Interpreting Oligonucleotide Microarray Data To Determine RNA Secondary Structure: Application to the 3′ End of *Bombyx mori* R2 RNA†

Shenghua Duan,‡ David H. Mathews,§,|| and Douglas H. Turner*,‡,§

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, and Center for Pediatric Biomedical Research, Department of Pediatrics, and Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, New York, 14642

Received December 22, 2005; Revised Manuscript Received May 11, 2006

ABSTRACT: A method to deduce RNA secondary structure on the basis of data from microarrays of 2′-O-methyl RNA 9-mers immobilized in agarose film on glass slides is tested with a 249 nucleotide RNA from the 3′ end of the R2 retrotransposon from *Bombyx mori*. Various algorithms incorporating binding data and free-energy minimization calculations were compared for interpreting the data to provide possible secondary structures. Two different methods give structures with 100 and 87% of the base pairs determined by sequence comparison. In contrast, structures predicted by free-energy minimization alone by Mfold and RNAstructure contain 52 and 72% of the known base pairs, respectively. This combination of high throughput microarray techniques with algorithms using free-energy calculations has potential to allow for fast determination of RNA secondary structure. It should also facilitate the design of antisense and siRNA oligonucleotides.

RNA serves many vital roles in life, including protein synthesis, gene silencing, gene editing, gene transcription control, and catalysis of chemical reactions (*1*−*11*). These different functions of RNA require different structures (*12*−*14*). Moreover, RNA−RNA interactions required for some functions such as those involving siRNA (*3*, *4*, *15*−*18*) and miRNA (*19*−*21*) are influenced by RNA structures (*22*−*27*).

Two- and three-dimensional structures are required to deduce RNA structure−function relationships (*12*). Most RNA three-dimensional structures are determined by nuclear magnetic resonance (NMR) (*28*−*32*) or X-ray diffraction (*33*−*36*), but it is not likely that these methods will match the rate at which interesting RNAs are sequenced. Therefore, fast and reliable prediction methods are needed.

Many RNA secondary structures are predicted by comparative sequence analysis (*37*, *38*), which is based on the principle that structures are more conserved than sequences during the evolution of RNAs with the same function (*38*). This method, however, requires substantial numbers of sequences and is difficult to automate.

In the absence of abundant homologous sequences, predicting RNA secondary structures from a single sequence is desirable. The most popular method for this application is free-energy minimization with a dynamic programming algorithm (*39*−*46*). Available programs include Mfold (*40*, *41*, *45*), RNAstructure (*42*), Sfold (*43*), and RNAFold (*44*). This approach works surprisingly well, but there are shortcomings. For example, the free energy of folding is estimated with a simple nearest-neighbor model for which many parameters have not been measured. Motifs that have not been studied extensively include bulge loops with more than one unpaired nucleotide, pseudoknots, and multibranch loops (*47*), which are common in large RNA structures. Chemical mapping (*42*) and enzymatic cleavage (*45*, *48*) have been employed to improve predictions, but these methods are time-consuming and labor intensive.

The accuracy of the RNA secondary-structure prediction from a single sequence has been tested with structures determined by sequence comparison (*42*, *45*). Current dynamic programming algorithms predict the lowest free-energy structure and a set of low energy structures, called suboptimal structures (*40*−*45*, *49*). Often there is a suboptimal structure that is more accurate than the predicted lowest free-energy structure. For example, the predicted lowest free-energy structures for *Escherichia coli* 5S rRNA and dog signal recognition particle (SRP) RNA are about 23% accurate, but the best suboptimal structure is over 90% accurate. Very accurate secondary structures can be generated by free-energy minimization algorithms, but usually they are not the predicted lowest free-energy structures. This offers an opportunity to develop experimental methods that allow for selection of the true secondary structure from among the thousands of predicted possible structures.

Here, we test a microarray approach (*50*) to determine RNA secondary structure by selecting structures compatible with oligonucleotide binding data. Short 2′-O-methyl RNA probes complementary to sequences in a large RNA are immobilized in agarose matrix on glass slides (*51*). The oligonucleotides can bind an RNA target to the chip. The

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

* To whom correspondence should be addressed. Telephone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.

‡ Department of Chemistry.

§ Center for Pediatric Biomedical Research, Department of Pediatrics.

|| Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry.

10.1021/bi052618x CCC: $33.50
Published on Web 07/21/2006

binding signal intensity is dependent upon both the free energy of the duplex formation between the probe and RNA target sequence and the intramolecular pairing in the RNA target. On the basis of the hybridization results, the sequence information can be interpreted into structural information with the aid of secondary-structure prediction by free-energy minimization. There are $4^9 = 262\,144$ possible 9-mer probes. Thus, short probe lengths make a universal microarray chip feasible (*52*). Short oligonucleotide lengths also enhance discrimination against mismatches (*53*, *54*) and prevent self-folding of the probes.

Interpretation of microarray binding data requires computer programs to generate all possible secondary structures and to compare them with the experimental data. The RNA structure 4.1 program (*42*) was modified to rigorously generate all possible secondary structures to eliminate the loss of secondary structures because of the sampling heuristic normally applied by RNAstructure and Mfold. The structures are evaluated by a comparison to the restraints derived from microarray experiments; structures that are consistent with restraints are kept, and structures that are not consistent are rejected. For example, if a 9-mer oligonucleotide binds the target RNA tightly, then a predicted secondary structure is retained only if at least a certain number of unpaired bases are in the region complementary to the 9-mer. The retained structures should be more accurate because they are consistent with the oligonucleotide binding data.

The method was tested by studying hybridization of the *Bombyx mori* R2 3′ untranslated region (UTR) RNA (*55*) to a 2′-O-methyl RNA array chip. R2 is a nonlong terminal repeat retrotransposable element that inserts specifically into the 28S rRNA gene of most insects (*56*). The R2 3′-UTR RNA is recognized by the R2 protein during the target-primed reverse transcription (TPRT) reaction that is responsible for insertion at the host target site (*56*).

MATERIALS AND METHODS

Synthesis of Oligonucleotides. The 2′-O-methyl RNA probes with 5′-end C6 amino linker or fluorescein dye were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method (*57*). Phosphoramidites [e.g., 6-(4-monomethoxytritylamino)hexyl-(2-cyanoethyl)-(*N*,*N*-diisopropyl)-phosphoramidite for 5′-end C6 amino linker and [(3′,6′-dipivaloylfluoresceinyl)-6-carboxamidohexyl]-1-*O*-(2-cyanoethyl)-(*N*,*N*-diisopropyl)-phosphoramidite for 5′-end fluorescein] and CPG support were purchased from Glen Research. Cleavage of the oligonucleotides from universal II CPG was done by incubating with 2 M ammonia in methanol (Sigma−Aldrich) at room temperature for 30 min. The solution was dried down, and 1:1 (v/v) concentrated ammonium hydroxide/ammonia in methylamine (AMA) (Sigma−Aldrich) was added; the sealed vial was incubated at 65 °C for 10 min or room temperature for 2 h. The solution was dried down, and the residue was purified on a large preparative Baker Si500F TLC Silica Gel Plate (20 × 20 cm, 500 μm thick) with a 60:35:10 (v/v/v) 1-propanol/ammonia/water running solution. The product was identified by UV shadowing and scraped from the plate. Oligonucleotides were extracted from the silica with distilled water. The water solution was dried down, and the purified oligonucleotides were dissolved in 5 mM ammonium bicarbonate at pH 7.0. The solution was loaded onto a Waters Sep-Pak C18 chromatography column to remove excess salts. Purities of oligonucleotides were checked with a Hewlett−Packard 1100 HPLC Chemstation. Molecular weights were determined by ESI−MS with a Hewlett−Packard 1100 LCMS Chemstation.

Microarray Chip Fabrication. Agarose film on glass slide was prepared by the Stefan Wolfi method II (*51*). Spotting buffer was 150 mM NaCl, 100 mM $NaHCO_3$ (pH 8.5), and 100 μM oligonucleotide with 5′ C6 amino linker. The spot volume varied from 1 to 0.5 μL. The incubation temperature was 37 °C, which gave better cross-linking than room temperature.

RNA Target Preparation. R2 3′-UTR RNA of *B. mori* was prepared following the protocol of Ruschak et al. (*55*). The *B. mori* R2 3′-UTR sequence was cloned into pUC19 plasmid and transfected into *E. coli* cells. The Qiagen Midi plasmid preparation kit protocol was used for plasmid extraction. The plasmids were linearized by XmnI (Promega). The plasmid DNA was transcribed with T7 RNA polymerase (Ambion MEGAScript), and the RNA was purified on an 8% polyacrylamide denaturing gel. RNA was excised from the gel and eluted by electroelution. The solution was desalted and concentrated on a filter with a 10 K cutoff membrane (Millipore).

The small RNA mimic was purchased from Dharmacon, deprotected, and HPLC-purified.

RNA Target Labeling and Purification. RNA 5′-terminus labeling followed the Invitrogen T4 kinase product protocol. After labeling, a Chroma Spin +STE-10 Column (Clonetech) was used to eliminate excess ATP. The R2 RNA was further purified by PAGE, recovered by crushing and soaking overnight with DEPC-treated water. The solution was desalted and concentrated on a filter with a 10 K cutoff membrane (Millipore).

Hybridization. The labeled RNA (5 pmol) was renatured at 45 °C for 30 min in buffer of 40 mM Hepes, 40 mM sodium Hepes, 110 mM NaCl, and 10 mM $MgCl_2$ at pH 7.4 and slowly cooled to room temperature. This gave more reproducible hybridization results than no renaturation or annealing at 37 °C for 1 h. RNA was incubated with oligonucleotide array chips at 4 °C for 12 h in a probe-clip press-seal incubation chamber (Sigma−Aldrich). For experiments to test cross hybridization, 2 μL of 1 mM free 2′-O-methyl RNA probe was added to 200 μL of the RNA solution after renaturation but before hybridization on the microarray chip.

After hybridization, the slide was removed from the chamber and washed with icy cold hybridization buffer for 5 min. The wet slide was immediately centrifuged to dry in a desktop centrifuge. The dry slide was placed on a phosphorimaging plate for 4 h, and the plate was scanned with a phosphorimager scanner (Molecular Dynamics). Signal intensity was quantified with Imagequant software. The intensities, I, were normalized to the lowest intensity, I_{min}, and plotted as $\log_2(I/I_{min})$. The same procedure was used for initial experiments with 7-, 9-, and 11-mers.

Native Gel-Shift Experiment. The R2 3′-UTR RNA was incubated at 45 °C for 30 min in the same buffer as for the microarray experiment and then slowly cooled to room temperature. Then, 2′-O-methyl RNA probe labeled on the 5′ end with fluorescein was added and incubated at 4 °C for

16 h. Then, the solution was loaded with glycerol onto a 4% polyacylamide native gel. The running buffer was 50 mM tris-acetate at pH 7.8.

Ribonuclease H Cleavage and Primer Extension. The R2 RNA (1 pmol) was incubated at 45 °C for 30 min in the same buffer as for the microarray experiment and then slowly cooled to room temperature. Then, DNA (1 nmol) was added to give a final concentration of 10 μM and incubated for 1 h at room temperature. Then, about 1 unit of Ribonuclease H and 1.25 units of RNAsin were added along with enough DTT to give a final concentration of 1 mM. The solution was left at room temperature for 20 min. Then, the RNA was ethanol-precipitated and resuspended in water to give a final RNA concentration of about 1 μM. The DNA primer in length between 17 and 19 nucleotides was phosphorylated with fresh [γ-^{32}P]ATP, and the solution was filtered with a Chroma Spin +STE-10 column (Clonetech). About 1.5 pmol of ^{32}P-labeled primer was added to a sample containing 1 pmol of RNA; sodium acetate was added to reach 0.5 M concentration in a total volume of 2 μL, and the solution was incubated at 70 °C for 5 min, room temperature for 5 min, and in ice for another 5 min. A mixture of 1 unit of AMV reverse transcriptase (Promega), 1× reverse transcription buffer, and 1 mM of each dNTP was added to the sample to give a final volume of 5 μL. For sequencing lanes, 1 mM of each ddNTP was also added. Reverse transcription reactions were run at 42 °C for 1 h and quenched by an equal volume of loading dye buffer (Ambion), and then, the solution was incubated at 90 °C for 3 min and kept on ice before loading onto an 8% polyacrylamide denaturing gel. The gel was run for 3 h at 100 W with a running buffer of 1× TBE.

UV Melting Experiments and Thermodynamics. Concentrations of single-stranded RNA oligonucleotides were calculated from the absorbance at 280 nm at 80 °C, and extinction coefficients were predicted from those of dinucleoside monophosphates (*58*). The 2′-O-methyl RNA oligonucleotide concentrations were calculated with RNA nearest-neighbor extinction coefficients (*59*). Small mixing errors for non-self-complementary duplexes do not appreciably affect thermodynamic measurements (*60*). Oligonucleotides were lyophilized and redissolved in 110 mM NaCl, 10 mM $MgCl_2$, 40 mM Hepes, and 40 mM sodium Hepes at pH 7.4. Curves of absorbance at 280 nm versus temperature were acquired with a heating rate of 1 °C/min with a Beckman Coulter DU640C spectrophotometer.

Melting curves were fit to a two-state model with MeltWin, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ (*61*−*63*). Additionally, the melting temperatures, T_M, at different concentrations were used to calculate thermodynamic parameters according to $T_M^{-1} = (R/\Delta H°)\ln(C_T/4) + (\Delta S°/\Delta H°)$ (*64*). T_M is the temperature in kelvin at which half of the strands are in duplex.

Generating Possible Structures. Mfold 3.2 for Linux was downloaded from Michael Zuker's webpage (http://www.bioinfo.rpi.edu/~zukerm/rna/mfold-3.html) and installed on a 2.0 GHz P4 Dell computer with the Redhat 8.0 operating system. A total of 1000 secondary structures were generated by setting P = 99, W = 0, and N = 1000.

RNAstructure 4.1 for Windows was downloaded from David Mathews' lab webpage (http://rna.urmc.rochester.edu/rnastructure.html) and installed on a 2.8 GHz P4 Dell computer with the Microsoft Windows XP Professional operating system. A total of 1000 secondary structures were generated by setting max energy difference = 99%, MAX structure number = 1000, and window size = 0.

To predict all suboptimal structures within a small free-energy increment from the lowest free-energy structure, the algorithm of Wuchty et al. (*65*) was implemented in RNAstructure. The free-energy function explicitly considers coaxial stacking of adjacent helices or helices separated by a single noncanonical pair. Coaxial stacking is stacking of the ends of two helices, which has been shown experimentally to stabilize the secondary structure (*66*−*68*). This is the same energy function as that used by RNAstructure for the lowest free-energy structure prediction (*42*) and the partition function calculation of base-pair probabilities (*69*). As with the partition function calculation, structures with the same set of base pairs but different coaxial stacking of helices are considered distinct structures.

To generate all suboptimal structures within a free-energy increment of the lowest free-energy structure, the dynamic programming recursions need to be nonredundant; i.e., each secondary structure must be considered once and only once. This is accomplished by calculating the lowest free energy possible for subsequences with defined structure and storing these results in arrays. Six arrays, V, W, WL, Wcoax, WMB, and WMBL, of size $O(N^2)$ are used, and the conformations are the same as those for the partition function calculation (*69*). The complete recursions are available as the Supporting Information. The program for generating all possible suboptimal structures will be available for download as part of RNAstructure at http://rna.urmc.rochester.edu.

Generating a Consensus Secondary Structure from Selected Reasonable Structures. A C++ program was written to calculate base-pair probability from the sum of base-pair frequency in each selected reasonable structure weighted by the predicted free energy. Base pairs with a probability higher than 65% were used to restrain the consensus secondary structure.

Prediction by OligoWalk of Free Energy of Binding to the RNA Target. The OligoWalk function in RNAstructure 4.1 was applied with options "break local structure" and RNA probes to calculate the binding free energy for every 9-mer probe complementary to R2 RNA. Because the probes cannot interact with each other after immobilization on the chip surface, the oligonucleotide concentration was set to 1×10^{-15} M, which makes intermolecular pairing of the probes used negligible.

RESULTS

Secondary Structure Prediction in the Absence of Experimental Restraints. Figure 1 shows the secondary structure of the 3′ UTR of *B. mori* R2 RNA as determined by sequence comparison and chemical mapping, along with free-energy minimization (*55*). The conserved helixes derived from the sequence comparison are numbered sequentially and contain a total of 46 base pairs.

RNAstructure 4.1 and Mfold 3.2 were used to generate a total of 1000 possible secondary structures for the R2 sequence. Parts a and b of Figure 2 show how accurately both programs predict the 46 conserved base pairs. Both programs use a heuristic to generate a representative set of

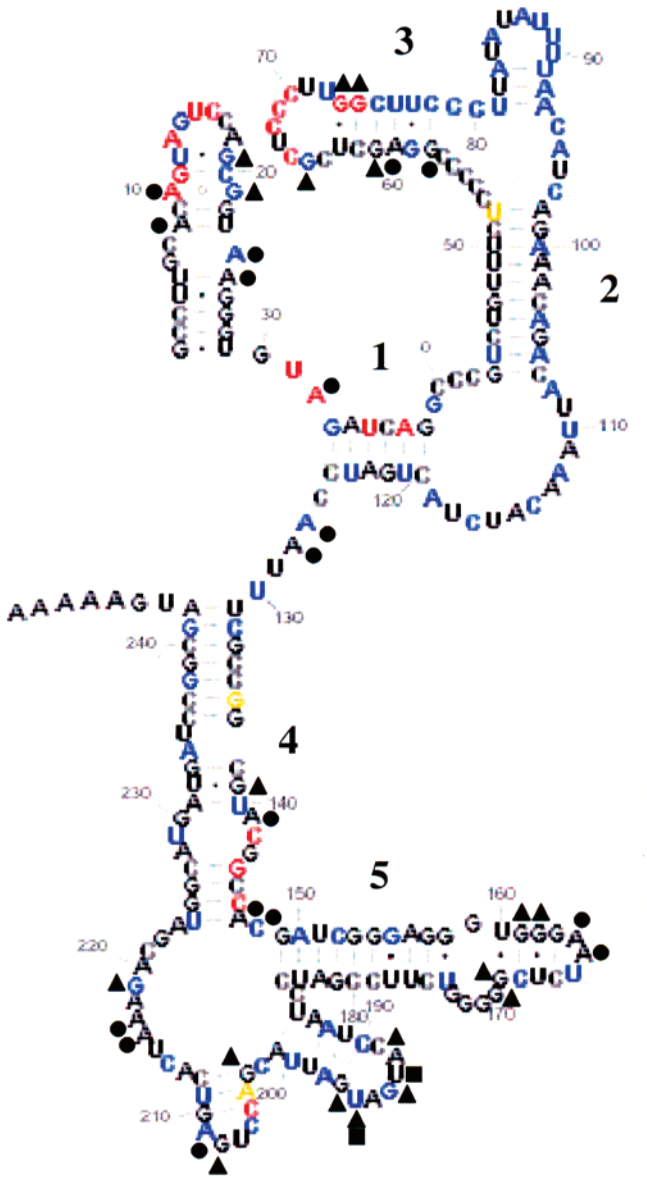

FIGURE 1: Secondary structure of *B. mori* R2 3′-UTR RNA (*55*). Conserved helixes are numbered. Note that helix 4 extends from nucleotides 131−147 and 224−242, but helix 5 only extends from nucleotides 149−158 and 175−184. Helixes without numbers are predicted by free-energy minimization. Nucleotides strongly modified by chemicals are indicated: kethoxal with ▲, CMCT with ■, and DMS with ● (*55*). Color annotation indicates the binding of R2 RNA to 9-mers on microarray, where the middle nucleotide of the nine nucleotides that are Watson−Crick complementary to the probe is colored to indicate binding intensity. Red, strong binding; orange, strong binding with a potential cross-hybridization site; blue, weak or no binding; black, not tested. The structure is drawn with software XRNA (http://rna.ucsc.edu/rnacenter/xrna/xrna.html).

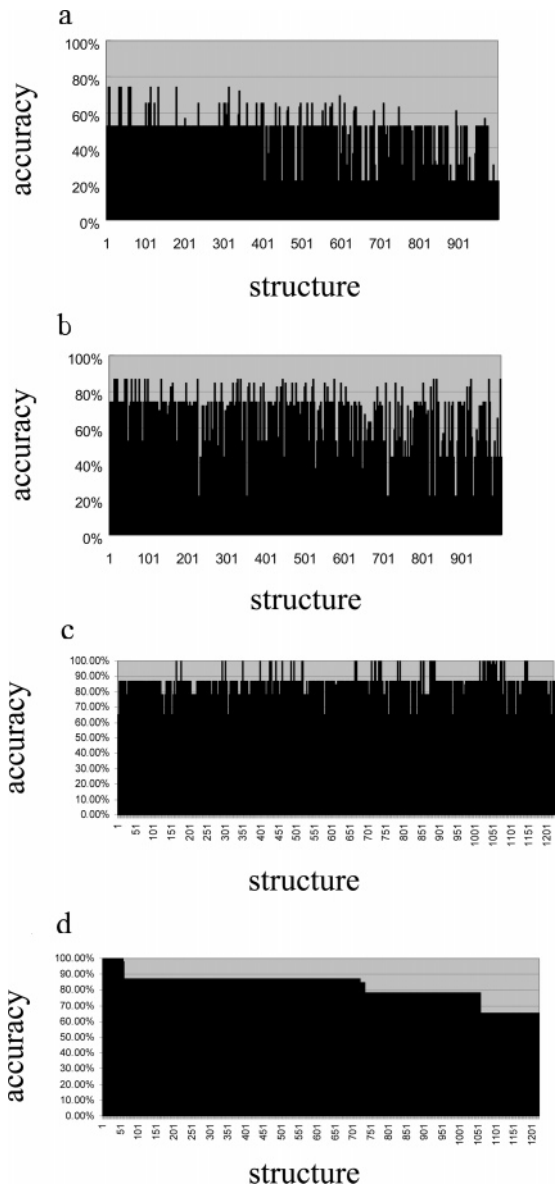

FIGURE 2: *B. mori* R2 RNA secondary-structure predictions compared to the 46 conserved base pairs determined by sequence comparison (*55*). A predicted pair is considered correct if it is identical or shifted by one nucleotide relative to the base pair in the structure determined by sequence comparison. (a) First 1000 structures predicted by Mfold with P = 99, W = 0, and MAX = 1000. The lowest free-energy structure (−70.5 kcal/mol) is 52% correct. The most correct structure, number 7 (−68.0 kcal/mol), is 74% correct. The 1000th structure (−49.5 kcal/mol) is 22% correct. (b) First 1000 structures predicted by RNAstructure 4.1 with max energy difference = 99%, window size = 0, and MAX structure number = 1000. The lowest free-energy structure (−71.1 kcal/mol) is 74% correct. The most correct structure, number 14 (−69.8 kcal/mol), is 87% correct. The 1000th structure (−62.1 kcal/mol) is 43% correct. (c) The 1221 reasonable structures selected from the 881 544 possible suboptimal structures generated by RNAstructure 4.1 within a 4 kcal/mol free-energy window by using oligonucleotide microarray restraints in a selection algorithm. The lowest free-energy structure (−68.1 kcal/mol) is 65% correct, although it is not obvious in the scale of the plot. The last structure (−67.1 kcal/mol) is 87% correct. (d) The 1221 reasonable structures sorted by 3-2-1 scoring of predicted binding to strong binding 9-mer probes. The first structure (−67.3 kcal/mol) is 100% correct. The last structure (−67.2 kcal/mol) is 65% correct. Some structures scoring similarly but with lower accuracy are also in the plot but not obvious in this scale.

suboptimal structures. This heuristic samples a small number of the suboptimal structures because the total number of structures increases exponentially as the free-energy difference increases (Figure 3).

Oligonucleotide Microarray Experiment. A 2′-O-methyl RNA/RNA duplex is more stable than a DNA/RNA duplex and similar in stability to an RNA/RNA duplex (*63*, *70*, *71*). Moreover, 2′-O-methyl RNA oligonucleotides are more chemically stable than RNA while still favoring A-form conformation (*72*). Thus, a variety of 7-, 9-, and 11-mer 2′-O-methyl RNA oligonucleotide probes complementary to the

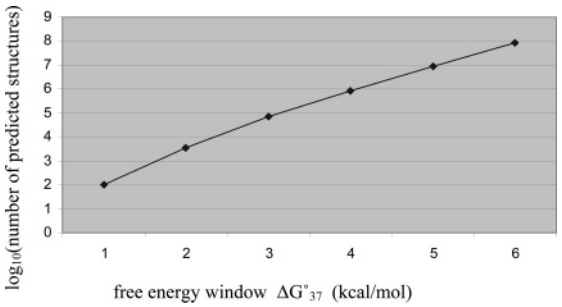

FIGURE 3: Log(number of predicted suboptimal structures) for *B. mori* R2 3′-UTR RNA versus the difference in free energy from the lowest free energy.

R2 3′ UTR were synthesized. Initial microarray experiments indicated that relatively few 7-mers bound R2 RNA but that 9-mers targeted to various regions did bind R2 RNA. Because the interpretation of binding to 9-mers is simpler than for 11-mers, binding of R2 RNA to microarrays of a series of 9-mers was measured. Most probes with more than five A and/or U nucleotides or with three or more consecutives Gs were avoided because of predicted weak binding to unstructured RNA (*71*) or difficulty in purification, respectively. A probe is numbered by the center nucleotide position of the complementary sequence in the RNA target. For example, the 9-mer probe 52 is Watson−Crick-complementary to nucleotides 48−56 in R2 RNA. The data are plotted as $\log_2(I/I_{min})$ in Figure 4 and summarized in Figure 1, where strong binding is defined as $\log_2(I/I_{min}) \geq 0.5 \log_2(I_{max}/I_{min})$.

Consideration of Cross Hybridization in Interpreting Microarray Data. Cross hybridization is binding by a probe to a site that is not completely Watson−Crick-complementary to the probe. The target RNA is folded; therefore, many potential cross-hybridization sites are intramolecularly base-paired and thus not available to bind probes. For every potentially strong cross-hybridization site, there is also a completely complementary probe targeting this site. If there is not strong binding to this probe, then cross hybridization to this region can usually be ignored (see below). Any site giving strong binding to its fully complementary probe, however, must be considered a potential site for cross hybridization.

To check for potential cross hybridization in the microarray experiment, a function was added to the RNAstructure 4.1 program to predict potential mismatches. This function predicts potential duplex-forming free energies for binding with mismatches by assuming that target RNA is totally unfolded. For this calculation, 2′-O-methyl RNA was assumed to behave the same as RNA and RNA/RNA thermodynamic parameters were used to approximate the free energy of the heteroduplex. This is consistent with thermodynamic results for RNA/2′-O-methyl RNA heteroduplexes (*71*). Table 1 lists the predicted free energies of complementary and mismatch binding for strong binding probes. A free energy more favorable than −10 kcal/mol is considered to indicate a possible mismatch binding site if the fully complementary probe binds the mismatch site strongly. For example, probe 68 (perfect match $\Delta G^\circ_{37} = -16.1$ kcal/mol) can mismatch site 77 (mismatch $\Delta G^\circ_{37} = -10.2$ kcal/mol), but probe 77 (perfect match $\Delta G^\circ_{37} = -16.3$ kcal/mol) (see Table S1 in the Supporting Information) has a much lower intensity than probe 68 (Figure 4). Thus, the strong binding of probe 68 is not due to binding to site 77. Probe 52 (perfect match $\Delta G^\circ_{37} = -15.0$ kcal/mol) can bind site 68 (mismatch $\Delta G^\circ_{37} = -10.0$ kcal/mol), and probe 68 binds strongly; therefore, cross hybridization of probe 52 to site 68 is possible. Such ambiguities can be resolved by additional experiments.

One experiment to check for cross hybridization of probe 52 is to add either free probe 52 or 68 to the hybridization solution to see if the binding signal changes because of free probe binding to the RNA. Free probe 52 or 68 at 10 μM prevents the binding of R2 RNA to both probes on the chip (data not shown). It is not clear, however, which probe binds to which sites.

Ribonuclease H cleavage and reverse transcription sequencing can locate probe-binding sites. The probe backbone is changed from 2′-O-methyl RNA to DNA because Ribonuclease H specifically cleaves RNA in a DNA/RNA heteroduplex. Figure 5 shows that probe 68 induces strong Ribonuclease H cuts at nucleotides 70−72, as expected. In contrast, probe 52 induces strong Ribonuclease H cuts at nucleotides 68−70 but not around site 52. Evidently, probe 52 binds at site 68 but not site 52. Thus, probe 52 is not used as a strong binding probe in evaluating predicted secondary structures.

Other strongly binding probes might also have cross hybridization (Table 1). For example, probe 204 can bind to site 142 (mismatch $\Delta G^\circ_{37} = -10.5$ kcal/mol) and binds less R2 RNA than probe 142 (Figure 4). Thus, the binding of probe 204 is not used as a restraint in structure evaluation. Probes with strong binding and potential cross hybridization that eliminates their use as restraints are colored orange in Figure 1.

Probes Binding to Both Sides of a Helix. If a probe has two binding regions in the target RNA and the two target sequences form a double helix, then two probes of identical sequence could open the helix by binding to both sides (Figure 6a). Usually such probes are almost or completely self-complementary and therefore form dimers in solution. This phenomenon is illustrated by comparing the binding of probes 136 and 238 to R2 RNA and to a 33 nucleotide mimic of helix 4 on a microarray and in solution (Table 2 and Figures 4 and 6−8). Both probes bind R2 RNA and the helix 4 mimic to a microarray (Figures 4 and 7c) even though they are predicted to not bind (Table 1 and Table S1 in the Supporting Information). In solution, fluorescein-labeled probe 136 does not bind either target but fluorescein-labeled probe 238 binds both (Table 2). The absence of solution binding by probe 136 is expected because the K_d for probe dimerization at 37 °C is predicted to be 4.6×10^{-13}, so that dimerization is favored over binding to the target. The predicted K_d for dimerization of probe 238, 5.2×10^{-6}, is considerably weaker, thus allowing for binding to the target. The binding of probe 238 to the helix 4 mimic in solution allows for a measurement of the stoichiometry of this association. As shown in Figure 8, the stoichiometry is two probes per helix 4 mimic. Because dimerization is not possible for probes immobilized on the microarray but 2:1 stoichiometry is possible, the binding of both RNA targets to probes 136 and 238 on the microarrays can be attributed to adjacent probes on the microarray binding each side of helix 4 (Figure 6a). Interestingly, however, probe 238 on a

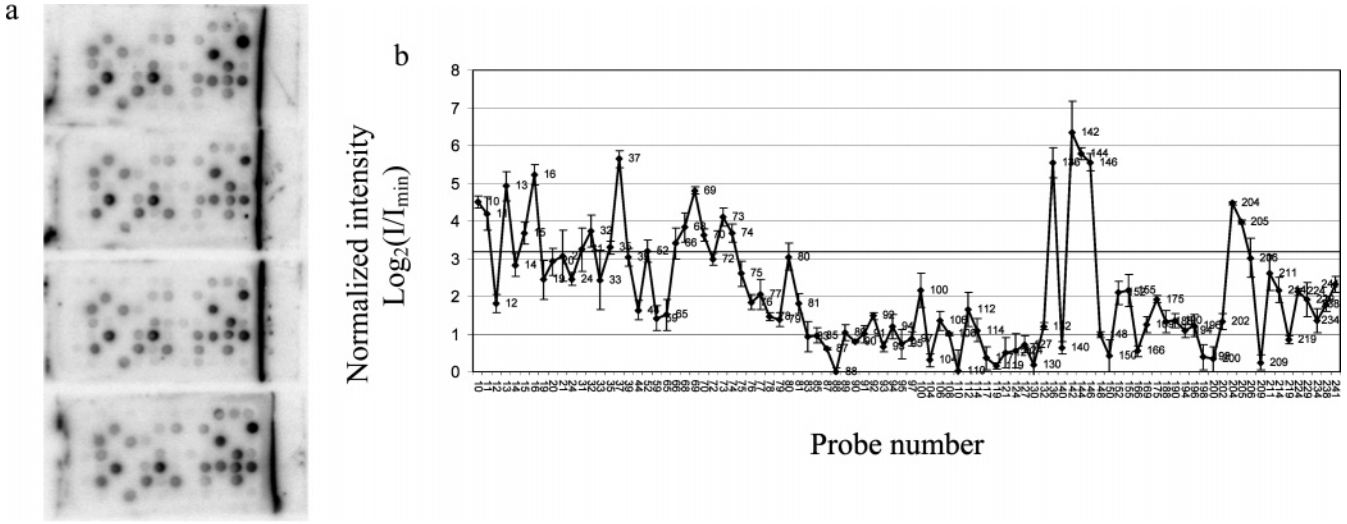

FIGURE 4: Microarray results for 2′-O-methyl RNA 9-mer probes hybridized to ^{32}P-labeled R2 RNA. (a) Images of four slides. (b) Average of the results for four slides. The purple line is drawn halfway between the minimum and maximum $\log_2(I/I_{min})$, where I is the intensity of the radioactivity. The probe numbers on the plot are to the right of the point representing that probe.

Table 1: Predicted Oligonucleotide Probe-Binding Free Energy When No Structure or Local Structure Is Considered for All Probes Strongly Binding to R2 RNA[a]

			no target structure[b]	after breaking local target structure[c]	no target structure	
					cross hybridization	
oligo sequence	used as restraints (R/N)	center of binding site	$\Delta G°_{37}$ (kcal/mol)	$\Delta G°_{37}$ (kcal/mol)	site	$\Delta G°_{37}$ (kcal/mol)
CUACUGUGC	R	10	−14.0	−13.0	-	-
ACUACUGUG	R	11	−12.0	−11.8	-	-
GGACUACUG	R	13	−15.8	−13.2	-	-
CUGGACUAC	R	15	−14.2	−13.2	-	-
GCUGGACUA	R	16	−16.2	−14.4	-	-
AUCUACACC	R	31	−12.2	−4.0	-	-
GAUCUACAC	R	32	−13.4	−3.0	-	-
CCUGAUCUA	R	35	−12.8	−8.2	152X	−11.1
GGCCUGAUC	R	37	−17.0	−11.5	152X	−10.2
GGGGAGAAA	N	52	−15.0	−6.4	68H	−10.0
GGGAGCGAG	R	66	−18.7	−12.5	75/76X	−10.2
AAGGGAGCG	R	68H	−16.1	−14.4	76/77X	−10.2
CAAGGGAGC	R	69	−17.1	−14.1	-	-
CCAAGGGAG	R	70	−16.0	−10.6	-	-
AAGCCAAGG	R	73	−13.9	−4.3	-	-
GAAGCCAAG	R	74	−14.2	−8.2	-	-
ACGCCGGCG	N	136	−19.0	2.0	238B	−12.3
GGCCGUACG	R	142	−18.3	−3.5	-	-
GUGGCCGUA	R	144H	−17.5	−6.0	-	-
UCGUGGCCG	R	146H	−18.0	−3.2	241X	−12.2
					140X	−10.7
					39X	−10.2
CAGGUCGUA	N	204	−14.7	−11.5	142	−10.5
UCAGGUCGU	R	205	−15.0	−13.7	-	-

[a] Each 2′-O-methyl RNA 9-mer is identified by the number corresponding to the number of the R2 RNA nucleotide complementary to the middle nucleotide of the 9-mer probe (center of the binding side). Alternative binding sites are defined by the middle nucleotide of the binding region on R2 RNA. R, restraint; N, not used as restraint; H, Ribonuclease H cleavage at this site; X, weak binding for the fully complementary probe; B, binding to both sides of a likely helix in R2; dash, no potentially strong cross-hybridization site. [b] Calculated with the bimolecular-folding option of RNAstructure 4.1 (*42*). [c] The structure in Figure 9a was used for this calculation by OligoWalk (*74*).

microarray binds R2 RNA relatively weakly (Figure 4). This may reflect the relative competition of many probes on the microarray for a limited concentration of R2 RNA or contribution of the fluorophore to binding in solution.

The results imply that immobilized probes can bind an RNA by binding to both sides of a helix but that this may not be evident from solution experiments because of dimerization such as observed with probe 136. This is a reason to not use restraints from some probes that have two potentially strong binding sites even if the cross-hybridization site does not exhibit binding to its Watson−Crick complementary probe. For the probe to be rejected, however, the site of potential cross hybridization must have the potential to form a helix with the site perfectly matched to the probe. Dot plots of energetically feasible base pairs provide a way to search for potential helixes (*73*). For example, from free-energy minimization alone, sites 136 and 238 are predicted to form helix 4 with base-pair probability more than 95% in a

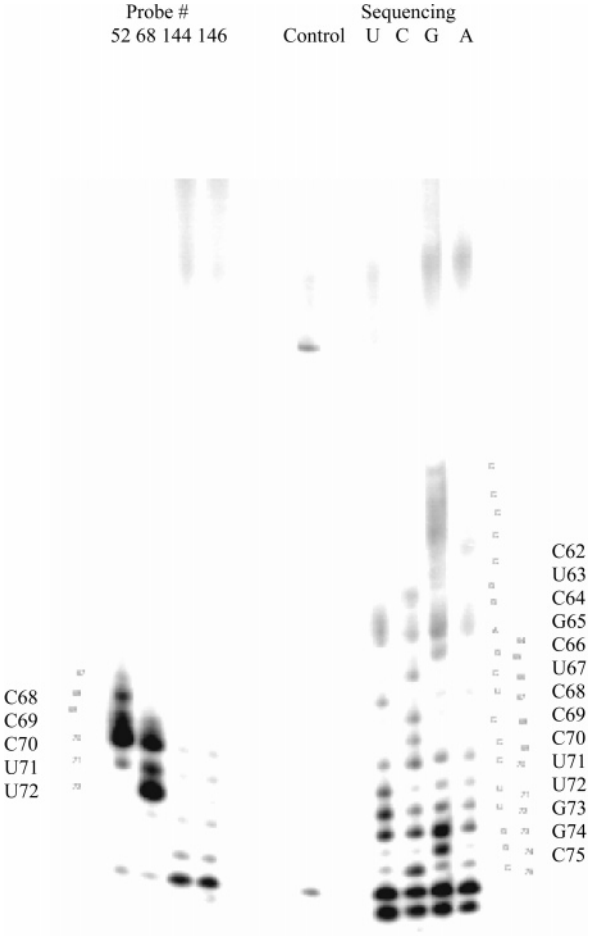

FIGURE 5: Ribonuclease H cleavage of R2 3′-UTR RNA. The cleavage positions indicate where the DNA probes bind to the target RNA. There are two control lanes with RNA but without Ribonuclease H or without DNA probes. The band at U77 for probes 144 and 146 can also be seen in the control lane without DNA probes and in all sequencing lanes. It is likely a stop in reverse transcription and not because of the cleavage by Ribonuclease H.

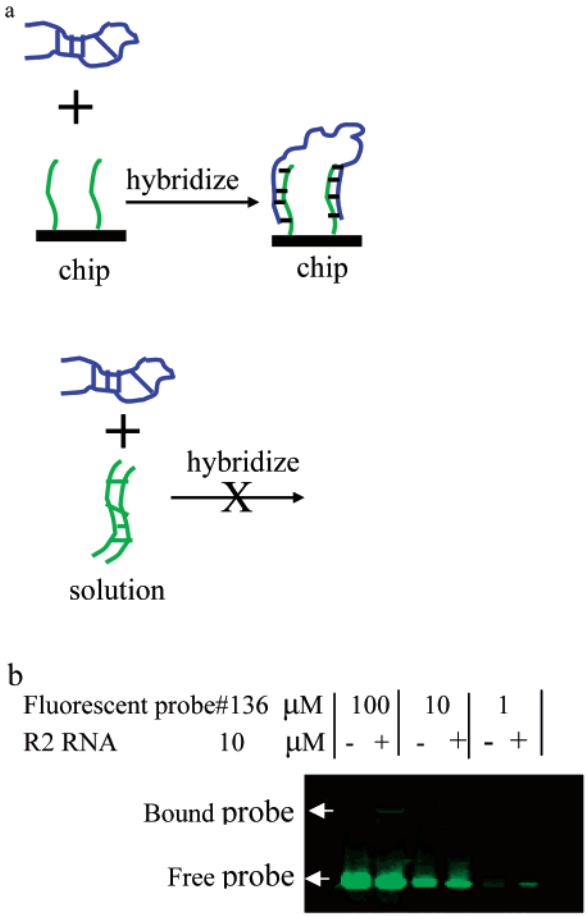

FIGURE 6: (a) Cartoon of target RNA hybridizing to the probe on the chip because of the binding of the probe to both sides of a helix. The probe may not hybridize to the RNA in solution if the probe dimerizes strongly. (b) Gel-shift experiment with fluorescent 2′-O-methyl RNA probe 136 and R2 RNA. When the probe 136 concentration is 100 μM, the (bound probe)/(total probe) ratio is 0.6%.

partition function dot plot (see the Supporting Information) (*69*). Thus, before applying final restraints in structure selection, it is possible to check whether a strongly binding probe with cross-hybridization sites might bind to both sides of a helix in the secondary structure.

Binding to both sides of a helix is likely to be important mostly at helixes that close a loop only on one end, e.g., helixes 1 and 4 in Figure 1. Opening such a helix will typically release an unfavorable loop. Binding to both sides of a helix that closes a loop at both ends, however, will result in a larger loop, which is usually unfavorable. This suggests that probes binding to both sides of a helix will be found primarily at the 5′ and 3′ ends of a structured RNA, e.g., helix 4.

Thermodynamics of 2′-O-Methyl RNA/RNA Duplexes in 110 mM NaCl and 10 mM $MgCl_2$ at pH 7.4. Interpretation of RNA binding to microarrays requires knowledge of the thermodynamics of binding. For example, relative binding of RNA to oligonucleotides on microarrays can be compared to the binding predicted by programs such as OligoWalk (*74*). OligoWalk uses thermodynamic parameters for RNA/RNA folding at 1 M NaCl to predict the free energy required to locally unfold an RNA target and bind a Watson−Crick complementary RNA oligonucleotide at 37 °C. Thermodynamic studies of Watson−Crick complementary 2′-O-methyl RNA/RNA duplexes in 0.1 M NaCl indicate that the sequence dependence is similar to that for RNA/RNA duplexes in 1 M NaCl (*71*). To test whether predictions with RNA/RNA parameters in 1 M NaCl approximate the

Table 2: Binding of RNA Targets to Probes 136 and 238 on the Microarray and in Solution

	probe			
	136		238	
	5′-ACGCCGGCG-3′		5′-UCGCCGGAU-3′	
target	microarray	solution	microarray	solution
R2	++[a]	−[b]	+[a]	++[c]
helix 4 mimic	++[d]	−[e]	++[d]	++[f]
dimer ΔG°_{37} (kcal/mol)	−17.5[g]		−7.5[g]	
K_d (dimer)	4.6×10^{-13}		5.2×10^{-6}	

[a] See Figure 4. [b] Less than 1% of the fluorescein-labeled probe is bound in a gel-shift experiment with 10 μM R2 RNA and 1, 10, or 100 μM probe 136 (Figure 6b). [c] In solution, 100 μM fluorescein-labeled probe 238 binds 20 μM R2 RNA strongly (data not shown). [d] See Figure 7c. [e] No binding of 25 μM fluorescein-labeled probe 136 to 100 μM helix 4 mimic in a gel-shift experiment (Figure 7b). [f] See Figures 7b and 8. [g] The dimers in solution are 5′-ACGCCGGCG-3′/3′-GCGGCCGCA-5′ and 5′-UCGCCGGAU-3′/3′-UAGGCCGCU-5′ for probes 136 and 238, respectively, where underlines indicate unpaired nucleotides.

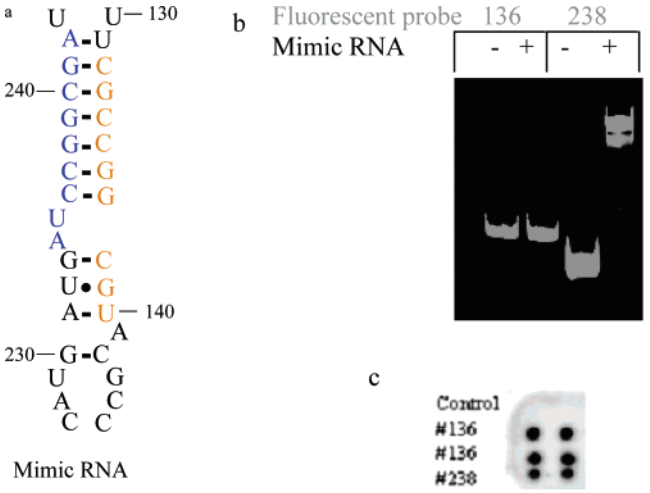

FIGURE 7: (a) Small RNA mimic of the secondary structure of part of helix 4. Bases in orange are the target site for probe 136; bases in blue are the target site for probe 238. (b) Native gel electrophoresis of the hybridization of 25 μM fluorescent 2′-O-methyl RNA probe with 100 μM small RNA mimic. The doublet observed with probe 238 may reflect two different conformations for the bound complex. (c) Microarray experiment with 0.1 μM small RNA mimic, where the control probe is 2′-O-methyl $(A)_9$. Both columns are identical.

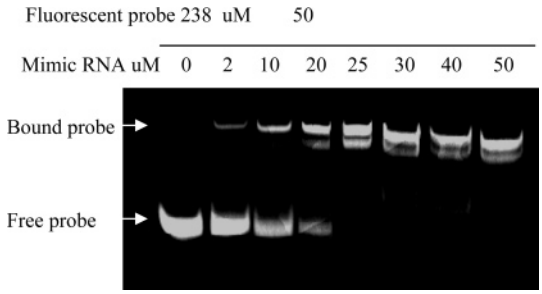

FIGURE 8: Concentration dependence of small RNA mimic binding to 50 μM fluorescent probe 238 indicates a binding stoichiometry of 1:2. The doublet observed for the bound complex may reflect two different conformations for the bound complex.

sequence dependence of 2′-O-methyl RNA/RNA duplexes in hybridization buffer (0.11 M NaCl and 10 mM $MgCl_2$), the thermodynamic parameters for several 2′-O-methyl RNA/RNA duplexes were measured by optical melting (Table 3). The duplexes measured corresponded to probe 114 and shortened versions of probes 146, 204, and 211 to keep melting temperatures below 60 °C. As shown in Table 3, the measured free-energy changes for complementary duplexes at 37 °C are on average 1.20 ± 0.53 kcal/mol more stable than predicted by parameters for RNA/RNA duplexes in 1 M NaCl. Sequence-independent effects because of 2′-O-methyl substitution and different salt are expected to go as the number of substitutions and the number of phosphate charges, respectively. When the average enhancement in stability in analyzed in these ways, it is 0.18 ± 0.09 kcal/mol per 2′-O-methyl substitution and 0.21 ± 0.12 kcal/mol per phosphate pair. Evidently, the 2′-O-methyl substitutions and change in salt concentrations do not have a major effect on the sequence dependence of Watson−Crick base-pair formation.

Probe 204 can cross-hybridize at site 142 by forming the duplex 5′-*CAGGUCGUA*-3′/3′-*ACCGGCA*-5′, which includes a GU pair and a terminal AA mismatch. To mimic this duplex, the shortened version of probe 204 was used to form the duplex, 5′-*AGGUCGU*-3′/3′-*ACCGGCA*-5′ and the thermodynamics were measured (Table 3). The mismatched duplex was 1.3 kcal/mol less stable than the perfectly matched duplex, whereas RNA/RNA parameters predict only a 0.6 kcal/mol loss in stability. It is possible that reduced flexibility because of 2′-O-methyl substitutions may increase the penalty for non-Watson−Crick pairings.

Algorithms for Using Restraints from Microarray Experiments. The secondary structure of the 3′ end of *B. mori* R2 RNA is known from sequence comparison (*55*). This provides a benchmark for testing algorithms to deduce secondary structure from microarray data. In Figure 4, the probes above the purple line are defined as strongly binding probes, where strong binding means that $\log_2(I/I_{min}) \geq 0.5 \log_2(I_{max}/I_{min})$. Thus, probes 10, 11, 13, 15, 16, 31, 32, 35, 37, 52, 66, 68, 69, 70, 73, 74, 136, 142, 144, 146, 204, and 205 are strongly binding probes. Considering cross hybridization and potential binding to both sides of a helix, probes 52, 136, and 204 are not used as restraints. This is summarized in Figure 1 where target sites for strongly binding probes are colored red or orange depending upon whether they were or were not used as restraints, respectively.

A study of *E. coli* 5S rRNA binding to a microarray of 2′-O-methyl RNA 7-mers suggested that the binding can be used to restrain predicted secondary structures by treating the middle nucleotide of a strong binding site as equivalent to a site of chemical modification (*50*). That is, the middle nucleotide cannot be in a Watson−Crick pair flanked by Watson−Crick pairs. If the middle nucleotide of strongly binding 9-mer probes for R2 RNA (excluding 52, 136, and 204) is input as a chemical-mapping restraint in RNAstructure 4.1, then the lowest free-energy structure has 35% of the 46 known base pairs. The poor prediction is due to the middle nucleotides of strongly binding probes 35 and 146 being in the middle of helixes. Treating the middle nucleotide of strongly binding probes as equivalent to a chemical modification restraint may only be useful for interpreting the binding of probes shorter than 9-mers.

In principle, the predicted ΔG°_{37} for binding of possible R2 RNA structures to each probe could be compared with the observed binding. This calculation takes too long for all possible 881 554 structures (about 7.7 GB in size) within a free-energy range of even 4 kcal/mol (Figure 3). Therefore, a simple initial screen was developed. Because RNAstructure generates each possible secondary structure, 10 sites that strongly bind a 9-mer are checked to see if there are at least three unpaired nucleotides with at least two adjacent. The structure is immediately rejected if this criterion is not met. The 10 sites were 10, 15, 32, 37, 66, 69, 74, 142, 146, and 205, but alternate choices gave similar results. Only 7184 structures (63.3 MB in size) were retained at this stage. Each secondary structure satisfying the criterion of three unpaired nucleotides with at least two adjacent at a strong binding site is later assigned a score by a JAVA program according to the following 3-2-1 scoring algorithm: each unpaired G or C in the predicted secondary structure of the target scores 3 points, each unpaired A or U scores 2 points, and each G in a G•U pair scores 1 point. All strongly binding probes used as restraints have scores higher than 8 for binding to the known secondary structure; therefore, 8 was chosen as the cut off indicating a reasonable structure. Only 1221 structures were left after this filter. When these structures

Table 3: Experimental Thermodynamic Parameters of Heteroduplex Formation between 2′-O-Me RNA and RNA Oligonucleotides in 110 mM NaCl and 10 mM $MgCl_2$

heteroduplex 5′-(RNA)^M-3′ 3′-RNA-5′	average of curve fits				T_M^{-1} versus log(C_T/4) plots			
	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_M[a] (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_M[a] (°C)
5′-(AGAUGUUUA)^M-3′ 3′-UCUACAAAU-5′	−89.21 ± 6.33[b]	−257.71 ± 19.63[b]	−9.28 ± 0.29[b]	46.9[b]	−72.68 ± 3.28[b]	−205.84 ± 10.30[b]	−8.84 ± 0.09[b]	47.2[b]
					(−64.42 ± 5.71)[c]	(−181.7 ± 17.7)[c]	(−8.05 ± 0.29)[c]	(46.9)[c]
					[−68.65 ± 15.23][d]	[−195.7 ± 46.8][d]	[−7.98 ± 0.70][d]	[45.7][d]
5′-(CGUGGC)^M-3′ 3′-GCACCG-5′	−77.68 ± 11.84[b]	−213.58 ± 35.35[b]	−11.44 ± 0.89[b]	57.9[b]	−70.33 ± 5.06[b]	−191.47 ± 15.33[b]	−10.94 ± 0.33[b]	57.8[b]
					(−57.14 ± 5.18)[c]	(−154.2 ± 16.0)[c]	(−9.30 ± 0.28)[c]	(55.6)[c]
					[−61.88 ± 14.44][d]	[−170.6 ± 44.4][d]	[−8.98 ± 0.66][d]	[52.2][d]
5′-(AGGUCGU)^M-3′ 3′-UCCAGCA-5′	−64.92 ± 6.91[b]	−176.39 ± 21.03[b]	−10.21 ± 0.41[b]	55.6[b]	−66.36 ± 7.08[b]	−180.93 ± 21.60[b]	−10.24 ± 0.42[b]	55.4[b]
					(−58.70 ± 5.35)[c]	(−158.5 ± 16.5)[c]	(−9.54 ± 0.28)[c]	(56.4)[c]
					[−62.48 ± 14.93][d]	[−171.6 ± 45.9][d]	[−9.30 ± 0.69][d]	[53.6][d]
5′-(AGGUCGU)^M-3′[e] 3′-ACCGGCA-5′	−46.56 ± 8.77[b]	−122.18 ± ±27.30[b]	−8.67 ± 0.40[b]	51.9[b]	−54.37 ± 5.84[b]	−146.42 ± 18.03[b]	−8.96 ± 0.31[b]	51.5[b]
					(−58.12)[c]	(−158.8)[c]	(−8.9)[c]	(50.1)[c]
5′-(GUGACUC)^M-3′ 3′-CACUGAG-5′	−75.69 ± 6.55[b]	−210.39 ± 20.67[b]	−10.44 ± 0.25[b]	53.9[b]	−87.61 ± 9.16[b]	−247.1 ± 28.2[b]	−10.96 ± 0.46[b]	53.5[b]
					(−64.99 ± 5.10)[c]	(−179.5 ± 15.8)[c]	(−9.28 ± 0.26)[c]	(53.3)[c]
					[−68.89 ± 14.79][d]	[−175.2 ± 45.4][d]	[−8.52 ± 0.68][d]	[49.7][d]

[a] Calculated for 10^{-4} M oligonucleotide concentration. [b] Measured in 110 mM NaCl, 10 mM $MgCl_2$, 40 mM HEPES, and 40 mM sodium HEPES at pH 7.4. The 2′-O-methyl RNA sequences are full-length probe 114 and shortened versions of probes 146, 204, and 211. [c] Values predicted for RNA/RNA duplexes in solution of 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA at pH 7. [d] Values predicted for 2′-O-methyl RNA/RNA duplexes in solution of 0.1 M NaCl at pH 7 using parameters from ref *71*. [e] This duplex mimics probe 204 binding to site 142. Underlined nucleotides indicate non-Watson−Crick pairs.

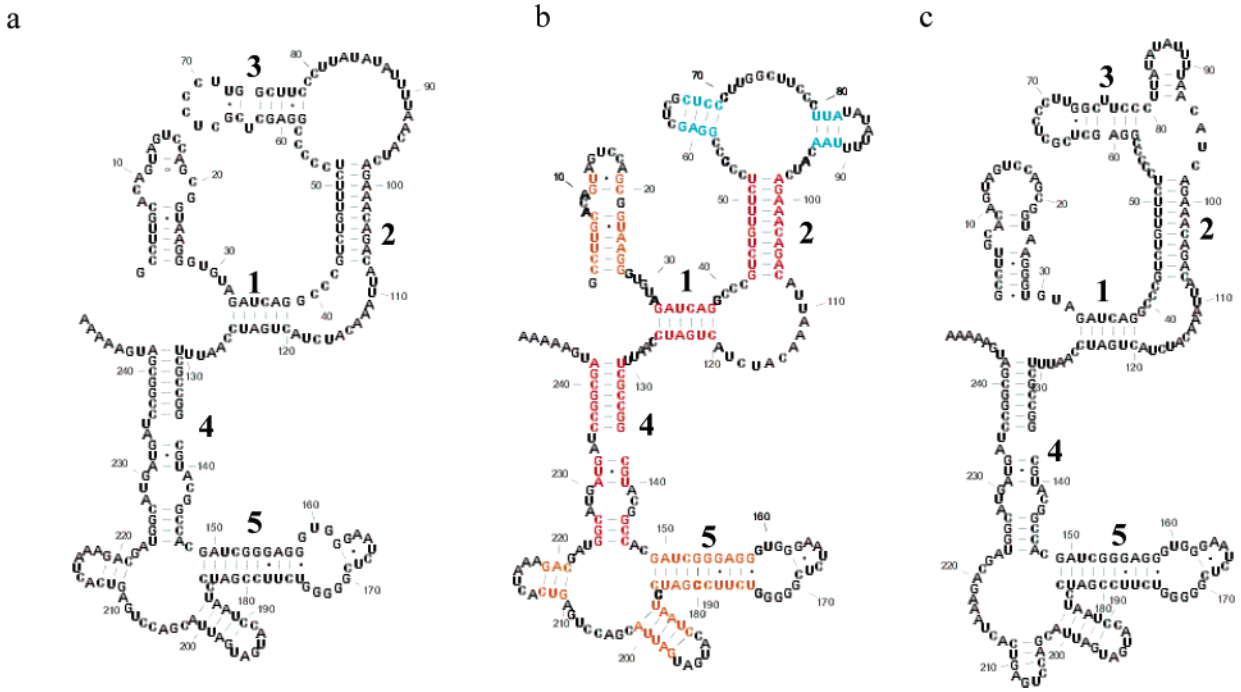

FIGURE 9: R2 RNA secondary structures deduced by different methods. Structures were drawn with XRNA (http://rna.ucsc.edu/rnacenter/xrna/xrna.html). (a) Secondary structure with highest score among 1221 structures by applying 3-2-1 scoring algorithm. It is 100% correct as compared to the 46 conserved base pairs from the sequence comparison (Figure 1) (*55*). Note that C64 of helix 3 is slipped in comparison to helix 3 in Figure 1 but is considered correct because it is slipped by only one nucleotide (*45*). (b) Consensus structure and base-pair frequency calculated from the 1221 structures consistent with microarray data. Each structure was weighted by its free energy of folding (*69*). Red indicates the base-pair frequency higher than 95%; orange indicates the base-pair frequency higher than 85%; and cyan indicates the base-pair frequency higher than 65%. The consensus structure has 87% of the 46 base pairs determined by sequence comparison. (c) Lowest free-energy structure applying chemical modification restraints (*55*). It is 100% correct in comparison to the 46 conserved base pairs from the sequence comparison. Note that helix 3 is considered the same as helix 3 in Figure 1 even though three nucleotides are slipped by one position.

were sorted by total scores for the potential to bind to 9-mers (Figure 2d), the structure with the highest score (Figure 9a) had 100% of the 46 base pairs known from sequence comparison (*55*). There is a slight difference in helix 3 in the predicted lowest free-energy structure and the comparative sequence analysis structure (Figures 1 and 9a). U63 is bulged in Figure 9a but not in Figure 1. When the two possible structures of helix 3 are compared, they are considered identical because base pairs are allowed to be slipped by up to one nucleotide on one strand and still be considered correct (*45*).

When the 1221 structures were sorted by free energy, the lowest free energy structure had 65% of the 46 base pairs known from sequence comparison, although most of the

structures at low free energies had 87% of the known base pairs (Figure 2c). This suggests that a partition function approach (*69*) may also be useful for interpreting microarray data. Partition functions allow for the generation of a consensus secondary structure on the basis of maximum base-pair probability. The base-pair probability is calculated by base-pair frequency in the reasonable structures weighted by the predicted free energy of each structure. Base pairs with a probability higher than 65% are used to generate a consensus secondary structure. The consensus secondary structure derived from the 1221 reasonable structures has 87% of the known base pairs (Figure 9b). For this case, the consensus structure is not as accurate as the structure with the highest score for probe binding, but this may not be general. For example, if there were no free-energy criterion, then the completely unfolded structure would have the highest score for potential binding of probes. It is possible, however, that limiting the free-energy window for structures considered may result in allowing the best structure to be selected by a comparison of measured and predicted binding to oligonucleotides. An alternative approach would be to compare oligonucleotide binding to centroid structures provided by the method of Ding et al. (*75*) to represent the relatively few classes of structures generated.

Constraints from Chemical-Mapping Data Gave Reasonable Structures for the 3′ End of B. mori R2 RNA. Chemical modification restraints improve RNA secondary-structure prediction accuracy, and this feature is incorporated in RNAstructure 4.1 (*42*). Sites of chemical modification are not allowed to be in Watson−Crick base pairs flanked by Watson−Crick base pairs. When bases strongly modified by chemical agents (*55*) are applied as chemical-mapping restraints for R2 RNA in RNAstructure 4.1, the lowest free energy structure is 100% correct (Figure 9c). This again demonstrates that free-energy minimization with the aid of chemical mapping can improve structure prediction accuracy. Note that in the predicted structure, a single base bulge of U76 (*76*) allows for the strong chemical modification at base G61, but this modification is in violation with helix 3 of the comparative analysis structure (Figures 1 and 9c) (*55*). Helix 3 is nevertheless considered correct because base pairs are slipped by only a single nucleotide (*45*). Interestingly, if bases that are moderately modified by chemical agents are added as additional constraints, then the predicted lowest free-energy structure is 87% correct. This may reflect the presence of alternate foldings in solution.

Application to E. coli 5S rRNA. To further test the utility of the methods for interpreting microarray data, the published microarray results from *E. coli* 5S rRNA (*50*) were also analyzed. For *E. coli* 5S rRNA, the lowest free-energy structure has a predicted ΔG°_{37} of −53.0 kcal/mol but only contains 27% of known base pairs (Figure 10). In a microarray experiment, the 5S rRNA binds to the 7-mer probes 27, 28, 29, 35, 36, 40, 41, and 44 without cross hybridization (*50*). When at least two adjacent unpaired bases are chosen as a constraint for allowing strong binding of 7-mer probes, only four structures are selected from 1.3×10^6 structures, generated within 6 kcal/mol. Figure 10d shows the accuracy of the four structures. The four structures remain in the same order whether ranked by free energy or total scores for expected binding to probes. The structure with the best score has a ΔG°_{37} of −47.1 kcal/mol and contains

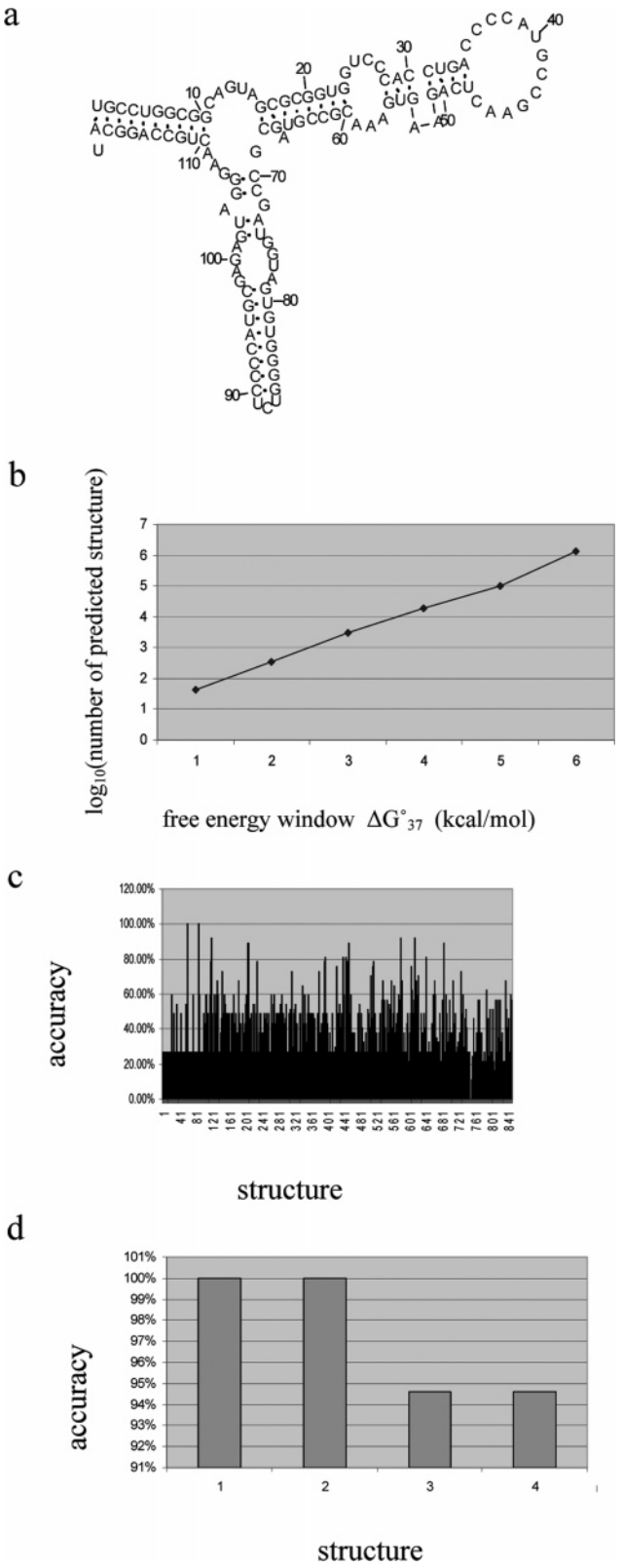

FIGURE 10: *E. coli* 5S rRNA secondary-structure prediction. (a) Secondary structure from the sequence comparison (*94*) and crystal structure (*34*). (b) Number of suboptimal structures generated versus the free-energy difference. (c) The 855 structures predicted by RNAstructure 4.1 with max energy difference = 99%, window size = 0, and MAX structure number = 1000. The lowest free-energy structure (−53.0 kcal/mol) is 27% correct. The most correct structure, number 61 (−47.1 kcal/mol), is 100% correct. (d) Only reasonable structures selected by a comparison to the binding observed on a microarray. The order is the same, whether ranked by the total score or by free energy.

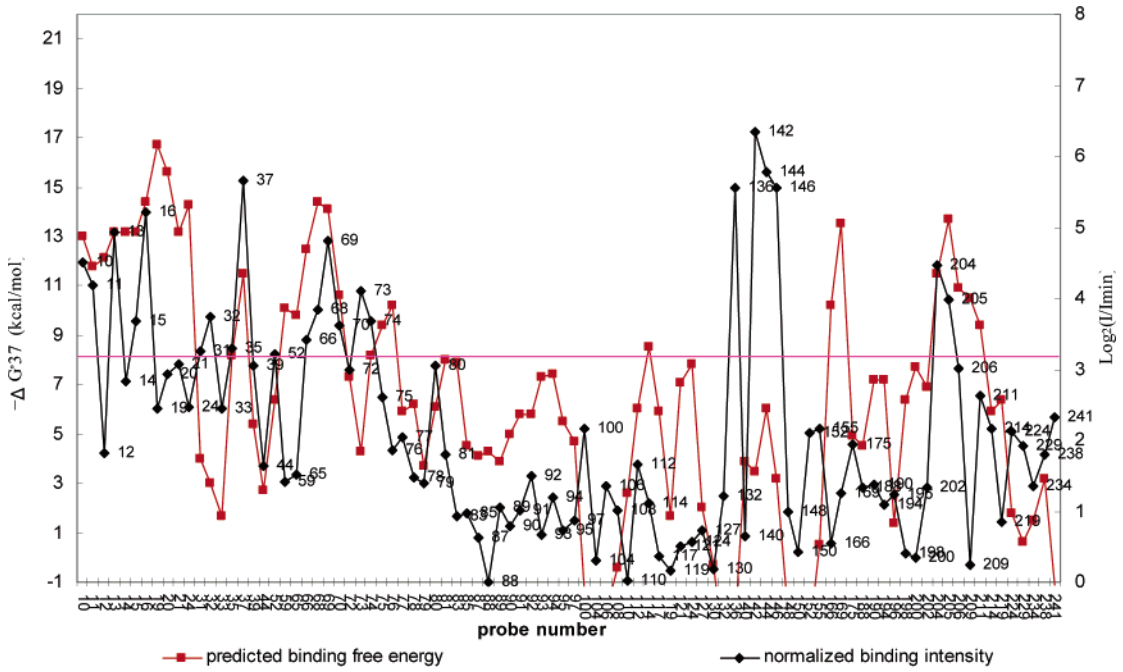

FIGURE 11: Correlation of predicted binding free energy at 37 °C (red) with microarray probe-binding intensity at 4 °C (black). The secondary structure used for OligoWalk was that scoring highest for the expected binding to 9-mers (Figure 9a) and contains all 46 base pairs derived from the sequence comparison (Figure 1). The purple line indicates the criteria for strong binding for both intensity and predicted binding free energy. Predicted binding was calculated at 37 °C because free-energy parameters for loops are available at 37 °C but not at 4 °C. The probe numbers on the plot are to the right of the experimental point representing that probe.

100% of known base pairs. The prediction on the basis of energetics alone is poor because an unusually stable loop E forms in the presence of Mg^{2+} (*77*), and this effect is not included in the current set of thermodynamic parameters (*42*). Evidently, oligonucleotide binding data can compensate for unknown factors in the energetics.

DISCUSSION

Free-energy minimization alone usually cannot predict a completely accurate secondary structure for RNA. Thus, experimental data are required to guide free-energy minimization. This is analogous to the use of NMR data to compensate for the limitations of force fields in generating three-dimensional structures (*28*, *32*, *78*). In this paper, we demonstrate that hybridization to microarrays of 9-mer 2′-O-methyl RNA oligonucleotides can provide experimental restraints useful for predicting accurate secondary structures. The secondary structure of the conserved helixes of *B. mori* R2 RNA could be deduced with 100 and 87% accuracy, respectively, by two different methods using microarray data. In contrast, structures predicted by free-energy minimization alone by Mfold and RNAstructure contain 52 and 72% of the known base pairs, respectively. The microarray method should be applicable to other RNA molecules as a fast tool to determine secondary structures.

Southern and co-workers first studied hybridization of RNA to immobilized oligonucleotides (*79*−*81*). They applied DNA dimers to dodecamers to bind tRNA on glass slides. Short oligonucleotides have several advantages for probing the RNA structure, including the absence of probe self-structure and greater specificity (*50*, *53*, *54*, *82*, *83*). Nevertheless, there are still ambiguities in interpreting binding data to 9-mers. The best ways for interpreting the data are just beginning to be investigated. This work reveals some important considerations and reports some new approaches for interpretation. Of course, much remains to be discovered about factors affecting the binding of RNA to oligonucleotides; therefore, approaches for interpreting the data are likely to become more sophisticated.

Choosing Restraints. For an RNA sequence with unknown secondary structure, it is important to have rules for choosing safe restraints. Here, strongly binding probes 10, 11, 13, 15, 16, 31, 32, 35, 37, 66, 68, 69, 70, 73, 74, 142, 144, 146, and 205 were used as restraints in selecting reasonable structures from all suboptimal structures generated (see red nucleotides in Figure 1). Strongly binding probes 52 and 204, however, were not chosen because of potential cross hybridization. Probes 52 and 204 have binding sites in the final structures that fit the selection criteria for accepting a structure. Probe 52 binds at site 68 as shown by Ribonuclease H cleavage, and probe 204 presumably binds at site 204 and possibly also at site 142. Probe 136, however, does not have a site fitting the selection criteria. It appears that two 136 probes on the microarray can bind to the two sides of helix 4, thus capturing the R2 RNA. For this to be a possibility, a probe must have at least one potentially strong cross-hybridization site that can form a helix with the completely matched site. Dot plots revealing potential helixes in the RNA target can be used to identify these exceptions. Such probes should not be used in the initial selection of restraints.

Correlation of Predicted Total Binding Free Energy and Microarray Probe Binding Intensity. The duplex-forming free energy of a given probe to the RNA target can be estimated with OligoWalk (*74*). If the target RNA concentration is in an appropriate range and equilibrium is reached, then the binding signal intensity and predicted duplex-forming free energy for each probe would correlate well with perfectly predicted $\Delta G°$ when target and probe structures are consid-

ered. The target RNA is folded; therefore, the calculation must include the free-energy cost of breaking local structure to allow for the binding to the probe. Such a correlation has been seen for DNA microarrays, although self-folding of target and probes was neglected (*84*). Figure 11 shows the correlation of predicted total binding free energy with microarray binding intensity. A total of 13 of 19 probes used as restraints are predicted to bind with free energy more favorable than −8 kcal/mol, but probes 31, 32, 73, 142, 144, and 146 are exceptions. The favorable binding of probes 31, 32, 73, 142, 144, and 146 may be due to coaxial stacking onto a helix pre-existing in the R2 RNA. Coaxial stacking by oligonucleotides is not included in the OligoWalk algorithm but is known to enhance binding (*66*−*68*, *85*). Moreover, the bimolecular part of the calculation is based on the thermodynamics of short oligonucleotides binding together (*63*). The entropy loss should not be as unfavorable for binding to R2 nucleotides in loops. The comparisons between measured and predicted binding indicate that more sophisticated algorithms will be required to definitively interpret binding to microarrays.

Several regions are predicted to strongly bind probes but do not (Figure 11). These include nucleotides 59, 65, 114, 166, 169, and 209. There are many possibilities for probes not binding as well as predicted. For example, the prediction is only based on the secondary structure and does not include elements of structure such as pseudoknots and tertiary interactions. Thus, inspection of regions that weakly bind oligonucleotides while predicted to bind strongly may provide insight into three-dimensional folding.

A comparison between binding to microarrays and chemical modification data may also provide insight into the three-dimensional structure. For example, the region including nucleotides 66−70 strongly binds to 9-mers but is not reactive to chemicals. This could result from the formation of a pseudoknot having only a few Watson−Crick base pairs, which can be opened by binding to a 9-mer. Three-dimensional modeling of RNA on the basis of known secondary structure and chemical modification has been remarkably successful in several cases (*86*−*88*). Additional insights from microarray data have potential for further refining these methods.

CONCLUSIONS

Several approaches were tested for interpreting the binding of the 249 nucleotide 3′ UTR of *B. mori* R2 RNA to a 2′-O-methyl RNA microarray of 9-mers to deduce secondary structure. Selecting and scoring possible structures by expected binding of strongly binding 9-mers gave a secondary structure containing 100% of the known canonical base pairs. A partition function method that considers only possible structures that are consistent with the binding to the microarray produced a consensus structure containing 87% of known base pairs. These results can be compared to 52 and 72% accuracy obtained by Mfold 3.2 and RNAstructure 4.1, respectively, on the basis of energetics alone. Comparisons between predicted binding, microarray, and chemical modification data may provide insight into pseudoknots and tertiary interactions. Interpretation of microarray binding should improve as more is learned about the factors affecting the binding of oligonucleotides to folded RNA. The microarray experiment should facilitate the design of siRNA and antisense oligonucleotides (*89*−*93*) and help reveal structure−function relationships of newly discovered RNA sequences.

SUPPORTING INFORMATION AVAILABLE

Details of nonredundant recursions for free-energy minimization, pictures of a dot plot, and gel from a Ribonuclease H cleavage experiment. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Miranda-Rios, J., Navarro, M., and Soberon, M. (2001) A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9736−9741.
2. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs, *Science 294*, 853−858.
3. Cullen, B. R. (2002) RNA interference: Antiviral defense and genetic tool, *Nat. Immunol. 3*, 597−599.
4. McManus, M. T., and Sharp, P. A. (2002) Gene silencing in mammals by small interfering RNAs, *Nat. Rev. Genet. 3*, 737−747.
5. Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) The structural basis of ribosome activity in peptide bond synthesis, *Science 289*, 920−930.
6. Hansen, J. L., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2002) Structural insights into peptide bond formation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 11670−11675.
7. Zamore, P. D., and Haley, B. (2005) Ribo-gnome: The big world of small RNAs, *Science 309*, 1519−1524.
8. Mattick, J. S., Claverie, J. M., and Pyle, A. M. (2005) Capping by branching: A new ribozyme makes tiny lariats, *Science 309*, 1527−1528.
9. Gesteland, R. F., Cech, T. R., and Atkins, J. F. (2005) *The RNA World*, 3rd ed., Cold Spring Harbor Press, Woodbury, NY.
10. Poy, M. N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P. E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., and Stoffel, M. (2004) A pancreatic islet-specific microRNA regulates insulin secretion, *Nature 432*, 226−230.
11. Bass, B. L. (2001) RNA editing by adenosine deaminases that act on RNA, *Annu. Rev. Biochem. 71*, 817−846.
12. Doherty, E. A., and Doudna, J. A. (2001) Ribozyme structures and mechanisms, *Annu. Rev. Biophys. Biomol. Struct. 30*, 457−475.
13. Moore, P. B., and Steitz, T. A. (2002) The involvement of RNA in ribosome function, *Nature 418*, 229−235.
14. Noller, H. F. (2005) RNA structure: Reading the ribosome, *Science 309*, 1508−1514.
15. Haley, B., and Zamore, P. D. (2004) Kinetic analysis of the RNAi enzyme complex, *Nat. Struct. Mol. Biol. 11*, 599−606.
16. Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA, *Nature 431*, 343−349.
17. Dorsett, Y., and Tuschl, T. (2004) siRNAs: Applications in functional genomics and potential as therapeutics, *Nat. Rev. Drug Discovery 3*, 318−329.
18. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi, *Cell 110*, 563−574.
19. Pfeffer, S., Zavolan, M., Grasser, F. A., Chien, M., Russo, J. J., Ju, J., John, B., Enright, A. J., Marks, D., Sander, C., and Tuschl, T. (2004) Identification of virus-encoded microRNAs, *Science 304*, 734−736.
20. Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005) Silencing of microRNAs in vivo with "antagomirs", *Nature 438*, 685−689.
21. Vaughn, M. W., and Martienssen, R. (2005) It's a small RNA world, after all, *Science 309*, 1525−1526.
22. Chiu, Y. L., and Rana, T. M. (2002) RNAi in human cells: Basic structural and functional features of small interfering RNA, *Mol. Cell. 10*, 549−561.
23. Schubert, S., Grunweller, A., Erdmann, V. A., and Kurreck, J. (2005) Local RNA target structure influences siRNA efficacy:

Systematic analysis of intentionally designed binding regions, *J. Mol. Biol. 348*, 883−893.
24. Heale, B. S. E., Soifer, H. S., Bowers, C., and Rossi, J. J. (2005) siRNA target site secondary structure predictions using local stable substructures, *Nucleic Acids Res. 33*, Feb 18, e30.
25. Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs, *Nature 433*, 769−773.
26. Bohula, E. A., Salisbury, A. J., Sohail, M., Playford, M. P., Riedemann, J., Southern, E. M., and Macaulay, V. M. (2003) The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript, *J. Biol. Chem. 278*, 15991−15997.
27. Kretschmer-Kazemi Far, R., and Sczakiel, G. (2003) The activity of siRNA in mammalian cells is related to structural target accessibility: A comparison with antisense oligonucleotides, *Nucleic Acids Res. 31*, 4417−4424.
28. Varani, G., and Tinoco, I., Jr. (1991) RNA structure and NMR spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
29. Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P., and Puglisi, J. D. (2000) Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function, *Nat. Struct. Biol. 7*, 1105−1110.
30. D'Souza, V., Dey, A., Habib, D., and Summers, M. F. (2004) NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus, *J. Mol. Biol. 337*, 427−442.
31. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat. Struct. Biol. 6*, 212−216.
32. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
33. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution, *Science 289*, 905−920.
34. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution, *Science 310*, 827−834.
35. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., and Noller, H. F. (2001) Crystal structure of the ribosome at 5.5 Å resolution, *Science 292*, 883−896.
36. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature 407*, 327−339.
37. Mears, J. A., Cannone, J. J., Stagg, S. M., Gutell, R. R., Agrawal, R. K., and Harvey, S. C. (2002) Modeling a minimal ribosome based on comparative sequence analysis, *J. Mol. Biol. 321*, 215−234.
38. Pace, N. R., Thomas, B. C., Woese, C. R. (1999) Probing RNA structure, function, and history by comparative analysis, *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 113−141, Cold Spring Harbor Laboratory Press, Woodbury, NY.
39. Rivas, E., and Eddy, S. R. (1999) A dynamic programming algorithm for RNA structure prediction including pseudoknots, *J. Mol. Biol. 285*, 2053−2068.
40. Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule, *Science 244*, 48−52.
41. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res. 31*, 3406−3415.
42. Mathews, D., Disney, M., Childs, J., Schroeder, S., Zuker, M., and Turner, D. (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.
43. Ding, Y., Chan, C. Y., and Lawrence, C. E. (2004) Sfold web server for statistical folding and rational design of nucleic acids, *Nucleic Acids Res. 32*, W135−W141.
44. Hofacker, I. L. (2003) Vienna RNA secondary structure server, *Nucleic Acids Res. 31*, 3429−3431.
45. Mathews, D., Sabina, J., Zuker, M., and Turner, D. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
46. 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.
47. Diamond, J., Turner, D., and Mathews, D. (2001) Thermodynamics of three-way multibranch loops in RNA, *Biochemistry 40*, 6971−6981.
48. Stiegler, P., and Zuker, M. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res. 9*, 133−148.
49. Ding, Y., and Lawrence, C. E. (2003) A statistical sampling algorithm for RNA secondary structure prediction, *Nucleic Acids Res. 31*, 7280−7301.
50. Kierzek, E., Kierzek, R., Turner, D. H., and Catrina, I. E. (2006) Facilitating RNA structure prediction with microarrays, *Biochemistry 45*, 581−593.
51. Afanassiev, V., Hanemann, V., and Wolfl, S. (2000) Preparation of DNA and protein micro arrays on glass slides coated with an agarose film, *Nucleic Acids Res. 28*, E66.
52. Van Dam, R. M., Quake, S. R., Sohail, M., Akhtar, S., and Southern, E. M. (2002) Gene expression analysis with universal *n*-mer arrays, *Genome Res. 12*, 145−152.
53. Roberts, R. W., and Crothers, D. M. (1991) Specificity and stringency in DNA triplex formation, *Proc. Natl. Acad. Sci. U.S.A. 88*, 9397−9401.
54. Herschlag, D. (1991) Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: More isn't always better, *Proc. Natl. Acad. Sci. U.S.A. 88*, 6921−6925.
55. Ruschak, A., Mathews, D., Bibillo, A., Spinelli, S., Childs, J., Eickbush, T., and Turner, D. (2004) Secondary structure models of the 3′ untranslated regions of diverse R2 RNAs, *RNA 10*, 978−987.
56. Luan, D. D., and Eickbush, T. H. (1995) RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element, *Mol. Cell Biol. 15*, 3882−3291.
57. Caruthers, M. H., Beaton, G., Wu, J. V., and Wiesler, W. (1992) Chemical synthesis of deoxyoligonucleotides and deoxyoligonucleotide analogs, *Methods Enzymol. 211*, 3−20.
58. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed. (Fasman, G. D., Ed.) p 597, CRC Press, Cleveland, OH.
59. Tsourkas, A., Behlke, M. A., and Bao, G. (2003) Hybridization of 2′-O-methyl and 2′-deoxy molecular beacons to RNA and DNA targets, *Nucleic Acids Res. 31*, 5168−5174.
60. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: Symmetric loops are more stable than asymmetric loops, *Biochemistry 30*, 6428−6436.
61. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−14089.
62. 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.
63. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., 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.
64. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices, *J. Mol. Biol. 86*, 843−853.
65. 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.
66. Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Muller, P., Mathews, D. H., and Zuker, M. (1994) Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding, *Proc. Natl. Acad. Sci. U.S.A. 91*, 9218−9222.
67. Kim, J., Walter, A. E., and Turner, D. H. (1996) Thermodynamics of coaxially stacked helixes with GA and CC mismatches, *Biochemistry 35*, 13753−13761.
68. Walter, A. E., and Turner, D. H. (1994) Sequence dependence of stability for coaxial stacking of RNA helixes with Watson−Crick base paired interfaces, *Biochemistry 33*, 12715−12719.

69. Mathews, D. H. (2004) Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization, *RNA 10*, 1178−1190.
70. Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes, *Biochemistry 34*, 11211−11216.
71. Kierzek, E., Mathews, D. H., Ciesielska, A., Turner, D. H., and Kierzek, R. (2006) Nearest neighbor parameters for Watson−Crick complementary heteroduplexes formed between 2′-O-methyl RNA and RNA oligonucleotides, *Nucleic Acids Res.*, in press.
72. Popenda, M., Biala, E., Milecki, J., and Adamiak, R. W. (1997) Solution structure of RNA duplexes containing alternating CG base pairs: NMR study of r(CGCGCG)$_2$ and 2′-O-Me(CGCGCG)$_2$ under low salt conditions, *Nucleic Acids Res. 25*, 4589−4598.
73. Zuker, M., and Jacobson, A. B. (1998) Using reliability information to annotate RNA secondary structures, *RNA 4*, 669−679.
74. Mathews, D., Burkard, M., Freier, S., Wyatt, J., and Turner, D. (1999) Predicting oligonucleotide affinity to nucleic acid targets, *RNA 5*, 1458−1469.
75. Ding, Y., Chan, C. Y., and Lawrence, C. E. (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble, *RNA 11*, 1157−1166.
76. Znosko, B. M., Silvestri, S. B., Volkman, H., Boswell, B., and Serra, M. J. (2002) Thermodynamic parameters for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges, *Biochemistry 41*, 10406−10417.
77. 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.
78. Furtig, B., Richter, C., Wohnert, J., and Schwalbe, H. (2003) NMR spectroscopy of RNA, *ChemBioChem 4*, 936−962.
79. Southern, E. M., Case-Green, S. C., Elder, J. K., Johnson, M., Mir, K. U., Wang, L., and Williams, J. C. (1994) Arrays of complementary oligonucleotides for analysing the hybridization behaviour of nucleic acids, *Nucleic Acids Res. 22*, 1368−1373.
80. Mir, K. U., and Southern, E. M. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays, *Nat. Biotechnol. 17*, 788−792.
81. Sohail, M., Akhtar, S., and Southern, E. M. (1999) The folding of large RNAs studied by hybridization to arrays of complementary oligonucleotides, *RNA 5*, 646−655.
82. Woolf, T. M., Melton, D. A., and Jennings, C. G. (1992) Specificity of antisense oligonucleotides in vivo, *Proc. Natl. Acad. Sci. U.S.A. 89*, 7305−7309.
83. Stein, C. A., and Krieg, A. M. (1994) Problems in interpretation of data derived from in vitro and in vivo use of antisense oligodeoxynucleotides, *Antisense Res. Dev. 4*, 67−69.
84. Held, G. A., Grinstein, G., and Tu, Y. (2003) Modeling of DNA microarray data by using physical properties of hybridization, *Proc. Natl. Acad. Sci. U.S.A. 100*, 7575−7580.
85. Milner, N., Mir, K. U., and Southern, E. M. (1997) Selecting effective antisense reagents on combinatorial oligonucleotide arrays, *Nat. Biotechnol. 15*, 537−541.
86. Michel, F., and Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis, *J. Mol. Biol. 216*, 585−610.
87. Masquida, B., and Westhof, E. (2005) A modular and hierarchical approach for all-atom RNA modeling, in *The RNA World*, 3rd ed. (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) Cold Spring Harbor Laboratory Press, Woodbury, NY.
88. Tanner, N. K., Schaff, S., Thill, G., Petit-Koskas, E., Crain-Denoyelle, A. M., and Westhof, E. (1994) A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses, *Curr. Biol. 4*, 488−498.
89. Disney, M. D., Haidaris, C. G., and Turner, D. H. (2003) Uptake and antifungal activity of oligonucleotides in *Candida albicans*, *Proc. Natl. Acad. Sci. U.S.A. 100*, 1530−1534.
90. Childs, J. L., Disney, M. D., and Turner, D. H. (2002) Oligonucleotide directed misfolding of RNA inhibits *Candida albicans* group I intron splicing, *Proc. Natl. Acad. Sci. U.S.A. 99*, 11091−11096.
91. Scherr, M., Rossi, J. J., Sczakiel, G., and Patzel, V. (2000) RNA accessibility prediction: A theoretical approach is consistent with experimental studies in cell extracts, *Nucleic Acids Res. 28*, 2455−2461.
92. Sohail, M., Hochegger, H., Klotzbucher, A., Guellec, R. L., Hunt, T., and Southern, E. M. (2001) Antisense oligonucleotides selected by hybridization to scanning arrays are effective reagents in vivo, *Nucleic Acids Res. 29*, 2041−2051.
93. Eckardt, S., Romby, P., and Sczakiel, G. (1997) Implications of RNA structure on the annealing of a potent antisense RNA directed against the human immunodeficiency virus type 1, *Biochemistry 36*, 12711−12721.
94. Fox, G. E., and Woese, C. R. (1975) 5S-RNA secondary structure, *Nature 256*, 505−507.

BI052618X