Published in final edited form as:

J Am Chem Soc. 2009 February 25; 131(7): 2541–2546. doi:10.1021/ja805460e.

Native-like RNA tertiary structures using a sequence-encoded cleavage agent and refinement by discrete molecular dynamics

Costin M. Gherghe¹, Christopher W. Leonard¹, Feng Ding², Nikolay V. Dokholyan^{2,*}, and Kevin M. Weeks^{1,*}

¹ Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599–3290, USA

² Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599–7260, USA

Abstract

The difficulty of analyzing higher order RNA structure, especially for folding intermediates and for RNAs whose functions require domains that are conformationally flexible, emphasizes the need for new approaches for modeling RNA tertiary structure accurately. Here, we report a concise approach that makes use of facile RNA structure probing experiments that are then interpreted using a computational algorithm, carefully tailored to optimize both the resolution and refinement speed for the resulting structures, without requiring user intervention. The RNA secondary structure is first established using SHAPE chemistry. We then use a sequence-directed cleavage agent, that can be placed arbitrarily in many helical motifs, to obtain high quality inter-residue distances. We interpret this in-solution chemical information using a fast, coarse grained, discrete molecular dynamics engine in which each RNA nucleotide is represented by pseudoatoms for the phosphate, ribose and nucleobase groups. By this approach, we refine base paired positions in yeast tRNA^{Asp} to 4 Å RMSD without any preexisting information or assumptions about secondary or tertiary structures. This blended experimental and computational approach has the potential to yield native-like models for the diverse universe of functionally important RNAs whose structures cannot be characterized by conventional structural methods.

Introduction

RNA molecules play a wide variety of functional roles inside the cell, from the tRNA intermediates that carry amino-acids,¹ to complex ribozymes,² and riboswitch regulator motifs.³ Knowledge of the underlying RNA structure in these and many other elements is a fundamental prerequisite to a complete understanding of RNA function. Methods such as X-ray crystallography and NMR spectroscopy offer impressive insight into the details of RNA structure-function relationships. However, both methods require highly structured RNAs and NMR is generally limited to small RNAs. Many important RNA motifs are poor candidates for structural analysis by these well established methods.^{4,5}

For many classes of RNA, three-dimensional structural model building offers a critically important approach. Significant recent progress has been made towards modeling short RNAs that do not contain extensive long-range tertiary interactions based on sequence information alone.^{6–8} For modeling RNAs with complex tertiary structures, a common theme among the

* To whom correspondence should be addressed: weeks@unc.edu, dokh@med.unc.edu.

Supporting Information. Two figures providing additional details regarding SHAPE and site-directed MPE cleavage experiments, plus a dataset file with all MPE constraint data.

most useful programs is the ability to incorporate both local secondary structure information and long-range inter-residue distance constraints. In general, existing programs for modeling complex RNAs either use computationally intensive all-atom reconstruction, which limits use to small RNAs, or employ overly simplified models that omit key structural details. Other challenges in many current approaches are the requirement for high levels of expert user intervention or comparative sequence information and the reliance on chemical intuition derived from preexisting information for tertiary interactions (reviewed in ref. ⁹). Here, we create an approach for accurate *de novo* determination of a complex RNA tertiary fold that does not require expert user intervention nor impose heavy computational requirements and that is efficient for large RNAs.

To successfully reconstruct any RNA tertiary structure, information that defines both the secondary structure and long-range inter-residue distances is required. SHAPE chemistry^{10, 11} (selective 2'-hydroxyl acylation analyzed by primer extension) is a powerful approach for analyzing secondary structure, at single nucleotide resolution, for RNAs of any length. SHAPE exploits the discovery that the 2'-OH group in unconstrained or flexible nucleotides reacts preferentially with hydroxyl-selective electrophilic reagents. In contrast, nucleotides constrained by base pairing or tertiary interactions are unreactive. The resulting reactivity information can be used, in concert with a secondary structure prediction algorithm, to obtain accurate secondary structures.¹²⁻¹⁶

Information regarding the three-dimensional arrangements of RNA helices can be obtained from instructive sets of pair-wise inter-residue distances. The accuracy of any refinement is closely linked to the quality and interpretation of this pair-wise constraint information. The most useful experimental information would come from a small, conformationally restricted probe that yields a large number of high-quality constraints per experiment. These requirements are well met by site-directed hydroxyl radical footprinting. Generally, an Fe(II)-EDTA moiety is tethered to an RNA via sulfur or 2'-amine groups.^{17,18} However, these derivatization steps are experimentally challenging and time consuming. We develop a rapid approach for placing a tethered Fe(II)-EDTA moiety at diverse sites in large RNAs via site-selective intercalation using the reagent methidiumpropyl-EDTA (MPE).¹⁹ We will show that MPE binds selectively at a small, but well-defined, motif. Many regions in a large RNA can thus be probed simply by making small changes to the RNA sequence.

We then use discrete molecular dynamics (DMD)^{8,20} to interpret this secondary and tertiary structural information in terms of a three-dimensional structure. DMD is a special type of molecular dynamics simulation in which pair-wise interactions are approximated by step-wise functions. This approximation enables DMD to sample conformational space more efficiently than with traditional molecular dynamics simulations. We use a coarse-grained RNA model in which each RNA nucleotide is represented as three pseudo-atoms corresponding to the phosphate, sugar and base moieties (see Figure 4A). Because the DMD engine is closely tuned to the resolution of biochemically-derived experimental information, we achieve both rapid refinement times and accurate RNA structures, based on readily obtained chemical reactivity information.

Results

The problem of refining accurate three-dimensional RNA structural models based on in-solution biochemical information can be divided into separate steps for determining the secondary and tertiary structure. Here, we focus our analyses on refining the structure of yeast tRNA^{Asp}. tRNA is a widely accepted benchmark in the use of three dimensional modeling for understanding RNA structure.^{8,21-24} A critical feature of this work is that all constraints for both the base paired secondary structure and for tertiary interactions are derived from efficient

biochemical experiments. No prior assumptions are made based on comparative sequence analysis, chemical intuition, or crystallographic information.

The secondary structure of tRNA^{Asp} is readily determined using SHAPE chemistry.^{15,16,25} SHAPE reactivities are converted to pseudo-free energy change terms and used to constrain a thermodynamically based minimum free energy prediction algorithm as implemented in the RNAstructure program.^{12,14,16} For tRNA^{Asp}, SHAPE-constrained prediction yields a secondary structure that is exactly correct (Figure 1B).

We used the intercalating cleavage agent, MPE,¹⁹ to obtain comprehensive sets of pair-wise inter residue distances (Figure 1A). MPE preferentially intercalates at CpG steps in RNA at sites adjacent to a single nucleotide bulge.^{26,27} In agreement with previous work,²⁸ we find that MPE preferentially binds at a simple motif comprised of three base pairs and a bulged A. The intercalated MPE occupies approximately the same space in a helix as a single RNA base pair.²⁹ Thus, the CGAG/C(C/U)G motif, containing the intercalated methidium group, replaces four continuously paired nucleotides (Figure 1A). MPE can in principle be placed at many helical sites in RNA by replacing four consecutive base pairs with the MPE binding site. MPE intercalates with a preferred orientation in this motif such that the Fe(II)-EDTA moiety points towards the bulged A (dashed circle, Figure 1A).

To obtain comprehensive three-dimensional structural constraints for tRNA^{Asp}, we designed four constructs that placed the MPE binding motif in each of the most stable, G-C rich helices in this RNA: in the acceptor stem (M4, M67), in the T-stem (M49), and in the anticodon stem (M29) (where Mx indicates the nucleotide where the Fe(II)-EDTA group is placed, Figure 1B).

The MPE motif is likely to be an isostructural replacement for many, but not all, four-base-pair sites in folded RNAs. We therefore directly tested whether these four MPE-binding sites are compatible with tRNA^{Asp} structure using SHAPE chemistry. We first measured the baseline SHAPE reactivity profile for the native sequence RNA. As expected,¹⁰ the T-, D- and anticodon loops are reactive, while nucleotides that participate in base pairing and tertiary interactions are unreactive (Figure 2, top). We then compared the native sequence SHAPE reactivity profile with those for the four mutant constructs. For M4, M49 and M67, reactivities were essentially identical to that for the native sequence. In addition, the position of the bulged A, introduced as part of the MPE binding site (Figure 1A), is seen clearly as a single reactive position, exactly at the expected position (Figure 2, central 3 panels, and Supporting Figure 1). The M29 mutant, however, showed marked differences in SHAPE reactivity as compared to the native sequence RNA (Figure 2, bottom) and clearly does not fold to a tRNA-like structure. The MPE binding site thus substitutes for many four base pair elements in a folded RNA; successful sites are readily confirmed by SHAPE. We focused our analysis of site-directed hydroxyl radical footprinting experiments on the three RNA constructs that fold to the correct tRNA structure.

In the presence of H₂O₂, the Fe(II)-EDTA moiety catalyzes the formation of highly reactive hydroxyl radicals that will cleave proximal regions of the RNA backbone. Cleavage intensity (*I*) at each position was calculated as a ratio relative to the mean value for all intensities ($\langle I \rangle$), after subtracting background cleavage observed for the native sequence tRNA that does not contain an MPE binding site. Cleavage ratios less than 1 were taken to be background. This approach eliminates the requirement for user interpretation of significant cleavages and the need to normalize between experiments. As expected, the strongest backbone cleavages occur adjacent to the position at which the Fe(II)-EDTA is located in the MPE binding site (see arrows in the left-hand panels, Figure 3).

Superposition of the backbone cleavages on the crystal structure³⁰ for tRNA^{Asp} shows clearly that the Fe(II)-EDTA moiety is selectively oriented within the MPE binding site (Figure 3).

For example, constructs M4 and M67, in which MPE is placed in different orientations in the acceptor stem, exhibit distinct patterns of cleavage. Critically, all strong backbone cleavages occur in a region in space around the MPE binding site. These experiments thus yield three sets of high quality, long-range, constraints that can be interpreted using DMD refinement (Figure 4).

Each observed cleavage event was defined as a stepwise potential function with a depth quantitatively proportional to the cleavage intensity [$E \propto \ln(I/\langle I \rangle)$], using a fully automated procedure (see Experimental Methods). The interaction potential features a “soft” energy wall at 25 Å, with smaller energy bonuses extending out to 35 Å (Figure 4B). The 25 Å barrier was derived from previous experiments with closely related cleavage agents.^{17,31} Pair-wise distances corresponding to strong cleavages, represented by deep wells, will generally be 25 Å or less. In contrast, pair-wise interactions reported by lower intensity cleavages, corresponding to shallower wells, can more readily achieve values greater than 25 Å without large energy losses. A critical feature of this approach is that no hard cutoff is imposed to distinguish reactive versus unreactive positions. All cleavage data is included and is simply weighted by its relative intensity. This approach has two advantages: (1) no user input is required to decide whether a given cleavage is significant or not and (2) structure refinement is highly tolerant of the measurement errors inherent in any hydroxyl radical footprinting experiment.

Two classes of information were input into the DMD algorithm: base pairing information as determined by SHAPE chemistry (Figure 1B) and long-range constraints derived from site-directed hydroxyl radical cleavage (Figure 3). Refinement is started using as input the RNA sequence and the base pairing information. Once all base pairs have formed, the refinement adds weighted potentials for the experimentally-determined long-range constraints (Figure 4C). The system is then cooled slowly to a DMD temperature of 0.15, corresponding roughly to room temperature,⁸ and finally equilibrated for 10⁵ DMD time units (tu). We generated one tRNA^{Asp} model for every 10 DMD tu, for a total of 10⁴ models per refinement.

We selected representative structures for each refinement using a simple automated algorithm that requires (1) structures to have potential energies lower than the mean value and (2) be at least 1000 tu apart to prevent analysis of consecutive, structurally similar, models. This algorithm consistently yielded sets of about 95 representative structures. We identified predominant RNA conformations using distance-based hierarchical clustering and without user intervention (Figure 4C).³² Final conformations were divided into 10 clusters, subject to the requirement that structures within a cluster agree to better than 5 Å root mean square deviation (RMSD) (Table 1). From the 10 clusters, we focused on the most highly populated ensemble, which contained 30–60% of the total representative models. We focused our analysis on the most central structure in each of these final clusters.

We first assessed the importance of adding long-range distance constraints to the DMD simulations. A baseline DMD simulation of tRNA^{Asp}, using only base pairing information and without any long range constraints, yields a structure in which the pair-wise stacking interactions between the acceptor and D-stems and between the anticodon and T-stems form correctly. This accurate reconstruction of stacking interactions reflects that the three pseudo-atom model for RNA and the inter-atom interactions implemented in our DMD algorithm (Figure 4A) robustly recapitulate local interactions in RNA. Whereas helix stacking interactions were correctly predicted, tertiary interactions in the T- and D-loops were not predicted and the L-shaped native fold of tRNA^{Asp} was not achieved. In the absence of tertiary constraints, tRNA^{Asp} refines to an RMSD of 11.0 Å over base paired positions as compared to the crystal structure (Figure 5A). For an RNA the size of tRNA^{Asp}, an 11.0 Å RMSD reflects a model that does not correspond to the correct three-dimensional structure.

When we implemented the long range-hydroxyl radical footprinting-derived constraints in the DMD simulations, we consistently refined accurate structures for tRNA^{Asp}: RMSD values were 3.8 Å to 5.0 Å over base paired positions (Table 1). Since we performed three independent site-directed cleavage experiments (Figure 3), DMD simulations could be performed using long-range constraints obtained from a single experiment or from any combination of experiments, for a total of seven possible refinements (Table 1). Regardless of which experimental constraints were used, we obtained highly similar predicted structures. In all cases, including our worst-case model (RMSD = 5.0 Å), stacking interactions and the native L-shaped tRNA^{Asp} structure were consistently predicted correctly (Figures 5B,C). This convergence indicates that, for a 75 nt long RNA, one site-directed hydroxyl radical experiment is potentially sufficient for a high-quality analysis of the overall three-dimensional fold.

Discussion

Using a concise blended experimental and computational approach, we show that it is possible to obtain native-like simplified three-dimensional RNA structures for our tRNA^{Asp} test case. No pre-existing or non-experimental information was required, other than the RNA sequence. The RNA secondary structure was established in a single experiment using SHAPE chemistry. Multiple high-quality, long-range, tertiary structure constraints were obtained using site-directed cleavage with MPE. It appears that MPE can be directed to bind at many helical sites to yield efficient tertiary structure information for diverse RNAs.

These experimental data are interpreted using a coarse-grained DMD engine whose underlying three-bead model is comparable to the resolution at which biochemical experiments yield structural constraints for RNA. tRNA^{Asp} structure was consistently refined to an average RMSD of 4.0 Å, relative to the crystallographically determined structure (Figures 5B,C & Table 1). Refinement accuracy is consistent with single nucleotide spatial resolution.

We anticipate that our melded biochemical and computational approach can be applied to a wide variety of RNAs, including large RNAs, and in the absence of prior knowledge of tertiary interactions. If tertiary interactions are known, this information can be readily incorporated into our DMD refinement algorithm by coding tertiary interaction constraints as either pseudo-base pairs or alternative long-range interactions. Accurate, blended biochemical and computational refinement creates many new opportunities for establishing structure-function relationships in large RNAs.

Experimental Methods

RNA constructs

RNAs (tRNA^{Asp} native sequence, 4 tRNA mutants, and a competitor RNA containing an MPE binding site) were synthesized using T7 RNA polymerase-mediated *in vitro* transcription using a single stranded DNA (IDT) with a double-stranded promoter region.³³ tRNA sequences were embedded in the context of 5' and 3' structure cassette¹⁰ sequences. The sequence of the bulge-containing competitor RNA was 5'-GGGCGGUUUU UCGAAGCUUG AGUCUCGUCC GGGCUUCGGC CUGGACGAAG ACUGACGCUC-3'.²⁸ RNAs were purified by denaturing polyacrylamide gel electrophoresis (PAGE), excised from the gel, and recovered by electroelution and ethanol precipitation. Purified RNAs were resuspended in TE [10 mM Tris (pH 8.0), 1 mM EDTA] and stored at -20 °C.

SHAPE analysis

RNAs were refolded by heating to 95 °C for 2 min, cooling on ice, and incubating at 37 °C for 10 min in reaction buffer [50 mM Hepes (pH 7.2), 100 mM NaCl, 10 mM MgCl₂] and then allowed to cool slowly to 22 °C over 15 min. Competitor RNA (folded separately) was added

and incubated with the tRNA for 10 min at 22 °C. Final concentrations (in 10 μ L) were: tRNAs, 0.5 μ M; competitor RNA, 7.2 μ M. RNAs were then treated with 1-methyl-7-nitro-isatoic anhydride¹³ (1M7; 2.5 μ L at 60 mM; in anhydrous DMSO) and allowed to react for 10 min at 22 °C. No-reagent control reactions contained 1 μ L neat DMSO. Modified RNAs were recovered by ethanol precipitation [80 μ L sterile H₂O, 10 μ L NaCl (5 M), 1 μ L glycogen (20 mg/mL), 400 μ L ethanol; 30 min at -80 °C] and resuspended in 5 μ L of TE.

Primer extension

Primer extension reactions were performed using a 5'-[³²P]-labeled primer as described,¹¹ with the exception that the extension reaction was incubated at 52 °C for 7 min. Dideoxy sequencing markers were generated using unmodified RNA. cDNA extension products were separated by gel electrophoresis and visualized by phosphorimaging.

MPE cleavage experiments

MPE was a generous gift from Peter A. Beal (University of California, Davis). tRNAs (either native sequence or mutants containing the MPE binding motif) were refolded as described above with the exception that the reaction buffer contained 100 mM bis-Tris (pH 7.0) in place of Hepes. The MPE-Fe(II) complex was pre-formed by incubating MPE and NH₄Fe(II) SO₄•6H₂O (1.2:1 ratio) for 5 min. The MPE-Fe(II) complex was then added, concurrently with the competitor RNA, to the tRNA solution to a final concentration of 7.0 μ M and allowed to intercalate for 10 min at 22 °C. Control reactions omitted the MPE complex. The cleavage reaction was initiated by addition of DTT and H₂O₂ (final concentrations 5 mM and 0.1%, respectively) and incubated at room temperature for 3 min. Reactions were quenched by ethanol precipitation. RNAs were resuspended in 5 μ L 1/2 $\times$ TE and subjected to primer extension, as described above. The resulting cDNAs were separated by denaturing PAGE (8% w/v, 72 W, 160 min) and quantified by phosphorimaging.

Biochemical data analysis

For all experiments, individual band intensities for the (+) and (-) reagent reactions were integrated using SAFA.³⁴ SHAPE reactivity profiles were obtained by subtracting the no-reagent background from the (+) reaction intensities. For the MPE experiments, baseline reactivity for the wild type and each mutant RNA was calculated by subtracting (-) MPE reaction from the (+) MPE reaction. MPE-specific cleavage in the MPE-binding mutants was calculated by subtracting the reactivity for the native sequence RNA from that for each mutant (for representative data showing native sequence cleavage versus that of an RNA containing the MPE binding site, see supporting Figure 2).

DMD algorithm

In our DMD algorithm for RNA, each nucleotide is represented as three pseudoatoms corresponding to the RNA phosphate (P), ribose (S) and base (B) groups.^{8,20} Atoms P and S are positioned at the center of mass for the phosphate and ribose; the B atom is placed at the center of the six-membered ring for all bases. Inter-pseudoatom interactions are governed by stepwise potential functions. Interactions between pseudoatoms include bonded (including bond angles and dihedrals), non-bonded (base pairing, stacking, hydrophobic, and phosphate-phosphate repulsion), and loop entropy terms.⁸ For this work, constraint potentials, derived from experimental information, are added to bias DMD simulations toward experimentally-supported RNA conformations. Two critical additions to the DMD algorithm were developed. First, additional base pairing interactions between base and phosphate¹⁷ (see Fig. 4A) are used to increase the rigidity of double-helices as observed in known RNA structures. We use the following potential function to model the interactions between base and phosphate:

$$E_{BP} = \begin{cases} 2\varepsilon, & d_{\min} \leq d < d_0 \\ \varepsilon, & d_0 \leq d < d_1 \\ 0, & d_1 \leq d < d_{\max} \\ \infty, & d < d_{\min}, d_{\max} < d. \end{cases}$$

Interaction parameters were determined by analyzing known RNA structures from the Nucleic Acid Database.³⁵ The values of $d_{\min}$, d_0 , d_1 and $d_{\max}$, are 8.90, 9.90, 10.60 and 12.30 Å for purines and 10.00, 11.00, 12.20 and 14.20 Å for pyrimidines, respectively; the repulsive potential, ε , is 0.4 kcal/mol. Second, long-range pair-wise tertiary interactions were modeled by variable-depth potentials in which the energetic component is proportional to the experimentally observed cleavage intensity (Fig. 4B). A favorable interaction was imposed between the appropriate base m in the MPE binding site and the ribose of the cleaved nucleotide i . Cleavage intensity I_i was taken to be proportional to the probability that nucleotide i is within the cleavage range of the MPE binding site. Thus, the attractive potential (Fig. 4B) between base m and ribose i is

$$E_{mi} = \varepsilon \ln (I_i / \langle I_m \rangle)$$

if $I_i / \langle I_m \rangle$ is greater than 1 and zero elsewhere. $\langle I_m \rangle$ is the mean cleavage intensity over all positions induced by MPE at position m ; ε is 1.0 kcal/mol. The full value of E_{mi} is imposed out to 25 Å, smaller energy bonuses extend to 35 Å, resulting in a “soft” attractive potential function (Fig. 4B).

DMD refinement

Sequence information and base pairs (as established by SHAPE^{10,25}) were subjected to one round of refinement by DMD ($T=0.3$, 10^4 tu), where T is the reduced temperature [units are kcal/(mol $\cdot k_B$)].⁸ After base pair formation was confirmed, long range interactions were added. The RNA was allowed to cool in 5 additional steps: (1) $T=0.3$, 10^4 tu; (2) $T=0.25$, 10^4 tu; (3) $T=0.2$, 10^4 tu; (4) $T=0.15$, 10^4 tu; and (5) $T=0.15$, 10^5 tu. This algorithm is efficient: computation time scales linearly with the number of nucleotides.³⁶ One complete three-dimensional refinement of the 75 nt tRNA^{Asp} requires ~2 hrs on an Linux workstation (Intel Pentium 4 CPU at 3.2 GHz, Fedora Core 4 OS). Structures from the final trajectory were then filtered by energy and simulation distance as described in the main text. The requirement that 1000 tu separate each model ensures that structures are not highly correlated but also affords a statistically significant number of models. These were then clustered using OC³² (available at <http://www.compbio.dundee.ac.uk/downloads/oc>) to identify the largest family of similar structures per refinement. We focused on the single most populated cluster because, by the Boltzmann distribution, these clusters represent the lowest free energy state, given our refinement model. RMSD values were computed based on superposition of backbone phosphate atoms at base paired positions. Structure images were composed with Pymol.³⁷

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by NIH grants AI068462 (to K.M.W.) and GM080742 (to N.V.D.).

References

1. Ibba M, Soll D. *Genes Dev* 2004;18:731–738. [PubMed: 15082526]
2. Vicens Q, Cech TR. *Trends Biochem. Sci* 2006;31:41–51. [PubMed: 16356725]
3. Edwards TE, Klein DJ, Ferre-D'Amare AR. *Curr. Opin. Chem. Biol* 2007;17:273–279.
4. Badorrek CS, Weeks KM. *Nature Chem. Biol* 2005;1:104–11. [PubMed: 16408007]
5. Badorrek CS, Weeks KM. *Biochemistry* 2006;45:12664–12672. [PubMed: 17042483]
6. Das R, Baker D. *Proc. Natl. Acad. Sci* 2007;104:14664–14669. [PubMed: 17726102]
7. Parisien M, Major F. *Nature* 2008;452:51–5. [PubMed: 18322526]
8. Ding F, Sharma S, Chalasani V, Demidov V, Broude NE, Dokholyan NV. *RNA* 2008;14:1164–1173. [PubMed: 18456842]
9. Shapiro BA, Yingling YG, Kasprzak W, Bindewald E. *Curr. Opin. Chem. Biol* 2007;17:157–165.
10. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM. *J. Am. Chem. Soc* 2005;127:4223–4231. [PubMed: 15783204]
11. Wilkinson KA, Merino EJ, Weeks KM. *Nature Protocols* 2006;1:1610–1616.
12. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. *Proc. Natl. Acad. Sci. USA* 2004;101:7287–7292. [PubMed: 15123812]
13. Mortimer SA, Weeks KM. *J. Am. Chem. Soc* 2007;129:4144–4145. [PubMed: 17367143]
14. Wilkinson KA, Gorelick RJ, Vasa SM, Guex N, Rein A, Mathews DH, Giddings MC, Weeks KM. *PLoS Biology* 2008;6:e96. [PubMed: 18447581]
15. Wang B, Wilkinson KA, Weeks KM. *Biochemistry* 2008;47:3454–3461. [PubMed: 18290632]
16. Deigan KE, Li TW, Mathews DH, Weeks KM. *Proc. Natl. Acad. Sci. USA* 2009;106:in press
17. Badorrek CS, Gherghe CM, Weeks KM. *Proc. Natl. Acad. Sci. USA* 2006;103:13640–13645. [PubMed: 16945907]
18. Joseph S, Noller HF. *Methods Enzymol* 2000;318:175–90. [PubMed: 10889988]
19. Hertzberg RP, Dervan PB. *J. Am. Chem. Soc* 1982;104:313–315.
20. Dokholyan NV, Buldyrev H, Stanley E, Shakhnovich EI. *Folding Des* 1998;3:577–87.
21. Levitt M. *Nature* 1969;224:759–63. [PubMed: 5361649]
22. Rietveld K, Pleij CW, Bosch L. *EMBO J* 1983;2:1079–85. [PubMed: 6628363]
23. Major F, Gautheret D, Cedergren R. *Proc. Natl. Acad. Sci. USA* 1993;90:9408–9412. [PubMed: 8415714]
24. Sorin EJ, Nakatani BJ, Rhee YM, Jayachandran G, Vishal V, Pande VS. *J. Mol. Biol* 2004;337:789–97. [PubMed: 15033351]
25. Wilkinson KA, Merino EJ, Weeks KM. *J. Am. Chem. Soc* 2005;127:4659–4667. [PubMed: 15796531]
26. White SA, Draper DE. *Nucl. Acids Res* 1987;15:4049–4064. [PubMed: 2438651]
27. White SA, Draper DE. *Biochemistry* 1989;28:1892–1897. [PubMed: 2655698]
28. Gooch BD, Krishnamurthy M, Shadid M, Beal PA. *ChemBioChem* 2005;6:2247–2254. [PubMed: 16245373]
29. Jain SC, Sobell HM. *J. Biomol. Struct. Dyn* 1984;1:1179–94. [PubMed: 6400817]
30. Westhof E, Dumas P, Moras D. *Acta Crystallogr. A* 1988;44:112–123. [PubMed: 3272146]
31. Maity TS, Leonard CW, Rose MA, Fried HM, Weeks KM. *Biochemistry* 2006;45:14955–14964. [PubMed: 17154533]
32. Barton, GJ. Barton, G. J. OC: A cluster analysis program. University of Dundee; Scotland, UK: 2002. (1993, 2002)
33. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. *Nucl. Acids Res* 1987;15:8783–8798. [PubMed: 3684574]
34. Das R, Laederach A, Pearlman SM, Herschlag D, Altman RB. *RNA* 2005;11:344–354. [PubMed: 15701734]
35. Berman HM, Olson WK, Beveridge DL, Westbrook J, Gelbin A, Demeny T, Hsieh S-H, Srinivasan AR, Schneider B. *Biophys. J* 1992;63:751–759. [PubMed: 1384741]

36. Sharma S, Ding F, Dokholyan NV. *Bioinformatics* 2008;24:1951–2. [PubMed: 18579566]
37. DeLano, WL. DeLano, W.L. The Pymol molecular graphics system. DeLano Scientific; South San Francisco, CA, USA:

Figure 2. SHAPE reactivity profiles for the native sequence tRNA^{Asp} and mutant constructs
The native sequence SHAPE profile is shown as a black line in the top panel and as a gray trace when superimposed on the SHAPE profiles for the mutants (blue lines). The site of the bulged A nucleotide, introduced as part of the MPE binding site, is labeled.

Figure 3. Site-directed MPE cleavage experiments

MPE cleavage data are illustrated (1) as histograms of intensity versus nucleotide position, after subtracting the background observed for the native sequence RNA and (2) as sites of cleavage (red spheres) superimposed on the tRNA^{Asp} tertiary structure.³⁰ Red spheres correspond to positions with $I/\langle I \rangle$ greater than 1. MPE insertion sites are illustrated with triangles and black spheres, respectively.

Figure 4. DMD refinement of RNA secondary and tertiary structure

(A) Three pseudoatom model in which an RNA nucleotide is represented by base (B), ribose (S) and phosphate (P) groups. Lines indicate pair-wise potentials between atoms: B_i-B_j , base pairing interactions, $B_{i(j)}-S_{j(i)}$ interaction energies for inter-nucleotide orientation dependences. The $B_{i(j)}-P_{j(i)}$ interaction between base and phosphate reflects rigidity of RNA helices. (B) Energy diagram for the long-range constraint potential. (C) Algorithm for DMD refinement.

Figure 5. tRNA^{Asp} structures obtained by blended experimental and computational refinement (A) tRNA^{Asp} structure predicted in the absence of inter-residue distance constraints. RMSD is 11 Å. (B). Superposition of the seven models refined using one, two, or three sets of constraints (see Table 1). All refinements yield self-consistent families of structures with native-like folds for tRNA^{Asp}. (C) tRNA^{Asp} models corresponding to the 3.8 and 5.0 RMSD refinements (rainbow), superimposed on the crystallographic structure for tRNA^{Asp} (in gray).

Table 1
DMD simulations using site-directed hydroxyl radical cleavage data from one, any two, or all three experiments.

Experiment	RMSD within most highly populated cluster (Å)	RMSD compared to crystal structure (Å)
M4	3.7	4.2
M49	3.8	5.0
M67	3.7	4.3
M4-49	4.0	4.9
M4-67	3.3	3.8
M49-67	3.1	4.0
M4-49-67	3.2	4.0
no constraints	4.6	11.2