1.03	0.82	20.4	43.9	49.2	40.8
K_{50}	2	0.96	0.81	15.6	40.2	46.9	37.3
K_{50}	4	0.94 ± 0.02[c]	0.81 ± 0.02	14.1 ± 2.0	32.7 ± 0.7	38.6 ± 0.6	29.0 ± 0.6
K_{50}	6	0.96	0.85	11.5	27.1	34.6	22.9
−	4				*d*		
Na_{10}	4				*d*		
Mg_2	4	0.63	0.53	15.9	43.7		
Mg_2	8	0.59	0.52	11.9	26.5		
Mg_5	4	0.57	0.45	21.0	49.3		
Mg_5	6	0.55	0.45	18.2	42.6		
Mg_5K_{50}	6	0.63	0.55	12.7	42.6		
Mg_5K_{100}	6				38.6		

[a] Experiments were in 1×TBE plus the indicated salt, present in both the electrophoresis buffer and the gel. K, Na, and Mg indicate KCl, NaCl, and $MgCl_2$, respectively. Subscripts indicate concentrations of salts in millimolar. [b] Percent hypomobility = $\{[R_f(\text{lower}) - R_f(\text{upper})]/R_f(\text{lower})\} \times 100$, and is defined as the percent decrease in mobility upon melting. Within experimental error, percent hypomobility was identical for the three sequences shown. [c] Errors are standard deviations for three separate measurements. [d] No transition.

loading. For perpendicular TGGE, 150 μL of sample was loaded into a single well across the top of the gel. For parallel TGGE, 2 μL of sample was loaded into each 10 mm wide lane. Each gel was run at constant voltage (350 V) until the bromophenol blue ran just below the aluminum plate (approximately 2 h). Every 60 min during electrophoresis, the upper and lower reservoirs were emptied, buffers mixed, and reservoirs refilled. The gel was dried on Whatman 3M paper and analyzed using a PhosphorImager (Molecular Dynamics). Any change in salt or urea concentration is noted in the appropriate legend.

Analysis of TGGE Data. In Table 1, relative mobility (R_f) values are reported for a variety of electrophoresis conditions. Values for R_f were calculated at the end of the 2 h run by dividing the distance from the bottom of the sample well to the center of the RNA band by the distance from the bottom of the sample well to the center of the bromophenol blue tracking dye. Values of R_f for lower and upper base lines were independent of the positions along each base line. Reported values for R_f were generally determined at a position near the center of the linear portion of the baseline. Melting temperatures (T_M) were determined by drawing a straight line through the lower base line and a straight line through the upper base line, and determining the point on the RNA transition that was equidistant from the two lines drawn through each base line.

UV-Melting Experiments. RNAs used in UV-melting studies were precipitated using NaCl and ethanol, and dissolved in melt buffer $P_{10}N_{100}E_{0.1}$ (10 mM sodium phosphate, 100 mM NaCl, 0.1 mM Na_2EDTA, pH 7.0). UV absorbance melting profiles were obtained at 260 nm using a Gilford Response II Spectrophotometer and 10, 5, or 1 mm path length cuvettes. At the start of each experiment, the desired RNA was renatured in melt buffer by heating for 3 min at 90 °C followed by cooling on the bench for at least 15 min. During the melting experiment, the heating rate was 0.93 °C/min. Similar results were obtained in an experiment for the sequence where XY = UA (see Figure 1A) using a heating rate of 0.24 °C/min, suggesting the transitions represent a thermodynamic equilibrium. To determine if any duplexes were present, the absorbance melting profile for each molecule was measured over a 100-fold range in concentration, 0.3−30 μM strand concentration. The similarity of the melting profiles indicated that the transitions correspond to unimolecular species. To favor unimolecular transitions, thermodynamic parameters were calculated using 0.3 and 0.6 μM oligomer concentrations only, and are the average of single melts of at least four independently prepared samples. Thermodynamic parameters were obtained by fit to a two-state model with sloping base lines using a nonlinear least-squares program (*17*).

RESULTS

Design of RNA Oligomers. Model RNA hairpins (Figure 1A) were designed to facilitate reverse transcription (RT) and polymerase chain reactions (PCR) necessary for in vitro selection experiments, to minimize alternative conformations upon folding, and to provide a reasonable change in electrophoretic mobility upon melting. The 5′- and 3′-ends contain six nucleotides that are engineered to be single-stranded to facilitate binding of RT-PCR primers to the stem of the hairpin. Moreover, T7 RNA polymerase can add extra nucleotides to the 5′- and 3′-ends of its transcripts (*18*−*20*). Positioning of the 5′- and 3′-ends of the RNA six nucleotides from the base of the stem thus minimizes the direct influence such heterogeneities have on the stability of the hairpin through either stacking on the end of the stem or base pairing with single-stranded nucleotides across the base of the stem.

The 5′-single-stranded region GGGAGA is an optimal T7 polymerase start sequence (*18*), and the 3′-single-stranded GCAAGG is designed to prohibit alternative foldings and interactions with the 5′-single-stranded region. In addition, the order and number of the AU and UA base pairs in the loop-proximal portion of the stem were designed to disfavor slippage of stem base pairs in potential alternative foldings induced by various nucleotide library insertions. For example, thermodynamic parameters calculated by free energy minimization with a window size of zero (*21*) reveal a 5.1 kcal/mol difference in ΔG°_{37} between the most stable and the second most stable folding for the hairpin in Figure 1A with XY substituted as UA. This calculation indicates that the most stable structure is favored by approximately 4000-fold at 37 °C.

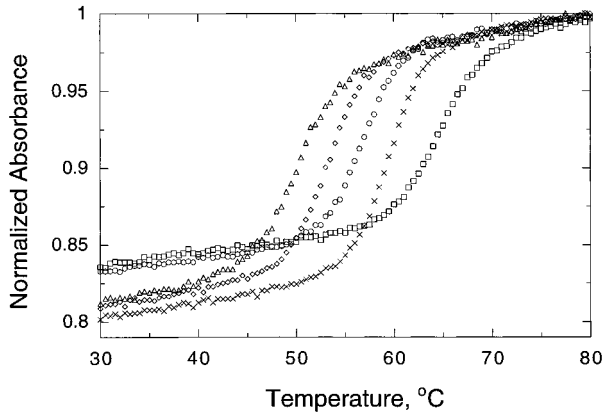

FIGURE 2: Normalized melting curves. Plot of normalized absorbance versus temperature. Substitutions of XY (Figure 1A) are denoted as follows: GA (triangle), GC (square), GG (diamond), GU (circle), and UA ("×"). Buffer was $P_{10}N_{100}E_{0.1}$ (10 mM sodium phosphate, 100 mM sodium chloride, 0.1 mM Na_2EDTA, pH 7.0).

The loop-proximal portion of the stem consists of a series of AU and UA base pairs. A set of six base pairs was chosen to give a helix long enough to provide a reasonable change in electrophoretic mobility upon thermal melting of the RNA. The percent hypomobility for the melting of the model oligomers found by perpendicular TGGE experiments ranged from 11.5 to 21% depending on salt and urea concentrations (Table 1). These are similar to percent hypochromicities found for UV melting of these model oligomers (Figure 2).

The loop-distal portion of the stem contains a *Bam*HI site near the 3′-end of the transcript to facilitate cloning for in vitro selection experiments; a second potential *Bam*HI site near the 5′-end of the transcript has been eliminated by incorporation of UG wobble pairs. The loop sequence "GAAA" is phylogenetically conserved (*22*) and has a compact structure (*23*).

Perpendicular TGGE. Temperature gradients perpendicular and parallel to the electric field were used to analyze the thermodynamic behavior of the model RNA hairpins. Results of experiments with a temperature gradient perpendicular to the electric field using GC-, GU-, and UA-substituted oligomers are shown in Figure 1A. Three well-separated transitions clustered near the center of the gel were observed by perpendicular TGGE. Upper and lower base lines were coincident for all three hairpins consistent with similar folded and unfolded structures. The steeply sloping base lines are due to a general increase in electrophoretic mobility with temperature, and were observed for molecules not undergoing a transition as well as for the tracking dyes used in the gel (data not shown).

For a mixture of the three oligomers, the order of stabilities observed by perpendicular TGGE is GC > UA > GU. The order of stabilities was determined by loading the three oligomers in a 1:2:3 ratio in a perpendicular TGGE experiment and quantitating to determine the identity of each transition (data not shown).

Thermodynamic parameters for model nucleic acids can be obtained by equilibrium UV-melting experiments (*6, 24*−*26*). Typical UV melts for model oligomers are shown in Figure 2, and thermodynamic parameters are reported in Table 2. A linear correlation was observed between the T_Ms obtained by the two methods (Figure 3), although differences in the magnitudes of the T_Ms exist. Additionally, a nonlinear

Table 2: Thermodynamic Parameters for Hairpin Formation Measured by UV Melting[a]

XY[b]	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (eu)	T_M[c] (°C)
GC	−7.7 ± 0.23	−93.6 ± 2.9	−277.0 ± 8.7	64.7
GU	−6.4 ± 0.07	−108.0 ± 1.2	−327.4 ± 3.7	56.7
GG	−5.3 ± 0.18	−107.3 ± 3.1	−328.8 ± 9.4	53.3
GA	−3.6 ± 0.10	−90.4 ± 2.0	−280.0 ± 6.2	49.9
CG	−8.0 ± 0.48	−94.2 ± 5.2	−278.0 ± 15.4	65.6
AU	−7.7 ± 0.11	−110.2 ± 1.4	−330.4 ± 4.3	60.4
UA	−7.2 ± 0.25	−105.1 ± 3.8	−315.7 ± 11.6	59.6

[a] Solutions are 10 mM sodium phosphate, 100 mM NaCl, 0.1 mM Na_2EDTA, pH 7.0. To favor unimolecular transitions, thermodynamic parameters were calculated at the lowest strand concentrations, between 0.3 and 0.6 μM oligomer. Parameters are the average of single melts of at least four independently prepared samples, and errors are standard deviations. The range in T_Ms among the four or more independently prepared low concentration samples was less than 0.5 °C. The first four sequences are members of the combinatorial library and listed in order of decreasing stability. [b] XY corresponds to substitutions in the model oligomer (Figure 1A). [c] Calculated for 10^{-4} M strand concentration.

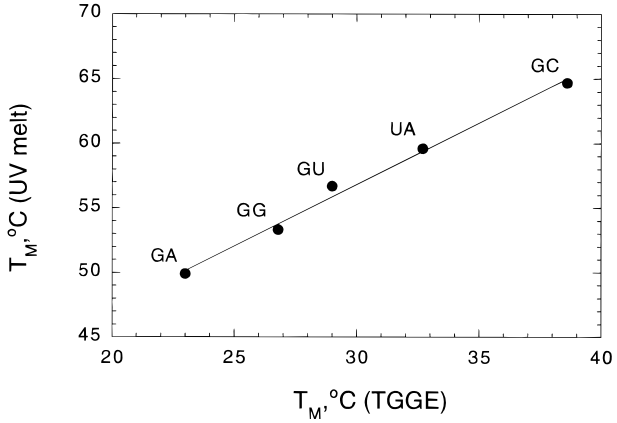

FIGURE 3: Comparison of melting temperatures by perpendicular TGGE and UV-melting. Plot of melting temperature (T_M) determined by UV-melting in melt buffer $P_{10}N_{100}E_{0.1}$ versus T_M determined by perpendicular TGGE in 1×TBEK_{50}/4M urea for five different oligomers. Substitutions of XY (Figure 1A) are denoted in the figure.

correlation was observed between the TGGE-derived T_Ms and the $\Delta G°_{37}$ determined by UV melting, suggesting that the ordering of melting determined by TGGE in urea-containing gels correlates with the ordering of free energies determined by UV-melting experiments.

The dependence of the perpendicular TGGE T_M on the concentration of urea in the gel was examined. As shown in Figure 4, a linear decrease in T_M with increase in urea concentration was observed. In addition, the urea dependencies of the T_Ms for the three oligomers tested are very similar. There is a modest decrease in the percent hypomobility of the hairpin with increase in urea concentration (Table 1), but the change in mobility remains adequate for separation of the oligomers. These data indicate that urea concentration can be used to tune the T_M for an oligomer into the experimentally optimal range for TGGE while maintaining the correct order of stabilities.

The effects of salt concentration and composition on melting behavior were also examined (Table 1). The buffer for all conditions is 1×TBE, and the UA-substituted oligomer was examined at all salt conditions tested. In a background

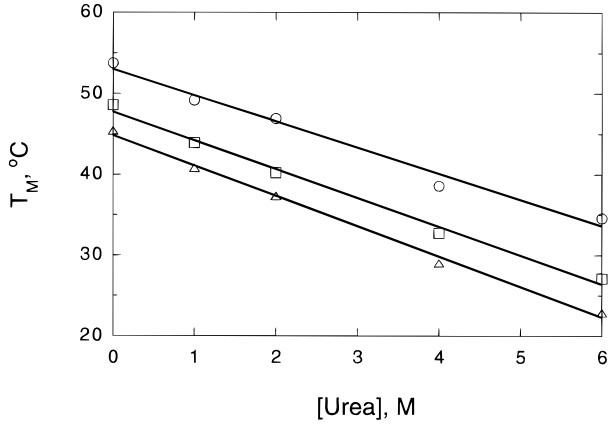

FIGURE 4: Dependence of melting temperature on urea concentration. Plots of melting temperature (T_M) versus urea concentration. As necessary, incomplete base lines for melts with very high or very low T_Ms were extrapolated from base lines observed for more stable or less stable oligomers of identical length run on the same gel. Substitutions of XY (Figure 1A) with GC, UA, and GU are denoted with a circle, square, and triangle, respectively. All experiments were done in 1×TBEK$_{50}$.

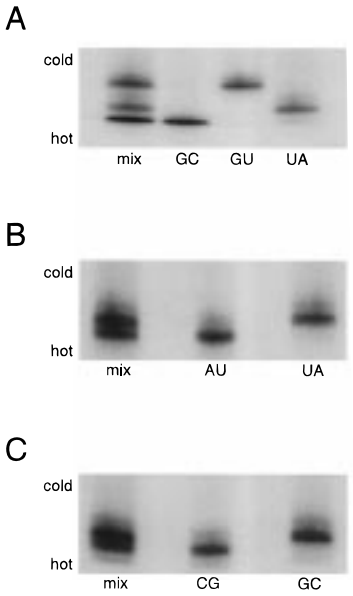

FIGURE 5: Separation of model RNA hairpins by parallel TGGE. RNA oligomers were 5′-^{32}P-labeled, with XY (Figure 1A) substituted as indicated. Buffer was 1×TBEK$_{50}$/4 M urea. (A) Separation of GC, GU, UA, and a mixture of the three. Low- and high-temperature baths were set at 25 and 35 °C. (B) Separation of AU, UA, and a mixture of the two. Low- and high-temperature baths were set at 29 and 36 °C. (C) Separation of CG, GC and a mixture of the two. Low- and high-temperature baths were set at 34 and 41 °C.

of 4 M urea with either no additional salt or only 10 mM NaCl, no melting transition was observed by perpendicular TGGE. Conditions of 50 mM KCl/4 M urea provided a melt for the UA-substituted oligomer with a T_m of 32.7 °C. In addition, melting of GN-substituted oligomers was also observed over this range.

Ribonucleic acids often rely on divalent metal ions for formation of tertiary structure, and they can form metal-specific binding sites (*27*, *28*). To test whether TGGE could potentially be used to look at such structures with these model oligomers, perpendicular TGGE was carried out at various magnesium chloride and urea concentrations. For the UA-substituted oligomer, a folding transition was observed with $MgCl_2$ concentrations of 2 and 5 mM. At 4 M urea, 2 mM $MgCl_2$ provided more stability than 50 mM KCl, and gave a similar T_M as conditions of 50 mM KCl at 1 M urea. Stability was increased upon raising the concentration of $MgCl_2$ to 5 mM at 4 M urea, although no further increase was observed upon addition of 50 mM KCl to 5 mM Mg^{2+} at 6 M urea. At 6 M urea, a decrease in stability was observed upon addition of 100 mM KCl to 5 mM Mg^{2+} (Table 2). Experiments in magnesium chloride revealed scattered counts below the transition curve which increased in intensity with temperature, presumably caused by Mg^{2+}-catalyzed hydrolysis of the RNA. Hydrolysis of a mixture of RNAs was approximately 20%, as measured in 2 mM Mg^{2+}/8 M urea gels at 50 °C in the gel.

Parallel TGGE. The low and high temperatures for the melting transition obtained by perpendicular TGGE experiments were determined by interpolation and used to calculate set temperatures for the water baths for parallel TGGE in identical buffer conditions of 1×TBEK$_{50}$/4 M urea. Shown in Figure 5 are results of experiments with a temperature gradient parallel with the electric field. On these gels, the most stable oligomers have the fastest mobility, and the least stable oligomers have the slowest mobility. GC-, GU-, and UA-substituted oligomers run primarily as a single band, consistent with a single conformation. In addition, a mixture of the three RNAs results in three bands coincident with the individual GC-, GU-, and UA-substituted oligomers. The order of stabilities is GC > UA > GU, consistent with the perpendicular TGGE and UV-melting experiments.

Parallel TGGE experiments were also carried out with transversions of AU and GC base pairs. The order of stabilities observed is AU > UA (Figure 5B), and CG > GC (Figure 5C), although the separation of the latter mixture is modest. UV-melting experiments were performed on the same oligomers, and thermodynamic parameters were extracted (Table 2). A correlation is observed between the stabilities obtained by the two methods. In addition, differences in $\Delta G°_{37}$ between transversions of AU or GC base pairs determined by UV-melting experiments using a melt buffer containing 100 mM NaCl were only 0.5 and 0.3 kcal/mol, respectively, demonstrating the sensitivity of this approach.

TGGE of the Combinatorial Library. To determine if TGGE methods were capable of separating an RNA library into stable and unstable members, a simple combinatorial library was prepared. Residues X and Y (Figure 1A) were replaced with G and N, respectively, where N is a mixture of the four bases. The library of RNAs was then subjected to perpendicular and parallel TGGE (Figure 6), and both methods resulted in four well-resolved species. Parallel TGGE of the library with preparations of each of the individual RNAs revealed that the order of stabilities is GC > GU > GG > GA (Figure 6B). Individual RNAs run primarily as a single band, consistent with a single conformation, and the RNA library resulted in only four bands, each coincident with one of the individual GA-, GC-, GG-, and GU-substituted oligomers (Figure 6B). In addition, a linear correlation was observed between TGGE- and UV-derived T_Ms for the combinatorial library (Figure 3), and a nonlinear correlation was observed between TGGE-derived T_Ms and the $\Delta G°_{37}$ determined by UV melting.

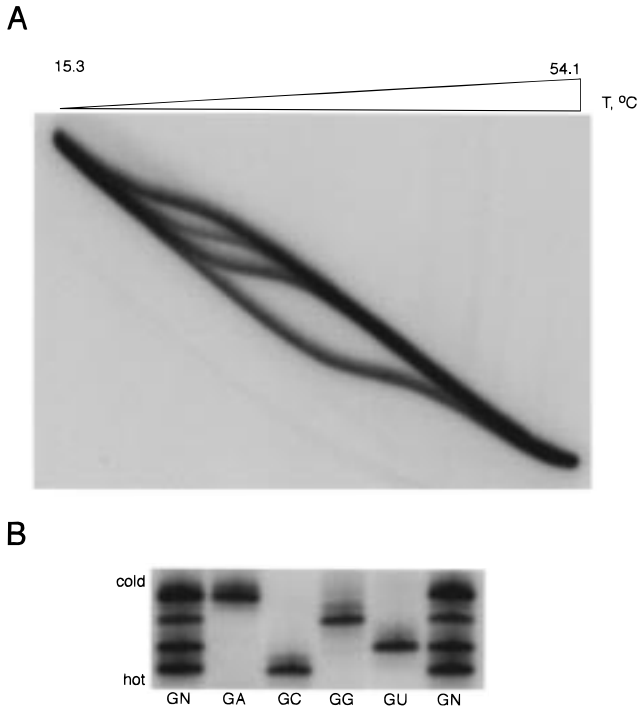

FIGURE 6: Separation of the combinatorial RNA library. RNA oligomers were 5′-^{32}P-labeled, with XY (Figure 1A) substituted as follows: X = G, and Y = N, a mix of A, C, G, and U. Buffer was 1×TBEK_{50}/4 M urea. (A) Perpendicular TGGE. The temperature in the gel ranged from 15.3 to 54.1 °C. (B) Parallel TGGE. Separation of the library is indicated in the first and last lanes. Positions of GA, GC, GG, and GU substitutions are indicated. Low- and high-temperature baths were set to 18 and 32 °C.

The finding that the GG-substituted oligomer was more stable than the GA-substituted oligomer was a surprising result, as the free energy minimization program *Mfold* v3.0 currently predicts the GA-substituted oligomer to have the same $\Delta G°_{37}$ as the GG-substituted oligomer (*21*). UV-melting experiments were conducted on the individual RNAs and indicated that the $\Delta G°_{37}$ and T_Ms are 1.7 kcal/mol and 3.4 °C more favorable for the GG-substituted oligomer than the GA-substituted oligomer (Figure 2, Table 2), consistent with the TGGE data. Recent data from Zhu and Wartell (*29*) using TGGE on long dsRNA with the same closing base pairs are also consistent with the order of stabilities observed here, GG > GA.

DISCUSSION

An enormous amount of nucleic acid sequence has been reported in recent years, including complete genomes from 15 organisms. The ongoing human genome initiative will eventually result in the determination of approximately 3 billion bases of sequence. One way to begin to make sense of this vast genomic information is to develop rapid and accurate structural models for nucleic acids. Prediction of RNA structure from sequence requires thermodynamic parameters for a wide variety of motifs, many of which are sufficiently complex as to require combinatorial approaches to study. We report here an approach for the analysis of the thermodynamics of combinatorial libraries in nucleic acids. Temperature gradient gel electrophoresis (TGGE) has been used to separate a simple RNA combinatorial library into stable and unstable members. Examination of a simple combinatorial library by parallel TGGE suggests new information about the relative stability of mismatches in RNA, namely, that for the flanking base pairs examined here, a GG-mismatch is more stable than a GA-mismatch, data that were subsequently confirmed by UV-melting. This illustrates the sensitivity of TGGE for resolving randomized nucleic acid libraries according to their equilibrium thermodynamic stabilities.

The ability of parallel and perpendicular TGGE to separate RNA hairpins differing in sequence at only one base pair has been demonstrated. Initial experiments involved perpendicular TGGE and showed well-resolved melting curves for GC-, GU-, and UA-substituted oligomers (Figure 1). The temperature extremes of the transitions were determined by linear extrapolation, and used as the low and high temperatures for parallel TGGE. Parallel TGGE experiments likewise resulted in well-resolved bands for lanes containing either single oligomers or mixtures of the three (Figure 5). Parallel TGGE offers the advantage that multiple lanes can be run on a single gel, allowing simultaneous examination of mixtures and individual components. Hairpins with only a single base pair or mismatch difference were resolved by parallel TGGE (Figures 5 and 6), including GC and AU base pair transversions differing by only 0.3−0.5 kcal/mol in $\Delta G°_{37}$ (Table 2).

Separation of DNA molecules that differ by only one base pair was initially demonstrated by denaturing gradient gel electrophoresis (DGGE), in which a gradient of urea/formamide linearly increases in the gel perpendicular to the electric field (*30*−*33*). TGGE was subsequently developed as a more robust method for investigating the melting behavior of a number of long DNAs and RNAs (*34*, *35*). In particular, a thermal gradient is more reproducible and easier to change than a denaturing gradient. Systems studied by TGGE have included viroids (small circular RNAs of about 350 nucleotides), viral dsRNAs (at least 335 bp), and DNA−protein interactions (*35*). TGGE has also been used to analyze the thermodynamic stability of 373 bp dsDNAs (*16*, *36*−*38*), and 345 bp dsRNAs (*29*). Melting of these long nucleic acids is non-two-state; however, by placing various base pairs, bulges, and mismatches in early melting domains, molecules have been separated based on small differences in thermodynamic stability (*16*, *29*, *36*−*39*). TGGE experiments on long nucleic acids were performed using low concentrations of TBE buffer in the absence of any additional salts.

To be able to identify the stable and unstable members from various complex combinatorial libraries by in vitro selection, we have adapted TGGE for use with a model system consisting of an RNA hairpin. Such model systems also have the obvious advantage that they allow the thermodynamics of various size hairpin loops to be examined. To facilitate two-state melting behavior, the RNA hairpins were designed to have 13 bp rather than several hundred. Experiments on these short RNA hairpins were shown to require the addition of salt to the TBE buffer to obtain a melting transition (Table 1), and gels were run with a variety of monovalent and divalent salts at concentrations typically required to fold such RNAs.

Several lines of evidence are consistent with two-state melting of model hairpins. Melts detected by two different methods, perpendicular TGGE (Figure 1B) and UV absor-

bance (Figure 2), show only one transition, providing no direct evidence for a third state. This is in contrast to large RNA and DNA systems which show non-two-state melting behavior (*16, 35, 39*). In addition, parallel and perpendicular TGGE melting experiments and UV-melting experiments are sensitive to both substitutions in the stem (Figures 1, 5, 6) and the loop of the hairpin (J. Bevilacqua et al., unpublished data), indicating that the observed melting transition is reporting on the stability of two disparate structural features.

To test whether samples behaved as hairpins or self-complementary duplexes, the concentration dependence of the UV melts was studied. Within experimental error, fits of UV-melts to a two-state model gave identical T_Ms from 0.3 to 30 μM strand concentrations. If oligomers were forming self-complementary duplexes, calculations using nearest-neighbor thermodynamic values determined at 1 M NaCl predict that the T_M should change by 4 °C for the range of oligomer concentrations used (*40*), well outside the error limits of the measured T_Ms. In addition, a mixture of GC-, GU-, and UA-substituted oligomers, and a mixture of GN-substituted oligomers, heated to 90 °C for 3 min and cooled on the bench for at least 10 min prior to electrophoresis, resulted in only three and four bands by parallel TGGE, respectively (Figures 5A, 6B). If the samples were duplexes, such treatment would likely result in a mixture of homo- and heteroduplexes, resulting in extra bands by parallel TGGE. In sum, all the data are consistent with a unimolecular, two-state transition.

The dependence of TGGE melting behavior on salt concentration revealed that 50 mM KCl is adequate to obtain a good melting transition (Table 1). Potassium was chosen as the cation since it is the predominant monovalent cation inside a typical mammalian cell (*41*). A similar order of stabilities is seen for a series of hairpins between 0.01 and 1 M NaCl (*12, 13*), suggesting that 50 mM is a reasonable salt concentration to choose.

Tertiary structure formation in RNA often requires divalent ions (*27, 28*). In particular, TGGE has been used with varying amounts of Mg^{2+} to investigate the thermodynamic stability of the tertiary structure from the P4−P6 domain of the *Tetrahymena* ribozyme (*42*). For the melting of RNA secondary structure in the constructs reported here, inclusion of $MgCl_2$ in the gel and electrophoresis buffer led to similar percent hypomobility shifts as KCl (Table 1). In addition, 2 mM $MgCl_2$ provided more folding stability than 50 mM KCl under conditions of 4 M urea, and increasing the [Mg^{2+}] to 5 mM resulted in further stabilization. Furthermore, addition of 100 mM KCl to 5 mM $MgCl_2$ at 6 M urea resulted in a decrease of 4 °C in T_M (Table 1). According to polyelectrolyte theory developed by Manning, this effect is caused by added KCl preferentially stabilizing the unfolded hairpin without affecting the entropy of mixing of Mg^{2+} ions associated with the folded hairpin (*43*−*45*).

Experiments carried out in Mg^{2+}-containing buffers resulted in hydrolysis of the RNA, especially at the high end of the temperature gradient. For example, there was approximately 20% hydrolysis in 2 mM Mg^{2+}/8 M urea gels at 50 °C in the gel. Although most of the cleaved transcripts will not be substrates for RT-PCR, there are generally many copies of each sequence; thus, 20% hydrolysis should not be a major deterrent to in vitro selection. In addition, the percent hydrolysis can potentially be suppressed by using an electrophoresis buffer with a lower pH, and tertiary structures often melt at lower temperatures than secondary structures (*42, 46, 47*).

Melting temperature data for perpendicular TGGE- and UV-melting experiments indicate that the T_M determined by TGGE correlates linearly with that determined by UV melting (Figure 3). There are differences in the magnitudes of the T_Ms observed by the two methods. Differences may be due in part to the presence of urea in the gels, although 0 M urea gels result in T_Ms that are still approximately 10 °C lower than by UV melting. Other effects may be caused by the lower concentrations of salt in TGGE (50 mM KCl) versus UV-melting experiments (100 mM NaCl), or due to nonequilibrium conditions in the gel. Despite differences in the magnitudes of the T_Ms, the linear correlation between T_Ms determined by the two approaches suggests that TGGE experiments resolve mixtures of oligomers in the same order as UV-melting experiments. In addition, the TGGE-determined T_Ms are linearly dependent on the concentration of urea in the gel, giving essentially identical slopes for the three sequences tested (Figure 4). Therefore, a linear correlation exists between T_Ms determined by UV melting and TGGE at all concentrations of urea.

For similar RNA sequences in which $\Delta S°$ does not change substantially between oligomers, changes in T_M between oligomers are proportional to $\Delta\Delta G°$ values for the oligomers (*29*). A general decrease in $\Delta G°_{37}$ with increase in TGGE-derived T_M was observed; however, the decrease leveled at higher T_Ms. Although the correlation was nonlinear under these conditions, TGGE in urea-containing gels resolves the mixture of similar oligomers studied here in the same order as the $-\Delta G°_{37}$ from the UV-melting experiment.

The thermodynamic analysis of the simple RNA combinatorial library reported herein lays the groundwork for the separation of complex nucleic acid libraries into populations of stable and unstable molecules. In particular, the model RNA hairpins studied here are suitable candidates for the RT-PCR and cloning steps of in vitro selection (*48, 49*) (J. Bevilacqua et al., unpublished data). RNA oligomers differing by only 0.3 kcal/mol in $\Delta G°_{37}$ have been separated by our method. Ultrastable RNA tetraloops, for example, are approximately −2.7 to −4.0 kcal/mol more stable in terms of $\Delta G°_{37}$ than other tetraloops (*12*), and should be separable. Use of in vitro selection with a separation step by TGGE will permit rapid and efficient searches of complex secondary and tertiary motifs for the identification of rare thermodynamically stable and unstable sequences. Thermodynamic parameters for such sequences may then be extracted by equilibrium experiments for use in structure prediction algorithms, and oligomers can be structurally characterized to reveal potential roles in RNA folding and structure formation. In addition, such studies may help reveal the roles stable and unstable RNA motifs play in biological activity. The thermodynamic properties of other polymers could potentially be studied by this approach, including DNAs and proteins.

ACKNOWLEDGMENT

We thank Janell Heffner, Jing Liang, David Proctor, Zhanyong Shu, Tom Cech, and Doug Turner for comments on the manuscript.

REFERENCES

1. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
2. Zuker, M. (1989) *Science 244*, 48−52.
3. Mathews, D. H., Andre, T. C., Kim, J., Turner, D. H., and Zuker, M. (1998) in *Molecular Modeling of Nucleic Acids* (Leontis, N. B., and SantaLucia, J., Jr., Ed.) pp 246−257, American Chemical Society, Washington, D.C.
4. Gultyaev, A. P., van Batenburg, F. H., and Pleij, C. W. A. (1995) *J. Mol. Biol. 250*, 37−51.
5. van Batenburg, F. H. D., Gultyaev, A. P., and Pleij, C. W. A. (1995) *J. Theor. Biol. 174*, 269−280.
6. Serra, M. J., and Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
7. Konings, D. A., and Gutell, R. R. (1995) *RNA 1*, 559−574.
8. Fields, D. S., and Gutell, R. R. (1996) *Folding Des. 1(6)*, 419−430.
9. Jaeger, J. A., Turner, D. H., and Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
10. Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Müller, P., Mathews, D. H., and Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9218−9222.
11. Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Jr., Brody, E. N., and Gold, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1364−1368.
12. Antao, V. P., Lai, S. Y., and Tinoco, I., Jr. (1991) *Nucleic Acids Res. 19*, 5901−5905.
13. Antao, V. P., and Tinoco, I., Jr. (1992) *Nucleic Acids Res. 20*, 819−824.
14. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
15. Bevilacqua, P. C., and Cech, T. R. (1996) *Biochemistry 35*, 9983−9994.
16. Wartell, R. M., Hosseini, S. H., and Moran, C. P., Jr. (1990) *Nucleic Acids Res. 18*, 2699−2705.
17. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
18. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8798.
19. Moroney, S. E., and Piccirilli, J. A. (1991) *Biochemistry 30*, 10343−10349.
20. Ferré-D'Amaré, A. R., and Doudna, J. A. (1996) *Nucleic Acids Res. 24*, 977−978.
21. Jaeger, J. A., Turner, D. H., and Zuker, M. (1990) *Methods Enzymol. 183*, 281−306.
22. Woese, C. R., Winker, S., and Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467−8471.
23. Heus, H. A., and Pardi, A. (1991) *Science 253*, 191−194.
24. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 256−263.
25. Puglisi, J. D., and Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304−325.
26. SantaLucia, J., Allawi, H. T., and Seneviratne, P. A. (1996) *Biochemistry 35*, 3555−3562.
27. Cate, J. H., and Doudna, J. A. (1996) *Structure 4*, 1221−1229.
28. Pyle, A. M. (1996) *Met. Ions Biol. Syst. 32*, 479−520.
29. Zhu, J., and Wartell, R. M. (1997) *Biochemistry 36*, 15326−15335.
30. Fischer, S. G., and Lerman, L. S. (1979) *Cell 16*, 191−200.
31. Fischer, S. G., and Lerman, L. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1579−1583.
32. Myers, R. M., Fischer, S. G., Maniatis, T., and Lerman, L. S. (1985) *Nucleic Acids Res. 13*, 3111−3129.
33. Myers, R. M., Fischer, S. G., Lerman, L. S., and Maniatis, T. (1985) *Nucleic Acids Res. 13*, 3131−3145.
34. Thatcher, D. R., and Hodson, B. (1981) *Biochem. J. 197*, 105−109.
35. Rosenbaum, V., and Riesner, D. (1987) *Biophys. Chem. 26*, 235−246.
36. Ke, S. H., and Wartell, R. M. (1993) *Nucleic Acids Res. 21*, 5137−5143.
37. Ke, S. H., and Wartell, R. M. (1995) *Biochemistry 34*, 4593−4600.
38. Ke, S. H., and Wartell, R. M. (1996) *Nucleic Acids Res. 24*, 707−712.
39. Steger, G., Po, T., Kaper, J., and Riesner, D. (1987) *Nucleic Acids Res. 15*, 5085−5103.
40. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
41. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994) in *Molecular Biology of the Cell*, p 508, Garland Publishing, Inc., New York.
42. Szewczak, A. A., Podell, E. R., Bevilacqua, P. C., and Cech, T. R. (1998) *Biochemistry 37*, 11162−11170.
43. Manning, G. S. (1972) *Biopolymers 11*, 951−955.
44. de Marky, N., and Manning, G. S. (1975) *Biopolymers 14*, 1407−1422.
45. Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179−246.
46. Banerjee, A. R., Jaeger, J. A., and Turner, D. H. (1993) *Biochemistry 32*, 153−163.
47. Gluick, T. C., and Draper, D. E. (1994) *J. Mol. Biol. 241*, 246−262.
48. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
49. Bartel, D. P., and Szostak, J. W. (1993) *Science 261*, 1411−1418.

BI981732V