Published online 13 July 2009

Nucleic Acids Research, 2009, Vol. 37, No. 18 ***e121***
doi:10.1093/nar/gkp587

Improvement of RNA secondary structure prediction using RNase H cleavage and randomized oligonucleotides

Andrew D. Kauffmann, Ryan J. Campagna, Chantal B. Bartels and Jessica L. Childs-Disney*

Department of Chemistry and Biochemistry, Canisius College, 2001 Main St., Buffalo, NY 14208, USA

Received April 12, 2009; Revised June 24, 2009; Accepted June 25, 2009

ABSTRACT

RNA secondary structure prediction using free energy minimization is one method to gain an approximation of structure. Constraints generated by enzymatic mapping or chemical modification can improve the accuracy of secondary structure prediction. We report a facile method that identifies single-stranded regions in RNA using short, randomized DNA oligonucleotides and RNase H cleavage. These regions are then used as constraints in secondary structure prediction. This method was used to improve the secondary structure prediction of *Escherichia coli* 5S rRNA. The lowest free energy structure without constraints has only 27% of the base pairs present in the phylogenetic structure. The addition of constraints from RNase H cleavage improves the prediction to 100% of base pairs. The same method was used to generate secondary structure constraints for yeast $tRNA^{Phe}$, which is accurately predicted in the absence of constraints (95%). Although RNase H mapping does not improve secondary structure prediction, it does eliminate all other suboptimal structures predicted within 10% of the lowest free energy structure. The method is advantageous over other single-stranded nucleases since RNase H is functional in physiological conditions. Moreover, it can be used for any RNA to identify accessible binding sites for oligonucleotides or small molecules.

INTRODUCTION

RNA folds into well-defined tertiary structures that lead to diverse functions including regulation of gene expression (1–3), cellular localization (4), catalysis (5–10) and serving as a structural scaffold (11,12). In many cases, misfolding of an RNA leads to functional incompetence. For example, misfolding leads to catalytic inactivity in group I introns (13–17) and RNase P RNAs (18–20), and to the inability of rRNAs to form scaffolds for ribosomal assembly (21–24). Thus, determination of the secondary and tertiary structure of an RNA and how structure affects function is important not only in order to understand biological function but also to understand how to modulate it.

RNA tertiary structure is a composite of secondary structure elements such as paired regions and various types of single-stranded regions (internal loops, hairpin loops, bulges and multibranch loops). Approximately 46% of the bases in RNA are unpaired or non-canonically paired (25). Often these sites are of functional importance—forming tertiary contacts (26–28), forming scaffolds for binding of other biomolecules (29,30) and metabolites (2,3,31,32) and direct involvement in catalysis (8).

Despite the importance of RNA structure in cellular processes, there are relatively few NMR and X-ray crystal structures of RNA in the Protein Data Bank (PDB; http://www.pdb.org) (33) when compared to available protein structures. This is due to inherent difficulty in crystallizing RNA and the overlapping signals in NMR spectra of nucleic acids. Fortunately, RNA secondary structure can be predicted accurately from sequence by free energy minimization or phylogenetic comparison, and there are a variety of experimental methods that can be used to improve prediction. The program RNAstructure (34,35) uses free energy minimization to predict an ensemble of possible secondary structures for a given RNA sequence. On average, the lowest free energy structure outputted by the program contains 73% of the base pairs predicted correctly. In an ensemble of 1000 suboptimal structures within 10% free energy of the lowest free energy structure, one structure has at least 87% of the base pairs predicted correctly (35). Experimental constraints aid in selection of that structure. The secondary structure then constrains the possible tertiary structures the RNA can adopt. In fact, a recently reported program uses secondary

*To whom correspondence should be addressed. Tel: +1 716 888 2342; Fax: +1 716 888 3112; Email: childsdj@canisius.edu

structures as an essential starting point to predict tertiary structures (36).

Insights into RNA secondary and tertiary structure can be gained by chemical modification and enzymatic mapping. In some cases, chemical modification and enzymatic mapping are conducted under conditions very different from those inside a cell. Such divergence could cause the RNA to fold into a non-native form, creating problems when trying to estimate *in vivo* structure. An alternative approach is to generate secondary structure constraints using a library of DNA oligonucleotides and RNase H, which cleaves DNA/RNA hybrids under a wide variety of conditions at pH 7 (Figure 1) (37). Previously, RNase H cleavage induced by binding of semi- (38,39) and fully randomized DNA libraries (40,41), semi-randomized DNA/2′–*O*–methyl RNA chimeras (42,43) and partially randomized tethered oligonucleotides (44) to RNAs have been used to design antisense oligonucleotides. In addition, designed oligonucleotides probes and RNase H cleavage have been used to identify accessible binding sites in whole cell extracts (45) and to study RNA folding pathways (20,46).

Herein, we describe a facile method to generate secondary structure constraints by identifying nucleotides that are subject to RNase H cleavage when bound to a member of fully randomized DNA libraries of different lengths (Figure 1). This method was applied to the 5S rRNA from *Escherichia coli* and yeast $tRNA^{Phe}$. The secondary structure of the 5S rRNA is poorly predicted by free energy minimization (27% of base pairs) while the tRNA is well predicted (95% of base pairs). The single-stranded constraints generated by RNase H cleavage improved the prediction of the 5S rRNA (only the phylogenetic structure is predicted) while the accuracy of prediction for the tRNA is unchanged. The single-stranded constraints for the tRNA, however, eliminate all other suboptimal structures that ranged in accuracy from 29 to 76%. This method not only generates secondary structure restraints but also identifies accessible binding sites within an RNA that may be appropriate for design of antisense oligonucleotides or small molecules that modulate RNA function.

MATERIALS AND METHODS

General

All experiments were completed with diethyl pyrocarbonate (DEPC) -treated nanopure water. Randomized oligonucleotides were purchased from Integrated DNA Technologies. Yeast $tRNA^{Phe}$ was purchased from Sigma Aldrich.

RNA secondary structure prediction

All secondary structures (Figures 2 and 4) were predicted using the RNAstructure program (version 4.5 or 4.6) (34,35) and the following suboptimal structure parameters: max. % energy difference = 10, max. no. of structures = 20, window size = 3. These parameters were chosen such that no more than five structures were generated in order to simplify design of single sequences. It is assumed that a DNA oligonucleotide is unable to significantly invade fully paired RNA helices. Nucleotides where cleavage occurred were constrained as single-stranded in secondary structure predictions.

OligoWalk

The $\Delta G^{\circ}_{binding}$ values for oligonucleotides binding to the phylogenetic structures, lowest free energy structures and suboptimal structures (Tables 1 and 2) were predicted using the OligoWalk program (47,48), which is part of RNAstructure. The following parameters were used: Mode = Break Local Structure; Oligo Length = 5, 6, 7 or 8; Oligo Concentration = 5 μM ($\Delta G^{\circ}_{binding}$ values are independent of concentration); Oligomer Chemistry = DNA; and Target Structure Limits for Walk: Start = 1 and Stop = 120.

Preparation of *E. coli* 5S rRNA

The *E. coli* 5S rRNA was prepared by overexpression from the pKK5-1 plasmid as previously reported (49), with the following modifications. The 5S rRNA was isolated by gel purification on a denaturing 8% polyacrylamide gel. The RNA was visualized by UV-shadowing, excised and extracted into 300 mM NaCl by tumbling overnight at 4°C. RNA was concentrated with 2-butanol

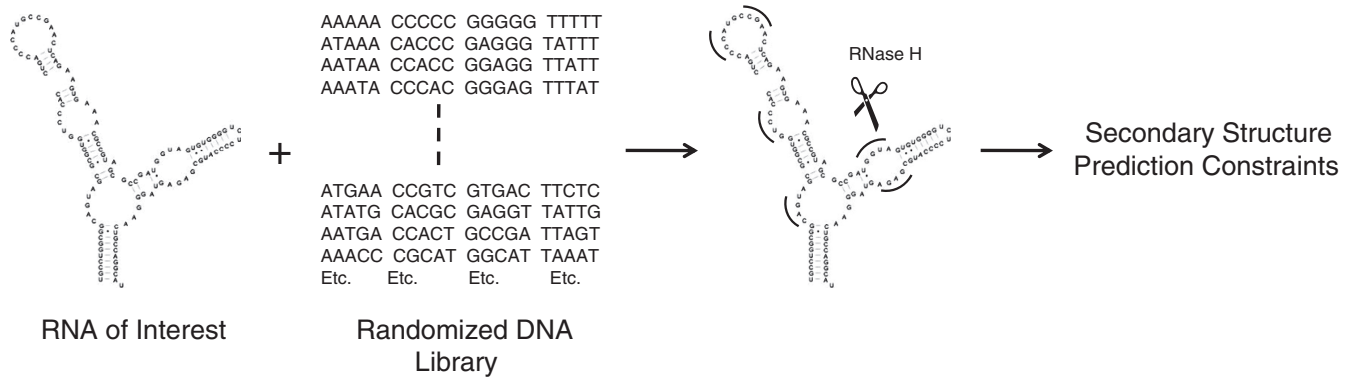

Figure 1. Schematic of the general method used in this study. Randomized DNA oligonucleotides are incubated with an RNA of interest. Only DNAs complementary to single-stranded regions bind, inducing RNase H cleavage of the RNA strand. Nucleotides which are subject to RNase H cleavage are used as single-stranded constraints in RNA secondary structure prediction.

and ethanol precipitated. Concentration was determined by measuring the absorbance at 260 nm and the corresponding extinction coefficient ($1.19 \times 10^6\,M^{-1}\,cm^{-1}$). The extinction coefficient was determined by the HyTher program (50), which is based on nearest neighbor parameters (51).

5′-End labeling of RNAs

The 5S rRNA and $tRNA^{Phe}$ were 5′-end labeled with [γ-^{32}P]ATP (PerkinElmer) and T4 polynucleotide kinase (New England BioLabs) as previously described (52). The RNA was purified and extracted as described above except the gel was exposed to a phosphor screen to visualize the RNA.

Design of DNA oligonucleotides for RNase H cleavage of *E. coli* 5S rRNA

Using the phylogenetic secondary structure of *E. coli* 5S rRNA and the secondary structures predicted by RNAstructure, DNA oligonucleotides were designed to distinguish between them. Negative controls were also designed that bind to a region that is double stranded in the phylogenetic structure and the predicted structures. These predictions were confirmed using OligoWalk.

RNase H cleavage experiments

The *E. coli* 5S rRNA was folded into Form A (53) in 50 µl of 1× assay buffer (150 mM NaCl, 4 mM $MgCl_2$, 10 mM Tris–HCl, pH 7.4) by heating at 65°C for 5 min and slow cooling (~0.5°C/min) to 37°C (53,54). For digestions using single sequences, DNA and DTT were added to final concentrations of 5 µM and 10 mM, respectively. For digestions with the randomized oligonucleotides, a final DNA concentration of 3.25 µM each possible 5-mer, 815 nM each possible 6-mer or 200 nM each possible 7-mer, and 10 mM DTT were used in a total volume of 150 µl. The RNA and DNA oligonucleotides were incubated for 15 min at 37°C. Then, 5 µl (5 units/µl) RNase H (New England BioLabs) were added, and the samples were incubated at 37°C for 10 h. The samples were ethanol precipitated and resuspended in 1× loading buffer (1 mM Tris, pH 7, 3.5 M urea and 1 mM EDTA). The products were separated on a denaturing 8% polyacrylamide gel. The gels were dried and exposed to a phosphor screen. Images were collected using a BioRad FX imaging system.

Yeast $tRNA^{Phe}$ was folded by heating at 95°C for 3 min in 1× assay buffer without $MgCl_2$ and then placing the sample on ice for 10 min. Then, 10 mM $MgCl_2$ and 200 nM final concentration of each possible 5-mer were added. The sample was allowed to equilibrate at 37°C for 15 min followed by addition of DTT to a final concentration of 10 mM and 25 units of RNase H. Time points were taken at 1, 2 and 4 h. The reaction was quenched by addition of 2.5 volumes of ethanol. Products were separated on a denaturing 10% polyacrylamide gel.

Hydrolysis and T1 ladders

The RNA was incubated in 5 µl of hydrolysis buffer (0.1 M $NaHCO_3$, pH 10 and 1 mM EDTA) for 1 min at 95°C. To stop hydrolysis, 5 µl of 2× loading buffer were added and the samples were stored at –80°C until use. Guanosine residues were identified by T1 ribonuclease cleavage under denaturing conditions. The RNA was incubated in 10 µl of 1× T1 Buffer (25 mM sodium citrate, pH 5, 7 M urea and 1 mM EDTA) and 2.5 units/µl of T1 nuclease at 55°C for 10 min. The samples were stored at –80°C if they were not immediately subjected to gel electrophoresis.

RESULTS

The secondary structure of the *E. coli* 5S rRNA was predicted by free energy minimization using the program RNAstructure (34,35). The output of the program includes the lowest free energy structure and suboptimal structures that are within 10% free energy of the lowest free energy structure. Figure 2 shows the phylogenetic structure, and the lowest free energy structure and the four suboptimal structures that were predicted by the program. The five predicted structures were then compared to the phylogenetic (accepted) structure. Interestingly, the phylogenetic structure has the highest free energy value. The lowest free energy structure has only 27% of the base pairs predicted correctly while the suboptimal structures have between 27 and 59%. These percentages are quite low when considering on average the lowest free energy structure predicted by RNAstructure contains 73% of base pairs found in the corresponding phylogenetic structure (35).

On the basis of the differences in structure, 11 DNA oligonucleotide probes were designed to differentiate between the structures when subjected to an RNase H cleavage assay (Table 1 and Figure 2). RNase H cleaves the RNA phosphodieser backbone of DNA/RNA hybrids. The DNA oligonucleotide should only bind to regions in the RNA that are single stranded. We assumed that the short DNA probes would be unable to invade paired regions. (This assumption is correct based on experimental results.) The oligonucleotides are complementary to the following nucleotides within the *E. coli* 5S rRNA: 2–8 (designed to bind a paired region in all structures and serves as a negative control), 23–28 (designed to bind the phylogenetic structure, the lowest free energy structure, and suboptimal structure #1), 26–32 (designed to bind suboptimal structures #3–5), 34–41 (designed to bind the phylogenetic structure), 38–44 (designed to bind the phylogenetic structure), 42–48 (designed to bind the phylogenetic structure and the lowest free energy structure), 50–54 (designed to bind all structures), 56–60 (designed to bind the phylogenetic structure, the lowest free energy structure and suboptimal structures #2–4), 70–74 (designed to bind suboptimal structures #2 and 4), 74–79 (designed to bind the phylogenetic structure and suboptimal structures #2 and 3) and 97–102 (designed to bind all structures except suboptimal structure #4).

The DNA oligonucleotides were incubated individually with the *E. coli* 5S rRNA in order to determine which sites are accessible. The RNA was first folded into Form A as

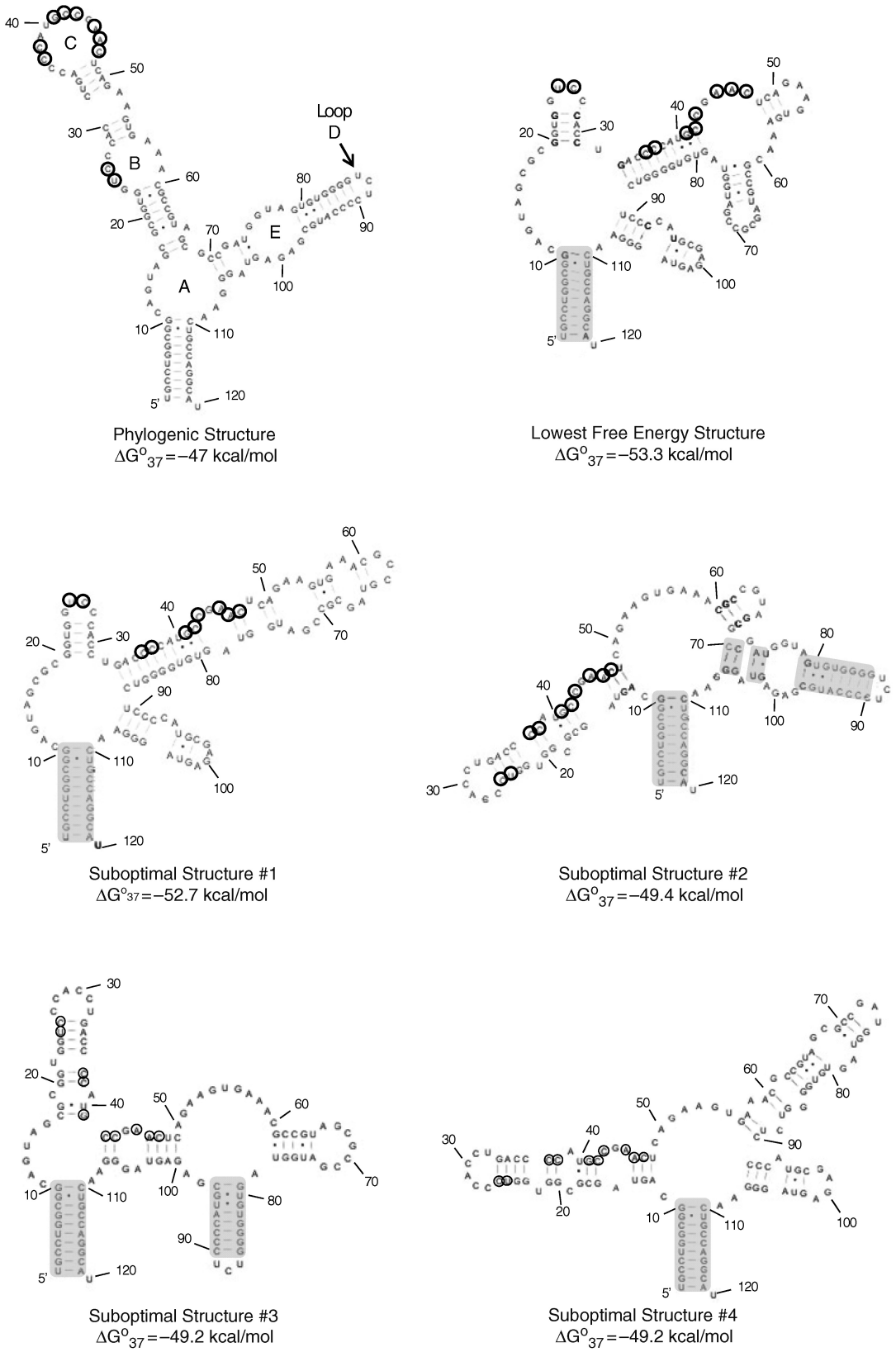

Figure 2. Phylogenetic secondary structure, the predicted lowest free energy secondary structure, and the four suboptimal structures of *E. coli* 5S rRNA. Loops A–E are labeled in the phylogenetic structure. Base-paired regions that are predicted correctly in the suboptimal structures are shaded. Nucleotides cleaved by RNase H cleavage are circled. The lowest free energy structure only has 27% of the base pairs present in the phylogenetic structure.

Table 1. Oligonucleotide sequences and binding free energies to the phylogenetic structure, lowest free energy structure and two suboptimal structures

Oligonucleotide sequence	Binding site	$\Delta G^{\circ}_{binding}$ to phylogenetic structure (kcal/mol)	$\Delta G^{\circ}_{binding}$ to lowest fee energy structure (kcal/mol)	$\Delta G^{\circ}_{binding}$ to suboptimal structure #1 (kcal/mol)	$\Delta G^{\circ}_{binding}$ to suboptimal structure #2 (kcal/mol)	$\Delta G^{\circ}_{binding}$ to suboptimal structure #3 (kcal/mol)	$\Delta G^{\circ}_{binding}$ to suboptimal structure #4 (kcal/mol)	RNase H cleavage observed using a single sequence?	RNase H cleavage observed in this region using randomized oligonucleotides?
5-mers									
GTTTC	56–60	4.1	3.2	1.7	–0.4	–1.9	3.9	No	No
ATCGG	70–74	1.0	–0.5	3.2	1.1	–0.5	3.4	No	No
6-mers									
GGGACC	23–28	–0.2	–3.5	–3.5	2.1	2.1	2.1	No	Yes
ACTTCT	50–54	5.5	–2.3	3.1	–3.5	–3.5	–3.5	No	No
CTACCA	74–79	2.1	10.0	2.5	2.1	9.7	0.2	No	No
CTCTCG	97–102	2.2	–0.3	–0.3	2.2	9.2	–0.3	No	No
7-mers									
GCCAGGC	2–8	15.6	15.6	15.6	15.6	15.6	15.6	No	No
AGGTGGG	26–32	–0.3	1.7	1.7	–4.0	–2.3	–4.0	No	No
CGGCATG	38–44	–6.4	4.9	4.0	2.6	3.3	2.6	Yes	Yes
AGTTCGG	42–48	–4.7	1.0	0.3	–0.1	5.4	0.2	Yes	Yes
8-mers									
CATGGGGT	34–41	–7.4	14.5	14.5	10.0	4.8	10.0	Yes	Yes

'Binding site' indicates the nucleotides in the *E. coli* 5S rRNA to which the oligonucleotide probe is complementary. All $\Delta G^{\circ}_{binding}$ values were generated using the OligoWalk program in 'break local' structure mode. Observation of RNase H cleavage is indicated for each oligonucleotide.

Figure 3. Representative gel autoradiogram of RNase H cleavage experiments to identify single-stranded regions in *E. coli* 5S rRNA. The numbers above the lanes indicate to which nucleotides in the RNA the oligonucleotide probe is complementary.

described previously (53) and then incubated with a DNA probe for 15 min at 37°C prior to addition of RNase H. Cleavage was only observed for three of the oligonucleotides—those that were complementary to nucleotides 34–41, 38–44 and 42–48 (Figure 3). Positions where cleavage occurred were entered as single-stranded constraints in secondary structure prediction by the RNAstructure program. This afforded only one structure, the phylogenetic structure; there were no other structures predicted within 10% free energy.

The same RNase H cleavage assay was then used to determine if randomized DNA oligonucleotides can be used to generate secondary structure prediction constraints. The 5S rRNA was incubated with DNA libraries containing all possible 5-mers (1024 unique probes), 6-mers (4096 unique probes) or 7-mers (16 384 unique probes). A sufficiently high concentration of DNA library was used to ensure that all library members were present at a final concentration of $\geq$200 nM. All three lengths of randomized DNA probes identified four regions accessible to binding and subsequent RNase H cleavage (Figure 3). These regions correspond to nucleotides 26–27 (Loop B), 36–37 (Loop C), 41–44 (Loop C) and 46–48 (Loop C), and are similar to those identified using single DNA sequences. Each is consistent with the phylogenetic structure. As for the single sequences, the positions of RNase H cleavage were used as single-stranded constraints for secondary structure prediction, resulting in the phylogenetic structure.

Cleavage was not observed in Loops A, D or E. For Loop E, this is consistent with the absence of cleavage for the single sequences complementary to nucleotides 74–79 and 97–102. The lack of cleavage in Loop E is consistent with another study that used nucleases to determine the secondary structure of *E. coli* 5S rRNA (55). In that study, only one nucleotide in Loop E (U77) was subject to cleavage by Nuclease S1. In fact, another nucleotide in the loop (G100) was cleaved by the double-stranded Nuclease V1 (cobra venom nuclease). A crystal structure of Loop E has been solved and revealed that Loop E is a highly structured loop containing non-canonical pairs (56). Likewise, cleavage may not be observed in Loop A as nucleotides 11 and 12 are subjected to cleavage by Nuclease V1 although Nuclease S1 does cleave nucleotides 13–16 (55). Loop D is only a three nucleotide hairpin. The stem of this hairpin contains eight base pairs; the four pairs immediately adjacent to the hairpin are GC. Such a helix is likely difficult to invade with a DNA oligonucleotide.

The OligoWalk program (47,48) was then used to determine if RNase H cleavage correlates with accessibility of the corresponding site in the RNA as indicated by the $\Delta G^{\circ}_{\text{binding}}$ value. It was expected that the more negative the value of $\Delta G^{\circ}_{\text{binding}}$, the more likely RNase H cleavage is. Large positive values for $\Delta G^{\circ}_{\text{binding}}$ should indicate a region where oligonucleotide binding would be unlikely. These results are summarized in Table 1. When individual sequences are used, cleavage is only observed if $\Delta G^{\circ}_{\text{binding}}$ was –4.7 kcal/mol or less, even if the region is predicted to be single stranded. The randomized oligonucleotides were able to induce cleavage when $\Delta G^{\circ}_{\text{binding}}$ was much higher, –0.2 kcal/mol (probe 23–28). It is interesting that the oligonucleotide that binds in the same region (26–32) has a similar $\Delta G^{\circ}_{\text{binding}}$ (−0.3 kcal/mol) but does not induce cleavage. However, the probe that binds nucleotides 23–28 forms one less base pair than the probe that binds nucleotides 26–32. It should be noted that 23–28 binds to four single-stranded nucleotides while 26–32 only binds two.

The secondary structure of yeast $tRNA^{Phe}$ (Figure 4) was also predicted using the RNAstructure program. In contrast to the *E. coli* 5S rRNA, $tRNA^{Phe}$ is well predicted with 95% of the base pairs in the phylogenetic structure present in the lowest free energy structure. The only difference between the phylogenetic structure and the lowest free energy structure is the formation of the U7–A66 base pair, which is not predicted by RNAstructure to form. Suboptimal structures #2, #3 and #4 contain 57, 76 and 29% of the base pairs in the phylogenetic structure, respectively (Figure 4). This RNA provides a test case to ensure that addition of constraints generated by RNase H cleavage does not negatively affect secondary structure prediction.

RNase H cleavage was observed for yeast $tRNA^{Phe}$ at the following nucleotides after a 1 or 2 h incubation when a DNA library of 5-mers was used (Figure 5): 18–21 (D-loop), 34–36 (anticodon-loop), 45 and 48 (multibranch loop) 56 and 58–59 (TΨC loop). These positions were then used as single-stranded constraints on secondary structure prediction. Interestingly, only one secondary structure was predicted—the lowest free energy structure, which has 95% of the base pairs predicted correctly. Although constraints do not improve prediction, importantly they do not decrease accuracy. It should be noted that cleavage at nucleotide C49 is also observed after a 4 h incubation period. This nucleotide forms a terminal GC pair (flanking the multibranch loop) and cleavage at this position was not expected. When included as a single-stranded restraint, only one structure is predicted. This structure is similar to the lowest free energy structure predicted without the use of experimental constraints except that the C49–G65 and the U51–A64 pairs are not predicted to form. Cleavage at this position could indicate that the structure of the terminal base pair is dynamic.

As in the analysis of the position of RNase H cleavage for the *E. coli* 5S rRNA, the OligoWalk program was also used to estimate the $\Delta G^{\circ}_{\text{binding}}$ of oligonucleotides that induce cleavage in yeast $tRNA^{Phe}$. As summarized in Table 2, these values are generally consistent with the phylogenetic structure. That is, if the position where cleavage is induced is assumed to be the middle nucleotide of the oligomer, most have negative ΔG° values. There are exceptions, however: the oligonucleotides where positions A21, C48 and U59 are the middle nucleotide. These oligonucleotides have $\Delta G^{\circ}_{\text{binding}}$ values of 2.0, 2.1 and 0.5 kcal/mol, respectively. These nucleotides are at or near the termini of helices. RNase H cleavage at the termini of helices has been observed previously in yeast $tRNA^{asp}$ (38). This suggests that loop dynamics, stacking of dangling ends (57) or coaxial stacking (58) could be important determinants of the sites accessible for binding DNA oligonucleotides (59).

DISCUSSION

The function of an RNA is intimately linked with its structure. Thus, having a reasonable estimate of RNA structure is important for understanding or predicting its function. Accurate structure prediction can be afforded by phylogenetic comparison if many related sequences are known. However, there are many cases in which a large data set is not available. In these cases, RNA secondary structure can be estimated via free energy minimization (35). The accuracy of secondary structure prediction can be improved if experimental constraints are used (35).

Many biochemical methods have been developed in order to gain insight into RNA secondary and tertiary structure including enzymatic mapping and chemical modification (60), equilibrium dialysis with short DNA oligonucleotides (61–63) and microarrays (54,64,65). One potential problem using enzymes is that some require 'non-biological' conditions that can affect RNA structure. For example, the optimal pH range for many nucleases specific for single-stranded regions is 4–5 or 9–10 (66). It is likely that a 2—3 unit pH difference (from pH 7) could cause the RNA to fold differently as the pK_a's of some functional groups present in RNA bases can be in this range (67). Perturbations of RNA structures due to pH changes have been observed in the Hepatitis δ virus (68) and *E. coli* RNase P RNA (18). Many single-stranded nucleases, such as S1, Mung Bean and the single-stranded

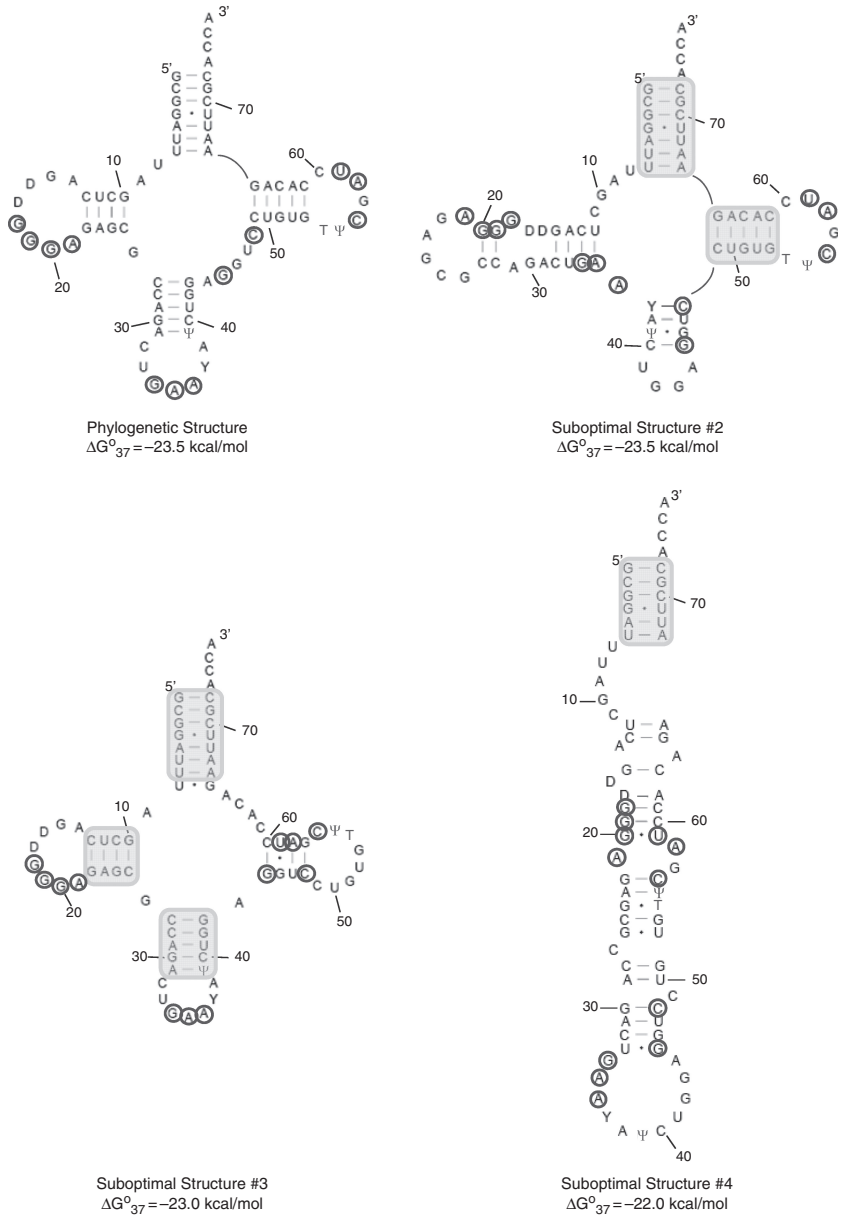

Figure 4. Phylogenetic secondary structure and three predicted suboptimal structures of yeast $tRNA^{Phe}$. The lowest free energy structure has 95% of the base pairs predicted correctly. Base paired regions that are predicted correctly in the suboptimal structures are shaded. Nucleotides cleaved by RNase H after 1 or 2 h incubation are circled. D denotes dihydrouracil while Y denotes wybutosine.

nucleases from *Neurospora crassa* and *Schizopyllum commune*, also require divalent transition metal ions such Zn^{2+} or Co^{2+} (66,69), which can also affect folding. In contrast, RNase H has optimal activity at pH 7.5–9.1, 100 mM monovalent metal ions, and 2–4 mM $MgCl_2$, which are physiological (37).

Numerous methods have been developed to identify accessible binding sites using DNA oligonucleotides and RNase H cleavage (38–45). For example, a 2′-*O*-methyl RNA–DNA chimeric library of 11-mers and RNase H cleavage (43) have been used to identify accessible sites in the human multidrug resistance-1 mRNA (43) and the Hepatitis C Virus (HCV) RNA (40,70,71). Other studies have used accessible binding sites identified by RNase H cleavage to design effective antisense agents (42), tethered oligonucleotide probes to disrupt RNA–protein interactions (44) and ribozymes (41). The positions of cleavage were not used as constraints in RNA secondary structure prediction in any of these cases.

Equilibrium dialysis was used to study the binding of a library of 46 RNA trimers and 57 RNA tetramers to the *E. coli* 5S rRNA, all of which were fully complementary to the RNA sequence (61). The trimers and tetramers that bound the 5S rRNA and are consistent with the phylogenetic structure were complementary to nucleotides 11–14, 39–49, 68–70, 79–81 and 95–98. The probe complementary to nucleotides 68–70 is also complementary to nucleotides 17–19 and 61–63. Thus, although the sequence of

accessible binding sites can be determined, the exact region in the structure to which a probe is binding cannot always be inferred due to the presence of multiple potential binding sites within the target RNA. The method reported herein can alleviate the problem of sequence degeneracy since the binding site is determined by RNase H cleavage.

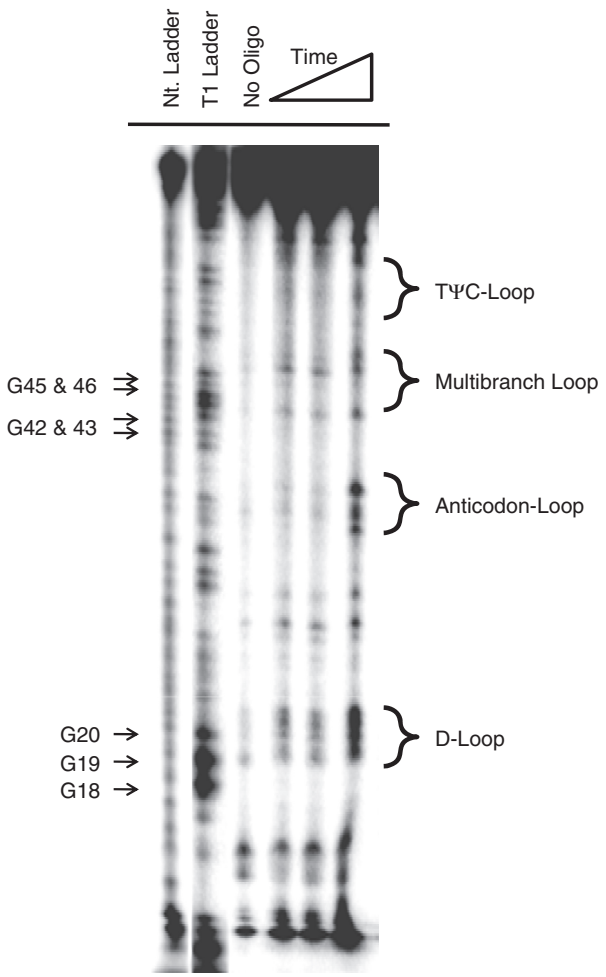

Figure 5. Representative gel autoradiogram of RNase H cleavage experiments to identify single-stranded regions in yeast $tRNA^{Phe}$ with randomized 5-mer oligonucleotides.

A library of 7-mers has also been used to interrogate the structure of *E. coli* 5S rRNA via microarray (54). In these studies, every possible 2′-*O*-methyl 7-mer that could bind the target RNA (114 unique probes) was individually synthesized with a 5′-amino group and immobilized onto microarrays. The arrays were then hybridized with radioactively labeled RNA. In good agreement with the results present in this report, strong binding was observed to probes complementary to nucleotides: 24–30, 25–31, 26–32, 33–39 and 34–40. Medium binding was observed to probes complementary to nucleotides 27–33, 32–38, 35–41, 37–43, 38–44 and 41–47. It should also be noted that the sequence degeneracy problem present in the equilibrium dialysis studies also occurs with microarray methods. For example, the 7-mer probe that binds to nucleotides 35–41 can also bind to nucleotides 90–96 and the probe that binds to nucleotides 36–42 can also bind to nucleotides 91–97. Thus, our results are comparable to the microarray studies except that the binding sites for probes that have the potential to bind multiple regions within the RNA can be deconvoluted.

Binding of oligonucleotides to tRNAs has been studied by equilibrium dialysis (62,63), microarray (59,72) and RNase H cleavage (38). All revealed that there are four accessible sites in tRNAs—the acceptor arm, the anticodon-loop, the D-loop and variable-length region. Two different studies used microarrays to identifying accessible binding sites in yeast $tRNA^{Phe}$ (59,72). The binding patterns were similar despite the differences in the oligonucleotides used. The study reported by Mir and Southern used DNA oligonucleotides ranging in length from monomers to 12-mers (59), while the study reported by Jenek and Kierzek used isoenergetic LNA/2′-*O*-methyl 7-mers (72). Consistent with the data reported herein, binding was observed for oligonucleotides complementary to the D-loop, the variable loop/TΨC stem and the anticodon-loop. Binding was also observed to the double-stranded acceptor stem. This could be in part due to the concentration of salt in the buffers (1 M NaCl) or to the high local

Table 2. Positions of RNase H cleavage of yeast $tRNA^{Phe}$ with randomized 5-mer oligonucleotides after a 1 or 2 h incubation and the corresponding $\Delta G^{\circ}_{binding}$ values for the phylogenetic and suboptimal structures

Cleavage position	Phylogenetic structure $\Delta G^{\circ}_{binding}$ (kcal/mol)	Suboptimal structure #2 $\Delta G^{\circ}_{binding}$ (kcal/mol)	Suboptimal structure #3 $\Delta G^{\circ}_{binding}$ (kcal/mol)	Suboptimal structure #4 $\Delta G^{\circ}_{binding}$ (kcal/mol)
18	−4.1	2.9	−4.1	1.7
19	−5.2	0.4	−5.2	0.6
20	−1.1	−2.0	−1.1	6.9
21	2.0	−0.4	2.0	6.3
34	−1.9	10.1	−1.3	3.7
35	−2.2	8.6	−2.2	−0.4
36	−1.9	9.2	−1.9	−1.9
45	−0.1	1.7	5.8	2.6
48	2.1	8.8	3.8	9.6
56	−1.2	−1.2	4.4	6.4
58	−1.9	−1.9	3.1	7.4
59	0.5	0.5	2.7	5.7

The $\Delta G^{\circ}_{binding}$ values are for the oligonucleotides in which the nucleotide where cleavage occurs is the middle nucleotide of the resulting duplex.

concentration of oligonucleotide immobilized on the array surface. Accessible binding sites in a related tRNA (yeast $tRNA^{Asp}$) were also determined using RNase H cleavage and a semi-randomized DNA library. This library was generated by digesting the corresponding DNA template with DNase I. In good agreement with this report, strong cleavage sites were observed at nucleotides A21 (end of the D-loop), U35 (middle of the anticodon-loop), C56 (middle of the TΨC loop) and G73 (single-stranded region of acceptor stem).

Not all single-stranded regions in the *E. coli* 5S rRNA and yeast $tRNA^{Phe}$ were identified using this method. This is perhaps not unexpected based on previous reports in which the RNAs were probed for binding to libraries of oligonucleotides in solution (61) or displayed on a microarray surface (54,59,72). Important factors in oligonucleotide hybridization are the base composition (73), stacking at helix termini (dangling ends) (57) and coaxial stacking (58) among others. Coaxial stacking has been observed for both RNAs (74–77). Cleavage was also observed at longer incubation times (4 h) for one nucleotide that is not single stranded in the phylogenetic structure of yeast $tRNA^{Phe}$, C49. Although this did not significantly affect secondary structure prediction accuracy, it is important to consider the relative amount of cleavage (weak or strong) and the incubation time required for RNase H cleavage to be observed.

Using the binding of randomized oligonucleotides and subsequent cleavage by RNase H to identify single-stranded regions is advantageous in many regards. For example, RNase H is active under conditions that are considered to mimic cellular pH and ionic strength. The randomized oligonucleotides can be used to map accessible sites in any RNA and does not require the individual syntheses of every possible DNA that might bind to the RNA of interest. The method is relatively fast and does not require specialized instrumentation such as an arrayer or microarray scanner. The instrumentation required for the RNase H method is common to most biochemistry laboratories.

Previous methods used to identify accessible sites in RNAs such as equilibrium dialysis (61–63) and microarrays (54,64) are able to determine which probes bind but they are not able to determine to which nucleotides. As mentioned previously, this can be problematic if a probe is complementary to more than one site as previously observed for the *E. coli* 5S rRNA using these methods (54,61). The RNase H method, however, also determines the binding site mitigating the inability to deconvolute sequence degeneracy.

Importantly, our method combines RNase H mapping with secondary structure prediction to improve prediction of both poorly predicted RNAs and to eliminate less accurate suboptimal structures from well predicted ones. As reported by other groups, the method can also be used to identify accessible sites for therapeutic intervention by oligonucleotides (antisense or RNA interference approaches) or small molecules. The use of a DNA library instead of designed sequences removes potential biases and prevents accessible sites not predicted by secondary structure prediction programs or programs that predict the stability of oligonucleotide binding from being overlooked. It is also likely that this approach can be applied to RNA folding studies. Previously, RNase H cleavage using single sequences has been used to study the folding of the *Tetrahymena thermophila* ribozyme (46) and the *E. coli* and *Bacillus subtilis* RNase P RNAs (20). By using randomized oligonucleotides in the form of DNA libraries, perhaps a clearer picture of RNA folding could be elucidated.

ACKNOWLEDGEMENTS

The authors thank Prof. Matthew Disney for helpful discussions and critical review of the manuscript.

FUNDING

Grants from Canisius College and the Howard Hughes Medical Institute (institutional grant). Funding for open access charge: Canisius College.

Conflict of interest statement. None declared.

REFERENCES

1. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
2. Winkler,W., Nahvi,A. and Breaker,R.R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, **419**, 952–956.
3. Winkler,W.C., Cohen-Chalamish,S. and Breaker,R.R. (2002) An mRNA structure that controls gene expression by binding FMN. *Proc. Natl Acad. Sci. USA*, **99**, 15908–15913.
4. Keenan,R.J., Freymann,D.M., Stroud,R.M. and Walter,P. (2001) The signal recognition particle. *Annu. Rev. Biochem.*, **70**, 755–775.
5. Zaug,A.J. and Cech,T.R. (1986) The intervening sequence RNA of Tetrahymena is an enzyme. *Science*, **231**, 470–475.
6. Stark,B.C., Kole,R., Bowman,E.J., Altman,S. and Altman,S. (1978) Ribonuclease P: an enzyme with an essential RNA component. *Proc. Natl Acad. Sci. USA*, **75**, 3717–3721.
7. James,B.D., Olsen,G.J., Liu,J.S. and Pace,N.R. (1988) The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell*, **52**, 19–26.
8. Qiao,F. and Cech,T.R. (2008) Triple-helix structure in telomerase RNA contributes to catalysis. *Nat. Struct. Mol. Biol.*, **15**, 634–640.
9. DeRose,V.J. (2002) Two decades of RNA catalysis. *Chem. Biol.*, **9**, 961–969.
10. Fedor,M.J. and Williamson,J.R. (2005) The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell Biol.*, **6**, 399–412.
11. Staley,J.P. and Guthrie,C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, **92**, 315–326.
12. 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.
13. Woodson,S.A. and Cech,T.R. (1991) Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the Tetrahymena intervening sequence RNA. *Biochemistry*, **30**, 2042–2050.
14. Nikolcheva,T. and Woodson,S.A. (1999) Facilitation of group I splicing *in vivo*: misfolding of the Tetrahymena IVS and the role of ribosomal RNA exons. *J. Mol. Biol.*, **292**, 557–567.
15. Pan,J. and Woodson,S.A. (1998) Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.*, **280**, 597–609.

16. Emerick,V.L. and Woodson,S.A. (1993) Self-splicing of the Tetrahymena pre-rRNA is decreased by misfolding during transcription. *Biochemistry*, **32**, 14062–14067.
17. Duncan,C.D. and Weeks,K.M. (2008) SHAPE analysis of long-range interactions reveals extensive and thermodynamically preferred misfolding in a fragile group I intron RNA. *Biochemistry*, **47**, 8504–8513.
18. Altman,S. and Guerrier-Takada,C. (1986) M1 RNA, the RNA subunit of *Escherichia coli* ribonuclease P, can undergo a pH-sensitive conformational change. *Biochemistry*, **25**, 1205–1208.
19. Pan,T. and Sosnick,T.R. (1997) Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nat. Struct. Biol.*, **4**, 931–938.
20. Zarrinkar,P.P., Wang,J. and Williamson,J.R. (1996) Slow folding kinetics of RNase P RNA. *RNA*, **2**, 564–573.
21. Hogan,J.J., Gutell,R.R. and Noller,H.F. (1984) Probing the conformation of 26S rRNA in yeast 60S ribosomal subunits with kethoxal. *Biochemistry*, **23**, 3330–3335.
22. Hogan,J.J., Gutell,R.R. and Noller,H.F. (1984) Probing the conformation of 18S rRNA in yeast 40S ribosomal subunits with kethoxal. *Biochemistry*, **23**, 3322–3330.
23. Hogan,J.J. and Noller,H.F. (1978) Altered topography of 16S RNA in the inactive form of *Escherichia coli* 30S ribosomal subunits. *Biochemistry*, **17**, 587–593.
24. Woese,C.R., Magrum,L.J., Gupta,R., Siegel,R.B., Stahl,D.A., Kop,J., Crawford,N., Brosius,J., Gutell,R., Hogan,J.J. *et al.* (1980) Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence. *Nucleic Acids Res.*, **8**, 2275–2293.
25. Gutell,R.R., Weiser,B., Woese,C.R. and Noller,H.F. (1985) Comparative anatomy of 16-S-like ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **32**, 155–216.
26. Torres-Larios,A., Swinger,K.K., Krasilnikov,A.S., Pan,T. and Mondragon,A. (2005) Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*, **437**, 584–587.
27. Kazantsev,A.V., Krivenko,A.A., Harrington,D.J., Holbrook,S.R., Adams,P.D. and Pace,N.R. (2005) Crystal structure of a bacterial ribonuclease P RNA. *Proc. Natl Acad. Sci. USA*, **102**, 13392–13397.
28. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
29. Vila-Sanjurjo,A., Ridgeway,W.K., Seymaner,V., Zhang,W., Santoso,S., Yu,K. and Cate,J.H. (2003) X-ray crystal structures of the WT and a hyper-accurate ribosome from *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **100**, 8682–8687.
30. 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.
31. Blount,K.F. and Breaker,R.R. (2006) Riboswitches as antibacterial drug targets. *Nat. Biotechnol.*, **24**, 1558–1564.
32. Blount,K.F., Wang,J.X., Lim,J., Sudarsan,N. and Breaker,R.R. (2007) Antibacterial lysine analogs that target lysine riboswitches. *Nat. Chem. Biol.*, **3**, 44–49.
33. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
34. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
35. Mathews,D.H., Disney,M.D., Childs,J.L., Schroeder,S.J., Zuker,M. and Turner,D.H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl Acad. Sci. USA*, **101**, 7287–7292.
36. Das,R. and Baker,D. (2007) Automated *de novo* prediction of native-like RNA tertiary structures. *Proc. Natl Acad. Sci. USA*, **104**, 14664–14669.
37. Berkower,I., Leis,J. and Hurwitz,J. (1973) Isolation and characterization of an endonuclease from *Escherichia coli* specific for ribonucleic acid in ribonucleic acid-deoxyribonucleic acid hybrid structures. *J. Biol. Chem.*, **248**, 5914–5921.
38. Matveeva,O., Felden,B., Audlin,S., Gesteland,R.F. and Atkins,J.F. (1997) A rapid *in vitro* method for obtaining RNA accessibility patterns for complementary DNA probes: correlation with an intracellular pattern and known RNA structures. *Nucleic Acids Res.*, **25**, 5010–5016.
39. Wrzesinski,J., Legiewicz,M. and Ciesiolka,J. (2000) Mapping of accessible sites for oligonucleotide hybridization on hepatitis delta virus ribozymes. *Nucleic Acids Res.*, **28**, 1785–1793.
40. Lima,W.F., Brown-Driver,V., Fox,M., Hanecak,R. and Bruice,T.W. (1997) Combinatorial screening and rational optimization for hybridization to folded hepatitis C virus RNA of oligonucleotides with biological antisense activity. *J. Biol. Chem.*, **272**, 626–638.
41. Birikh,K.R., Berlin,Y.A., Soreq,H. and Eckstein,F. (1997) Probing accessible sites for ribozymes on human acetylcholinesterase RNA. *RNA*, **3**, 429–437.
42. Ho,S.P., Bao,Y., Lesher,T., Malhotra,R., Ma,L.Y., Fluharty,S.J. and Sakai,R.R. (1998) Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat. Biotechnol.*, **16**, 59–63.
43. Ho,S.P., Britton,D.H., Stone,B.A., Behrens,D.L., Leffet,L.M., Hobbs,F.W., Miller,J.A. and Trainor,G.L. (1996) Potent antisense oligonucleotides to the human multidrug resistance-1 mRNA are rationally selected by mapping RNA-accessible sites with oligonucleotide libraries. *Nucleic Acids Res.*, **24**, 1901–1907.
44. Cload,S.T. and Schepartz,A. (1994) Selection of structure-specific inhibitors of the HIV Rev-Rev response element complex. *J. Am. Chem. Soc.*, **116**, 437–442.
45. Scherr,M. and Rossi,J.J. (1998) Rapid determination and quantitation of the accessibility to native RNAs by antisense oligodeoxynucleotides in murine cell extracts. *Nucleic Acids Res.*, **26**, 5079–5085.
46. Zarrinkar,P.P. and Williamson,J.R. (1994) Kinetic intermediates in RNA folding. *Science*, **265**, 918–924.
47. Mathews,D.H., Burkard,M.E., Freier,S.M., Wyatt,J.R. and Turner,D.H. (1999) Predicting oligonucleotide affinity to nucleic acid targets. *RNA*, **5**, 1458–1469.
48. Lu,Z.J. and Mathews,D.H. (2008) OligoWalk: an online siRNA design tool utilizing hybridization thermodynamics. *Nucleic Acids Res.*, **36**, W104–W108.
49. Moore,P.B., Abo,S., Freeborn,B., Gewirth,D.T., Leontis,N.B. and Sun,G. (1988) Preparation of 5S RNA-related materials for nuclear magnetic resonance and crystallography studies. *Methods Enzymol.*, **164**, 158–174.
50. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J. Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A•A, C•C, G•G, and T•T mismatches. *Biochemistry*, **38**, 3468–3477.
51. Puglisi,J.D. and Tinoco,I. Jr. (1989) Absorbance melting curves of RNA. *Methods Enzymol.*, **180**, 304–325.
52. Disney,M.D., Testa,S.M. and Turner,D.H. (2000) Targeting a *Pneumocystis carinii* group I intron with methylphosphonate oligonucleotides: backbone charge is not required for binding or reactivity. *Biochemistry*, **39**, 6991–7000.
53. Ciesiollka,J., Lorenz,S. and Erdmann,V.A. (1992) Different conformational forms of *Escherichia coli* and rat liver 5S rRNA revealed by Pb(II)-induced hydrolysis. *Eur. J. Biochem.*, **204**, 583–589.
54. Kierzek,E., Kierzek,R., Turner,D.H. and Catrina,I.E. (2006) Facilitating RNA structure prediction with microarrays. *Biochemistry*, **45**, 581–593.
55. Speek,M. and Lind,A. (1982) Structural analyses of *E. coli* 5S RNA fragments, their associates and complexes with proteins L18 and L25. *Nucleic Acids Res.*, **10**, 947–965.
56. Lu,M. and Steitz,T.A. (2000) Structure of *Escherichia coli* ribosomal protein L25 complexed with a 5S rRNA fragment at 1.8-A resolution. *Proc. Natl Acad. Sci. USA*, **97**, 2023–2028.
57. Freier,S.M., Alkema,D., Sinclair,A., Neilson,T. and Turner,D.H. (1985) Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helixes. *Biochemistry*, **24**, 4533–4539.
58. 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. USA*, **91**, 9218–9222.
59. Mir,K.U. and Southern,E.M. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.*, **17**, 788–792.
60. Ehresmann,C., Baudin,F., Mougel,M., Romby,P., Ebel,J.P. and Ehresmann,B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.*, **15**, 9109–9128.
61. Lewis,J.B. and Doty,P. (1977) Identification of the single-strand regions in *Escherichia coli* 5S RNA, native and A forms, by the binding of oligonucleotides. *Biochemistry*, **16**, 5016–5025.
62. Uhlenbeck,O.C. (1972) Complementary oligonucleotide binding to transfer RNA. *J. Mol. Biol.*, **65**, 25–41.
63. Uhlenbeck,O.C., Baller,J. and Doty,P. (1970) Complementary oligonucleotide binding to the anticodon loop of fMet-transfer RNA. *Nature*, **225**, 508–510.
64. Milner,N., Mir,K.U. and Southern,E.M. (1997) Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat. Biotechnol.*, **15**, 537–541.
65. Duan,S., Mathews,D.H. and Turner,D.H. (2006) Interpreting oligonucleotide microarray data to determine RNA secondary structure: application to the 3′ end of *Bombyx mori* R2 RNA. *Biochemistry*, **45**, 9819–9832.
66. Desai,N.A. and Shankar,V. (2003) Single-strand-specific nucleases. *FEMS Microbiol. Rev.*, **26**, 457–491.
67. Moody,E.M., Brown,T.S. and Bevilacqua,P.C. (2004) Simple method for determining nucleobase pK(a) values by indirect labeling and demonstration of a pK(a) of neutrality in dsDNA. *J. Am. Chem. Soc.*, **126**, 10200–10201.
68. Wadkins,T.S., Shih,I., Perrotta,A.T. and Been,M.D. (2001) A pH-sensitive RNA tertiary interaction affects self-cleavage activity of the HDV ribozymes in the absence of added divalent metal ion. *J. Mol. Biol.*, **305**, 1045–1055.
69. Ando,T. (1966) A nuclease specific for heat-denatured DNA in isolated from a product of *Aspergillus oryzae*. *Biochim. Biophys. Acta*, **114**, 158–168.
70. Smith,R.M. and Wu,G.Y. (2004) Secondary structure and hybridization accessibility of the hepatitis C virus negative strand RNA 5′-terminus. *J. Viral. Hepat.*, **11**, 115–123.
71. Smith,R.M., Walton,C.M., Wu,C.H. and Wu,G.Y. (2002) Secondary structure and hybridization accessibility of hepatitis C virus 3′-terminal sequences. *J. Virol.*, **76**, 9563–9574.
72. Jenek,M. and Kierzek,E. (2008) Isoenergetic microarray mapping–the advantages of this method in studying the structure of Saccharomyces cerevisiae $tRNA^{Phe}$. *Nucleic Acids Symp. Ser. (Oxf.)*, 219–220.
73. Ratmeyer,L., Vinayak,R., Zhong,Y.Y., Zon,G. and Wilson,W.D. (1994) Sequence specific thermodynamic and structural properties for DNA.RNA duplexes. *Biochemistry*, **33**, 5298–5304.
74. Robertus,J.D., Ladner,J.E., Finch,J.T., Rhodes,D., Brown,R.S., Clark,B.F. and Klug,A. (1974) Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature*, **250**, 546–551.
75. Kim,S.H., Suddath,F.L., Quigley,G.J., McPherson,A., Sussman,J.L., Wang,A.H., Seeman,N.C. and Rich,A. (1974) Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science*, **185**, 435–440.
76. Goringer,H.U. and Wagner,R. (1986) Does 5S RNA from *Escherichia coli* have a pseudoknotted structure? *Nucleic Acids Res.*, **14**, 7473–7485.
77. 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.