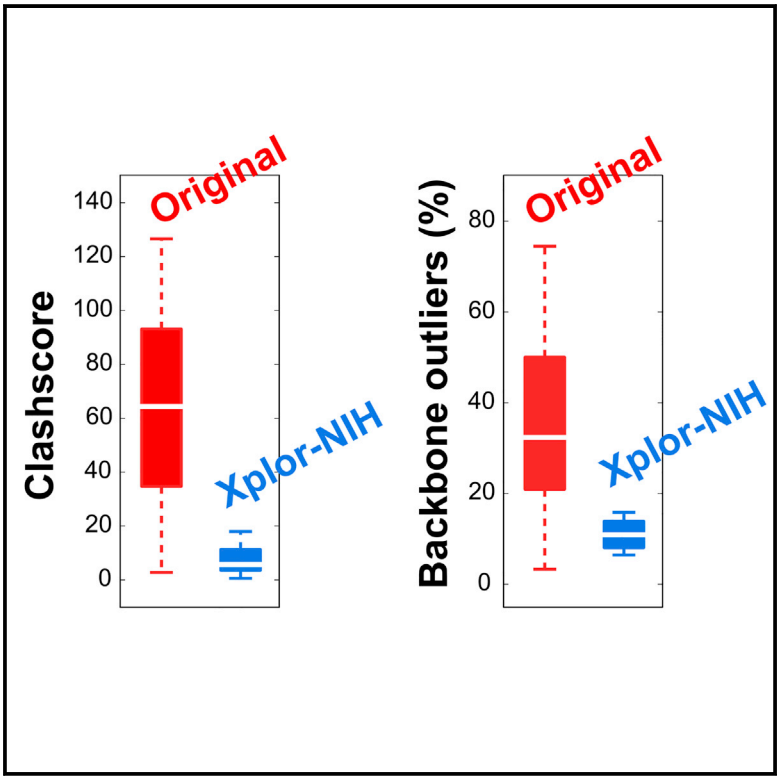
Resource

Structure

Improving NMR Structures of RNA

Graphical Abstract



Authors

Guillermo A. Bermejo, G. Marius Clore, Charles D. Schwieters

Correspondence

charles.schwieters@nih.gov

In Brief

Bermejo et al. show that typical NMR structures of RNA exhibit more steric clashes and conformational ambiguities than X-ray models. They introduce RNA-ff1, a force field for structure calculation with Xplor-NIH. Tested on seven NMR datasets, RNA-ff1 improves geometry, steric contacts, conformation, and residual dipolar coupling cross-validation.

Highlights

- Typical NMR RNA models have more clashes and conformational outliers than X-ray ones
- Force field RNA-ff1 adopts realistic atom radii and a new statistical torsional term
- RNA-ff1 improves geometry, steric contacts, conformation, and RDC-based accuracy

Bermejo et al., 2016, Structure *24*, 806–815
May 3, 2016 ©2016 Elsevier Ltd All rights reserved
http://dx.doi.org/10.1016/j.str.2016.03.007

CellPress

CellPress

Structure
Resource

Improving NMR Structures of RNA

Guillermo A. Bermejo,[1] G. Marius Clore,[2] and Charles D. Schwieters[1,*]
[1]Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892-5624, USA
[2]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA
*Correspondence: charles.schwieters@nih.gov
http://dx.doi.org/10.1016/j.str.2016.03.007

SUMMARY

Here, we show that modern solution nuclear magnetic resonance (NMR) structures of RNA exhibit more steric clashes and conformational ambiguities than their crystallographic X-ray counterparts. To tackle these issues, we developed RNA-ff1, a new force field for structure calculation with Xplor-NIH. Using seven published NMR datasets, RNA-ff1 improves covalent geometry and MolProbity validation criteria for clashes and backbone conformation in most cases, relative to both the previous Xplor-NIH force field and the original structures associated with the experimental data. In addition, with smaller base-pair step rises in helical stems, RNA-ff1 structures enjoy more favorable base stacking. Finally, structural accuracy improves in the majority of cases, as supported by complete residual dipolar coupling cross-validation. Thus, the reported advances show great promise in bridging the quality gap that separates NMR and X-ray structures of RNA.

INTRODUCTION

RNA plays a critical role in the storage and transfer of genetic information, enzymatic catalysis, molecular recognition, and genetic regulation (Gesteland et al., 2006). As a result, knowledge of RNA three-dimensional structure has become an important goal toward understanding its diverse biological functions. However, atomic-level structural studies are hampered by the inherent molecular complexity of RNA, where each of its constituent nucleotide residues has seven torsional degrees of freedom (six along the sugar-phosphate backbone and one around the glycosidic bond) and two major pucker conformations for the sugar ring (C2′-endo and C3′-endo). To complicate matters further, RNA is generally ill behaved from an experimental standpoint, particularly when compared with protein. For example, typically, crystallographic X-ray studies are performed at resolutions worse than 2.5 Å (Keating and Pyle, 2010), and solution-state nuclear magnetic resonance (NMR) suffers from sparseness of its main experimental structural probe (interproton distances) as a result of low proton density and severe chemical shift degeneracy and overlap.

In an attempt to curtail structure inaccuracies caused by the aforementioned hurdles, the most recent report from the X-ray Validation Task Force of the worldwide PDB advocates the use of several quality measures to validate and help correct experimental RNA models (Read et al., 2011). These include the evaluation of conformational plausibility in terms of combinations of backbone torsion angles and sugar pucker configuration, and the assessment of atomic packing by quantification of steric clashes, tools for which are conveniently available in the MolProbity software (Chen et al., 2010; Davis et al., 2007). Because of their generality, these quality criteria apply not only to X-ray models but RNA structures generated by any means, and have been recommended recently to the structural NMR community by the NMR Validation Task Force of the worldwide PDB (Montelione et al., 2013).

NMR spectroscopy has emerged as an important player in the structural characterization of RNA, with solution-state techniques accounting for more than 40% of isolated RNA structures in the PDB, and novel solid-state methodology promising further contributions (Marchanka et al., 2015). The standard approach for nucleic acid (and protein) structure determination by NMR is to minimize a target energy function comprising terms for the experimental data and a priori chemical information. The latter, referred to as the "force field," typically involves bond lengths and angles, torsion angles, planarity, chirality, and interatomic repulsions. Similar to the other covalent energy terms, the torsion angle term—which sometimes is altogether omitted from calculations—can be parameterized using specific consensus torsion values obtained from detailed analysis of high-resolution X-ray structures (e.g., see Parkinson et al., 1996). An alternative nonparametric formulation of the term involves the estimation of torsion angle probability densities from a large database of X-ray structures. By applying the negative logarithm, the densities are subsequently converted into energy potentials, usually referred to as "statistical" or "knowledge-based" (Kuszewski et al., 1996).

The power of statistical torsional potentials to improve conformational quality criteria in protein NMR structures, such as the reduction of Ramachandran outliers, has been thoroughly documented (Bermejo et al., 2012; Bertini et al., 2003; Clore and Kuszewski, 2002; Kuszewski and Clore, 2000; Kuszewski et al., 1996, 1997; Mertens and Gooley, 2005; Yang et al., 2012). On the other hand, to our knowledge there is only one comprehensive RNA study on the subject, performed on a single system (Clore and Kuszewski, 2003). It showed that a statistical torsional term (in combination with another statistical potential that takes into account base–base interaction preferences [Kuszewski et al., 2001]) improved the structural accuracy of an aptamer/theophylline complex, as indicated by complete residual dipolar coupling (RDC) cross-validation. However, despite

 ©2016 Elsevier Ltd All rights reserved

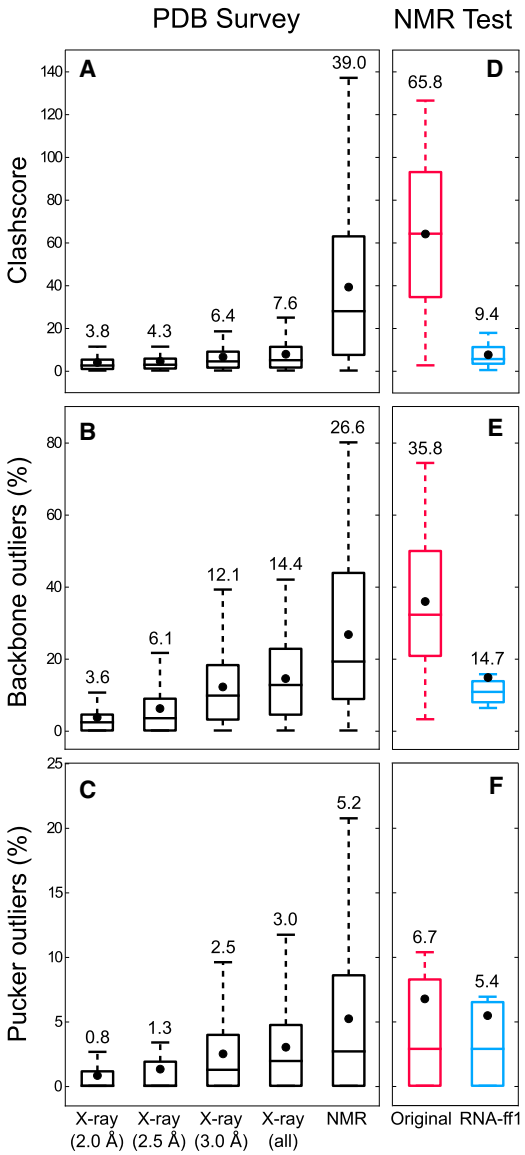


Figure 1. Steric Clash and Backbone Conformational Analysis of RNA Structures

MolProbity statistics are displayed as box plots, where, given a class of structures (e.g., "NMR") and a distribution (e.g., "clashscore"), the following parameters are indicated: minimum and maximum values (box "whiskers"), lower quartile (box bottom), median (line within box), average (dot), and upper quartile (box top). Average values are also shown in numerical form on top of each box. Outliers are omitted to improve readability. (A) and (D) display distributions of clashscore (the number of serious steric overlaps per 1,000 atoms), (B) and (E) the percentage of suites with outlier backbone conformation, and (C) and (F) the percentage of sugar rings with suspicious puckers.

(A–C) X-ray and solution NMR structures deposited in the PDB between January 1, 2010 and June 2, 2014. X-ray models are grouped according to the lower-limit resolution indicated in the *y* axis (e.g., "2.0 Å," 2.0-Å resolution or better; "all," all resolutions).

(D–F) Original NMR models (red; see Table 1) are recomputed in test calculations with the RNA-ff1 force field (blue). The structures involved are the same as those in Figure 4, except that only representative models are considered instead of whole structure bundles.

this encouraging result, analysis of the backbone conformation was not performed, as the corresponding validation tools had not yet been introduced (Murray et al., 2003).

Here, we find that recently solved NMR models of RNA compare unfavorably with their X-ray counterparts under modern structure validation criteria. In this light, using the program Xplor-NIH (Schwieters et al., 2003, 2006), multiple components of the overall NMR energy function are revised. A new force field is introduced with updated covalent parameters and a more realistic representation of atomic radii to improve nonbonded repulsions. Prompted by recent success with proteins (Bermejo et al., 2012), a new statistical torsion angle potential is implemented, significantly smoother than that previously available (Clore and Kuszewski, 2003) and, therefore, less prone to trap calculations in local minima. Based on a benchmark of published NMR datasets for seven different RNA systems in aqueous solution, the new approach improves covalent geometry, conformation, and atomic contacts. Notably, the statistical torsional potential significantly reduces unfavorable backbone configurations while retaining or improving structural accuracy, as supported by complete RDC cross-validation.

RESULTS

X-Ray versus NMR Structures: Steric Clashes and Backbone Conformation

To shed light on the quality of current RNA structure determination efforts, we analyzed the steric contacts and backbone conformation of structures released by the PDB in a ~4.5-year period between 2010 and 2014 (see Experimental Procedures for details) with MolProbity (Chen et al., 2010; Davis et al., 2007). The results are summarized in Figures 1A–1C and explained below.

First, as shown in Figure 1A, X-ray structures of 2.0-Å resolution or better have an average "clashscore" (the number of serious steric overlaps per 1,000 atoms) of 3.8, a value that increases to 7.6 when all resolutions in the set (4.0 Å, the lowest) are considered. In stark contrast, the solution NMR structures released during the same time period have an average clashscore of 39.0 and values as high as 136, which indicates that many structures suffer from poor atomic packing.

Second, it has been proposed that the backbone of RNA, particularly when viewed in terms of the base-to-base conformational unit called "suite" (Figure 2), favors specific torsion angle combinations or rotamers, in a behavior similar to that of protein side chains (Murray et al., 2003) (for standard definitions of RNA torsion angles see Figure 2). Embracing this concept, the structural biology community has compiled a library of consensus rotamer types, many of which have identified functional roles (Richardson et al., 2008). Thus, suite backbone conformations far from those of any rotamer family in the library reflect potential structural errors. Figure 1B shows that high-resolution X-ray structures typically have few such backbone outliers (average of 3.6% for 2.0-Å resolution or better), the proportion of which grows with decreasing resolution, although never as much as that in NMR structures, with an average of 26.6% outliers, and instances with up to 80% nonrotameric suites.

Finally, RNA sugar rings adopt almost exclusively two pucker conformations in high-quality crystallographic data, the

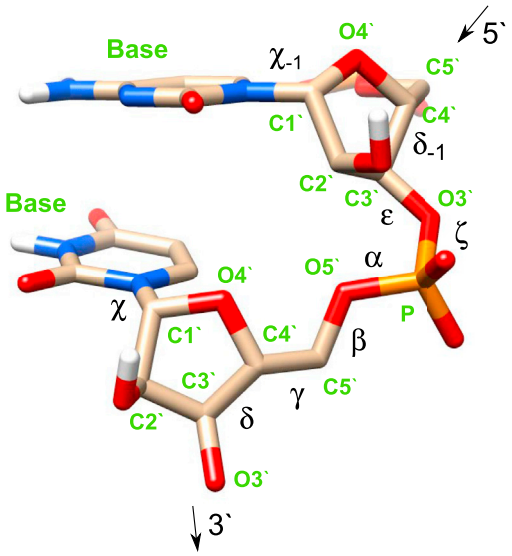


Figure 2. The Suite Conformational Unit

Torsion angles in the base-to-base unit (i.e., the suite) are indicated next to the corresponding rotatable bonds. A "−1" subscript, used to remove name ambiguities, denotes the first residue of the suite (i.e., χ_{-1}, δ_{-1}, ε, and ζ belong to the first residue, all other torsions to the second one). Bases, atom names, and the 5′ and 3′ ends are indicated.

predominant C3′-endo (found in A-form RNA) and the less common C2′-endo (Murray et al., 2003). They can be identified by their distinctive values of both the δ torsion and the "base–phosphate distance" (the length of the line that starts perpendicularly from that traced through the glycosidic bond of the ring in question and ends at the phosphorus atom of the following residue) (Davis et al., 2007). When the pucker indicated by these two criteria conflict with one another, or when either of their values falls outside of their primary ranges, the ring conformation is considered suspicious—an out-of-range value for the ε torsion also indicates ring anomaly. MolProbity pucker evaluation of the subset of the PDB considered here yields a trend by now familiar: on average, the fraction of suspected errors correlates inversely with X-ray resolution, and NMR structures occupy the bad end of this spectrum, with the largest spread (Figure 1C).

RNA-ff1: An Improved Xplor-NIH Force Field for NMR-Based Calculation of RNA Structures

The results of the previous section suggest that, typically, NMR structures display poorer steric and backbone conformational scores than X-ray models. The goal of the present study is to mitigate these shortcomings by improving the conventional molecular dynamics/simulated annealing NMR structure calculation procedure. To this end, a new force field, called RNA-ff1, has been introduced into Xplor-NIH (Schwieters et al., 2003, 2006), as discussed below.

RNA-ff1 uses the covalent parameters compiled by Parkinson et al. (1996), considered the standard for nucleic acids (Read et al., 2011). Our implementation closely follows that of the CNS force field "dna-rna-allatom" (version 1.2) (Brünger et al., 1998), which specifies bond lengths and angles, planarity, and chirality, but lacks torsion angle information. The latter is provided to RNA-ff1 by a newly developed statistical potential described below. Atomic radii, which affect the repulsive-only nonbonded interactions, were changed relative to dna-rna-allatom and previous Xplor-NIH force fields to those used by the MolProbity validation program, as listed elsewhere (Word et al., 1999a). In addition, RNA-ff1 employs a statistical potential that accounts for base–base interaction preferences, as described by Clore and Kuszewski (2003).

It has been shown recently that, in contrast to earlier methods used to implement statistical torsion angle potentials in Xplor-NIH (Kuszewski and Clore, 2000), adaptive kernel density estimation (KDE), coupled with cubic interpolation, is able to reproduce fine features of the energy landscape without sacrificing smoothness (Bermejo et al., 2012). The latter is a desirable quality for molecular dynamics calculations, as rough surfaces may preclude the system from reaching the global minimum (see Discussion). Here, the KDE-based strategy was applied to the development of a new statistical torsional potential for RNA, using a custom database of suite fragments (see Figure 2 for an example) filtered by quality criteria, such as *B*-factor and steric clashes, from the RNA09 set of X-ray models, a successor of earlier structure sets (Murray et al., 2003; Richardson et al., 2008). The new potential, called torsionDB$_{RNA}$ (the RNA version of a similar term recently developed for proteins [Bermejo et al., 2012]), provides the a priori torsion angle knowledge otherwise missing from the RNA-ff1 force field. By default, torsionDB$_{RNA}$ acts on all torsions within each suite of the RNA molecule under study (with the exception of sugar-ring torsions other than δ); the energy surfaces of its terms are considerably smoother than those of the preexisting RNA statistical torsional potential in Xplor-NIH (Clore and Kuszewski, 2003) (Figure 3).

The developments discussed above were tested on the structure calculation of seven RNA systems (Table 1), using published NMR restraints for interatomic distances (from nuclear Overhauser effects [NOEs] and hydrogen bonds), torsion angles, and RDCs. In addition to a standard approach that included all RDCs, the latter were also used in a complete cross-validation scheme, where a randomly selected subset of the data (the test set, in this case comprising 30% of the restraints) was excluded from the calculations. Thus, structural accuracy can be judged by the fit to the test set via R_{free}, an RDC *R*-factor (Clore and Garrett, 1999). To reduce bias, we performed ten such random selections of test set RDCs and averaged the results of the corresponding calculations. Unless otherwise stated, the presented results stem from computations that included the full RDC dataset (the fit to which is measured by the "overall" RDC *R*-factor), while cross-validation is reserved only to assess structure accuracy.

Figure 4 compares the statistics of structures computed with RNA-ff1 (in blue) with those generated by the previous default Xplor-NIH force field (in yellow), henceforth referred to as the "old force field," which includes the older statistical torsional potential (Clore and Kuszewski, 2003). Except for minor differences, the same standard molecular dynamics/simulated annealing protocol was used with both force fields (see Experimental Procedures). Figure 4 also includes analysis of the structures originally published along with the NMR data (here referred to as the "original" models; in red), as well as of control structures calculated with RNA-ff1 stripped of torsionDB$_{RNA}$ to

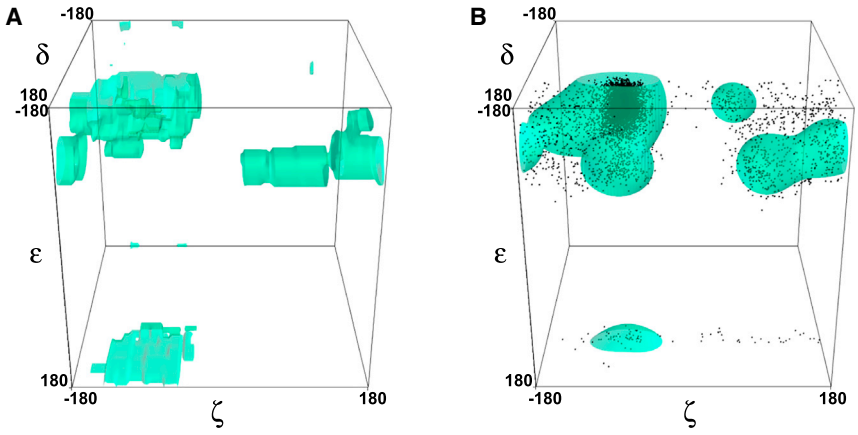


Figure 3. Representative Isoenergetic Surfaces of RNA Statistical Torsion Angle Potentials in Xplor-NIH

Isoenergetic surfaces for the three-dimensional δ–ε–ζ term of the statistical potential are shown (green). Units on all axes are degrees. For details on the effect of surface smoothness in structure calculations, see Discussion and Figures S5 and S6.

(A) Old statistical torsional potential in Xplor-NIH (Clore and Kuszewski, 2003).

(B) New statistical torsional potential, torsionDB$_{RNA}$. Instances from the suite database used to generate the potential are shown (black dots).

ascertain the effect of the latter (in gray). In what follows, for didactical reasons, results are presented in different subsections, and structures are simply referred to by the PDB codes of the original models (see Table 1 for details). In addition, selected statistics from Figure 4 are collected in Figures 1D–1F to facilitate comparison with the larger PDB survey of the previous section (Figures 1A–1C).

Covalent Geometry

RNA-ff1 results in improved covalent geometry relative to both the old force field and the original PDB models, as the corresponding structures reproduce the bond lengths and angles of Parkinson et al. (1996) more faithfully. This is suggested by lower root-mean-square deviations (RMSDs) from bond length and angle equilibrium values in the vast majority of cases (Figures 4A and 4B).

Steric Clashes

Consistent with the larger steric contact analysis of the PDB presented above (Figure 1A), the original NMR structures display relatively poor clashscores, with the exception of PDB: 2KOC (Figures 1D and 4C). The serious clashes, which also plague structures calculated here with the old force field, are significantly reduced, however, by RNA-ff1 (Figures 1D and 4C), mostly through its use of the MolProbity atomic radii (see Experimental Procedures). As an example, Figure 5 provides a side-by-side comparison of clashes in the molecular context of representative structures from the original PDB: 2L1V bundle and RNA-ff1.

Backbone Conformation

As expected from Figure 1B, the original NMR structures suffer from a large proportion of backbone conformational outliers, the most notable exception being PDB: 1O15, which included the older Xplor-NIH statistical torsional potential in its refinement (Clore and Kuszewski, 2003) (Figures 1E and 4D). Indeed, such a potential is very effective at improving the backbone conformation, as suggested here with torsionDB$_{RNA}$: RNA-ff1 calculations (that include torsionDB$_{RNA}$ by default) show a fraction of outliers that is, on average, 56% smaller than control RNA-ff1 calculations without any torsional potential (Figure 4D). Similar conclusions can be drawn from the assessment of sugar-ring pucker conformation (Figures 1F and 4E).

Although the fractions of backbone conformational outliers yielded by the old force field and RNA-ff1 are comparable (Figure 4D), a more detailed analysis of the nonoutlier conformations suggests that RNA-ff1 reproduces more closely the rotamers most commonly observed in nature (thus, presumably the most favorable). This is indicated by improved "suiteness" (Figure 4F), a rotamer match score that ranges from 0 (for an outlier) to 1 (at the center of the distribution of the corresponding rotamer family in the consensus rotamer library) (Richardson et al., 2008). Specific examples of backbone conformational improvements by RNA-ff1 are provided below.

Fit to Experiment

Structures calculated with both the old force field and RNA-ff1 show acceptable agreement with the enforced experimental restraints (Figures 4G–4I). While satisfying RDCs similarly (Figure 4I), however, RNA-ff1 fits experimentally measured interatomic distances and torsion angles better than the old force field, as suggested by smaller deviations from the corresponding restraints in the vast majority of cases (Figures 4G and 4H). It is noteworthy that the original NMR models generally fit all restraint types better than structures computed here. This is expected, as a result of possible differences in weighting of energy terms in the molecular dynamics target functions employed, and the present use of statistical potentials, which are known to conflict with the experimental restraints to some extent (e.g., see Bermejo et al., 2012).

Certainly, the RDCs randomly excluded from the present calculations under the cross-validation scheme should be (and are) better satisfied by the original PDB models (Figure 4J), as the latter included them in their respective computations as part of the entire RDC dataset (i.e., cross-validation applies only to present calculations and not to those of the original structures). The RDCs cross-validated here, which provide information on the orientation of interatomic vectors within a molecular frame, serve an important function: the independent assessment of structural accuracy (Clore and Garrett, 1999). Figure 4J suggests that RNA-ff1 produces more accurate structures than the old force field, as they generally exhibit smaller R_{free} values (where R_{free} of 0% and 100% denote perfect and no agreement with the experimental RDC values, respectively). Furthermore, the significant improvements in backbone conformation by torsionDB$_{RNA}$ (see above) are associated with more accurate structures, as supported in most cases by a drop in R_{free} (by 21%, on average) upon addition of torsionDB$_{RNA}$ to RNA-ff1 calculations with no torsional potential (Figure 4J).

Table 1. RNA Systems Used in Test Calculations

PDB ID	Description	Residues[a]	Reference
2KOC	14-mer cUUCGg tetraloop hairpin	14	Nozinovic et al., 2010
2L5Z	A730 loop of the VS ribozyme from *Neurospora*	26	Desjardins et al., 2011
2M24	5′ splice site of the group IIB intron *Sc*.ai5γ from *Saccharomyces cerevisiae*	29	Kruschel et al., 2014
1O15	high-affinity theophylline-binding aptamer bound to theophylline	33 + ligand	Clore and Kuszewski, 2003[b]
2M57	domain 5 of intron 5 from *Azotobacter vinelandii*	35	Pechlaner et al., 2015
2L1V	$PreQ_1$ riboswitch aptamer domain bound to $PreQ_1$ from *Bacillus subtilis*	36 + ligand	Kang et al., 2009
2LU0	κ–ζ region of domain 1 of group II intron *Sc*.ai5γ from *Saccharomyces cerevisiae*	49	Donghi et al., 2013

[a]Number of residues (presence of small-molecule ligands also indicated).
[b]Study based on NMR data reported elsewhere (Sibille et al., 2001; Zimmermann et al., 1997).

Analysis of Known Structural Motifs

Analysis of RNA structures in the ever-growing public database has enabled the accumulation of detailed knowledge on recurring motifs (Hendrix et al., 2005), such as the A-form double helix, which can be used to assess a molecule of interest in a piecewise manner. The program DSSR (Lu et al., 2015) (part of the 3DNA software suite [Lu and Olson, 2003, 2008]) was used to evaluate the stacking configuration of successive base pairs (i.e., "steps") within the helical stems of the systems in the present calculations. The most interesting trends are observed for the base-pair step parameters slide (Figure 4K) and rise (Figure 4L), which respectively measure an in-plane dislocation and the vertical displacement of a step relative to a local mid-step reference frame (Lu and Olson, 2003; for analysis of all step parameters, see Figure S1). Relative to A-form parameters in high-resolution X-ray structures (Olson et al., 2001) (Figures 4K and 4L, dashed lines), the average slide of all but one of the original NMR models (PDB: 1O15) is small in absolute value (Figure 4K). In general, calculations performed with both the old force field and RNA-ff1 bring the slide closer to the expected X-ray value of −1.53 Å, with RNA-ff1 yielding a slightly better improvement (e.g., see PDB: 2M24 in Figure 4K). Moreover, four out of the seven original PDB models display an average rise considerably larger than the expected 3.32 Å (the van der Waals separation distance between bases, not to be confused with the helical rise, measured relative to the helical axis, expected to be 2.83 Å for A-form [Olson et al., 2001]). In contrast, structures generated by both the old force field and RNA-ff1 are closer to the target step rise, again with RNA-ff1 holding a slight advantage (Figure 4L).

As an example, the single stem of PDB: 2KOC's representative structure, assumed to be an A-form helix (Nozinovic et al., 2010), displays a particularly large separation between base pairs C3–G12 and A4–U11 (rise: 4.33 Å) that is visually evident when compared with that of the RNA-ff1 representative model (rise: 3.33 Å) (Figure 6A). Indeed, this base-pair step defies conformational classification by DSSR in the PDB: 2KOC structure, while it is assigned as A-form (along with the rest of the stem) in the RNA-ff1 structure. The latter enjoys more extensive favorable van der Waals interactions between bases, as suggested by small-probe contact dots (Word et al., 1999a), in another manifestation of improved base stacking (compare Figures 6B and 6C). (Note that although, for clarity, Figures 6B and 6C exclude unfavorable atomic contacts, both structures have similar, low clashscores, suggesting that the larger favorable interaction surfaces do not occur at the expense of overpacking.) In addition, the stem's backbone in the PDB: 2KOC model presents several suites with outlier conformations, while that in the RNA-ff1 structure has all suites in A-form, thus perfectly matching the conformational assignment based on base-pair step parameters (Figure 6A).

The backbone conformation of other structural motifs, such as tetraloops, one of the most characterized types of RNA hairpin loops (Hendrix et al., 2005), also improves with the RNA-ff1 force field, as exemplified with the UUCG tetraloop of structure PDB: 2LU0 (Donghi et al., 2013). Figure 7 shows backbone torsion angle differences relative to a prototypical reference structure of the UUCG tetraloop (the TL1 loop in X-ray structure PDB: 1F7Y) (Ennifar et al., 2000), for the representative model of PDB: 2LU0 and that computed with RNA-ff1. Using the modular nomenclature provided by the program Suitename (Richardson et al., 2008), the sequence and suite backbone conformation of the tetraloop X-ray reference can be succinctly written (as on top of Figure 7): C**1a**U**1z**U**2[**C**6n**G**1a**G, where the two bases flanking the tetraloop are included, and the conformational assignment label for each suite, sandwiched between the corresponding bases, is indicated in bold. For example, the "**1a**" label between C and U denotes A-form for δ–ε–ζ–α–β–γ–δ, where the first three torsions belong to the cytidine and the last four to the uridine, and together make up the suite backbone. With this in mind, every suite within the UUCG tetraloop in the original PDB: 2LU0 structure is an outlier ("**!!**" label), consistent with relatively large deviations from the reference tetraloop for one or more torsions within each suite (Figure 7). On the other hand, the UUCG tetraloop in the RNA-ff1 structure presents no suite outliers, smaller torsion deviations, and an overall conformation identical to that of the reference, with the exception of the last suite (Figure 7). Interestingly, the conformation of the latter is **1c** instead of the reference **1a**, both related by a compensatory crankshaft-like backbone motion (involving primarily torsions α and γ; see Figure 7) that leaves the bases similarly stacked (Richardson et al., 2008). Inspection of the whole structure bundle produced with RNA-ff1 (20 models) reveals a mixture of **1c** and **1a** conformations for this suite, which suggests that the torsionDB$_{RNA}$ potential is not extremely biasing, as it allows access to several of its minima that are consistent with the NMR data. Conformational analysis of other tetraloops is provided in Figures S2–S4.

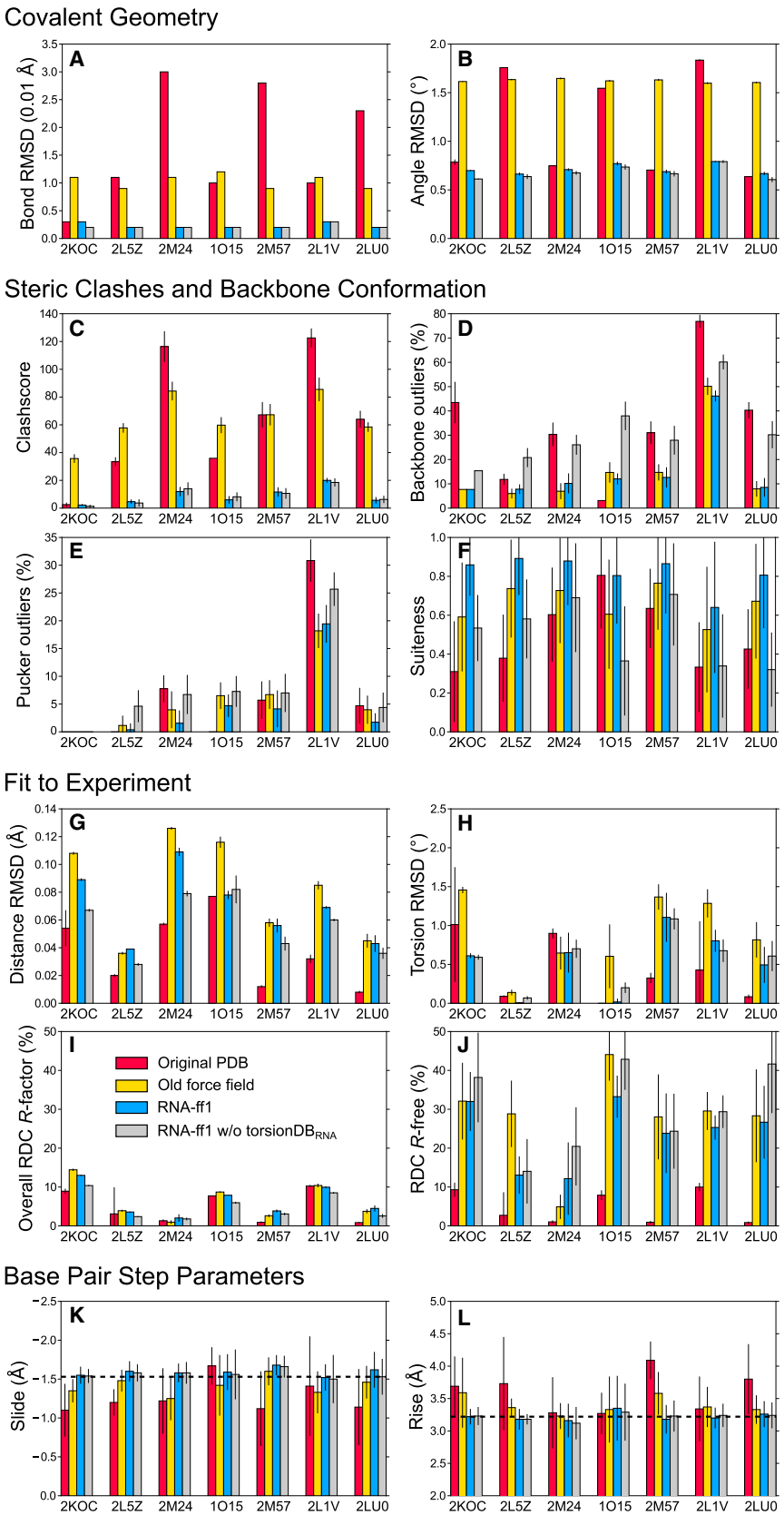


Figure 4. Structural Statistics of Original PDB and Xplor-NIH NMR Models

Statistics on original PDB models (red) and Xplor-NIH structures calculated with the old force field (yellow), RNA-ff1 (blue), and RNA-ff1 without the torsionDB$_{RNA}$ potential (gray) are grouped by the PDB code of the corresponding original NMR bundle (*y* axes). Each statistic represents the average over the structure bundle (SD indicated as error bar). Xplor-NIH calculations included the full NMR dataset, except for (J), associated with the RDC cross-validation scheme. See also Figure S1.

(A) RMS deviation from bond length equilibrium values of Parkinson et al. (1996).

(B) RMS deviation from bond angle equilibrium values of Parkinson et al. (1996).

(C) Clashscore (number of serious steric overlaps per 1,000 atoms).

(D) Percentage of suites with outlier backbone conformation.

(E) Percentage of sugar rings with suspicious puckers.

(F) Suiteness (rotamer match quality score that ranges from 0, for an outlier, to 1, at the center of the distribution of the corresponding rotamer family in the consensus rotamer library [Richardson et al., 2008]).

(G) RMS deviation from distance restraint bounds (NOEs and hydrogen bonds).

(H) RMS deviation from torsion angle restraint bounds.

(I) *R*-factor for the full RDC dataset.

(J) *R*-factor for the test set of RDCs used here under the cross-validation scheme. These RDCs were used as restraints in the computation of the original NMR structures (red bars), such that they fit considerably better than calculations performed here, where the bars represent R_{free} values.

(K) Average slide of base-pair steps within helical stems (a dashed line indicates the A-form value observed in high-resolution X-ray structures [Olson et al., 2001]).

(L) Average rise of base-pair steps within helical stems (a dashed line indicates the A-form value observed in high-resolution X-ray structures [Olson et al., 2001]).

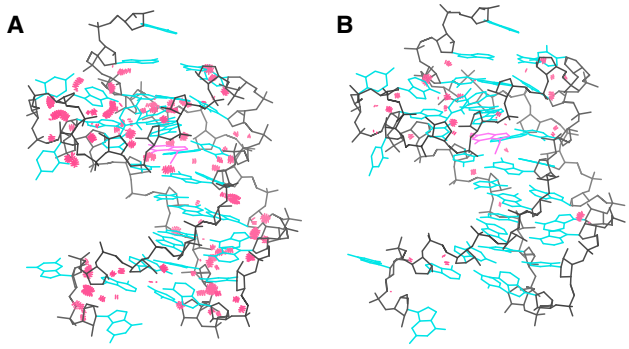


Figure 5. Reduction of Steric Clashes in PDB Structure 2L1V

Lilac spikes indicate serious atomic overlaps. The heavy-atom "stick" molecular representation shows bases in cyan, backbone in black, and ligand in magenta.

(A) Original PDB structure 2L1V (model 1).

(B) Structure with the lowest experimental energy calculated with the RNA-ff1 force field.

DISCUSSION

The majority of the original NMR models associated with the present test calculations (Table 1) systematically display poor quality of atomic contacts and conformation, consistent with typical structure determination efforts, as suggested by our larger survey of the PDB (Figure 1). There are, however, a few exceptions, such as the favorable backbone and base-pair step conformations of PDB: 1O15 (Clore and Kuszewski, 2003), an Xplor-NIH structure originally refined with the old statistical torsional potential (used here in calculations with the old force field) and a statistical base–base positional potential (used here throughout; see Experimental Procedures). While these potentials are the likely reason for the improved conformation, the original torsional potential presents rough surfaces (Figure 3A) that may frustrate molecular dynamics simulations, as intermediate results suggest (Figure S5). This prompted us to develop the torsionDB$_{RNA}$ statistical potential, which, with its smooth surfaces (Figure 3B), not only showed no evidence of compromised sampling but also led to significantly faster calculations (Figure S6).

Another exception in the original NMR structure set is the few steric clashes in the PDB: 2KOC bundle, which has an average clashscore of 2.55, comparable with that of high-resolution X-ray models (Figure 1A). Indeed, PDB: 2KOC, which represents a 5-bp A-form stem capped by a UUCG tetraloop, has been referred to as a "high-resolution" NMR bundle, primarily due to the wealth of structural data collected on this model system (Nozinovic et al., 2010). However, with 43.5% backbone conformational outliers (mostly on the stem; see Figures 6A and S3), PDB: 2KOC seems far from typical high-resolution X-ray values for this metric (average of 3.6% outliers at 2.0-Å resolution). Instead, it seems more at home in the NMR category, where almost 75% of the instances have a smaller proportion of outliers than PDB: 2KOC (Figure 1B). Mainly through torsionDB$_{RNA}$, RNA-ff1 reduces the fraction of backbone outliers to 7.7% in this system, a value close to that of the average 2.5-Å-resolution X-ray structure.

An additional noteworthy case is that of PDB: 2L1V, where although RNA-ff1 significantly reduces the proportion of backbone and pucker outliers relative to the original PDB deposition, they remain relatively high at 46.1% (Figure 4D) and 19.4% (Figure 4E), respectively. This suggests a conflict between torsionDB$_{RNA}$ and the PDB: 2L1V NMR data (where the latter wins), the particular reasons for which are beyond the scope of the present study. However, the observation of such discrepancies is interesting per se, because it suggests problems with the experimental data and/or truly unusual conformations (thus underrepresented in the torsionDB$_{RNA}$ database). Possible experimental issues not only include trivial, correctable errors but also the difficult-to-disentangle effects of inherent molecular motion, under which the conventional premise (followed here) of a unique structure solution becomes less valid. Instead, a structure ensemble may better describe such a dynamical system, the individual members of which should by and large adopt favorable conformations (e.g., as those represented in torsionDB$_{RNA}$). In this regard, Xplor-NIH is particularly well suited, allowing an ensemble representation of the system, as previously exemplified with the refinement of the Dickerson DNA dodecamer (Schwieters and Clore, 2007), an approach that can readily incorporate the torsionDB$_{RNA}$ potential.

Concluding Remarks

We have presented RNA-ff1, a new Xplor-NIH force field for the NMR-based determination of RNA structures. Compared with the older Xplor-NIH force field, RNA-ff1 considerably reduces steric clashes, and modestly but consistently improves covalent geometry, conformation, and fit to experiment. RNA-ff1's use of the new statistical torsional potential, torsionDB$_{RNA}$, speeds up calculations and has a marked effect in discouraging backbone conformational outliers, while concomitantly improving structural accuracy, as supported by complete RDC cross-validation. The smoothness of torsionDB$_{RNA}$ makes it particularly well suited for molecular dynamics simulations, at the heart of most NMR structure calculation schemes. Relative to the original NMR benchmark structures, those based on the RNA-ff1 force field display significantly better validation statistics for steric clashes and both backbone and base-pair step conformation in the vast majority of cases. Thus, the reported advances show great promise in bridging the quality gap that separates NMR and X-ray structures of RNA.

EXPERIMENTAL PROCEDURES

Survey of X-Ray and NMR Structures from the PDB: Steric Clashes and Backbone Conformation

The structures considered in Figures 1A–1C result from a PDB search for entries released between January 1, 2010 and June 2, 2014 that satisfied the following criteria: (1) RNA content only (i.e., no protein, DNA, or RNA/DNA hybrid), and (2) X-ray crystallography or solution NMR as the experimental method. Exclusion of biopolymers other than RNA aimed at simplifying the subsequent structural analysis. The search yielded a total of 173 X-ray and 98 NMR structures. If several models were provided in an NMR entry, only the first one was considered. For meaningful assessment of steric contacts, hydrogen atoms were added to X-ray structures with Reduce (Word et al., 1999b), using "electron-cloud" bond lengths. Entries PDB: 3P4A, 2X2Q, 2M39, and 2KWG were excluded from the analysis due to failed runs of the MolProbity software (Chen et al., 2010; Davis et al., 2007).

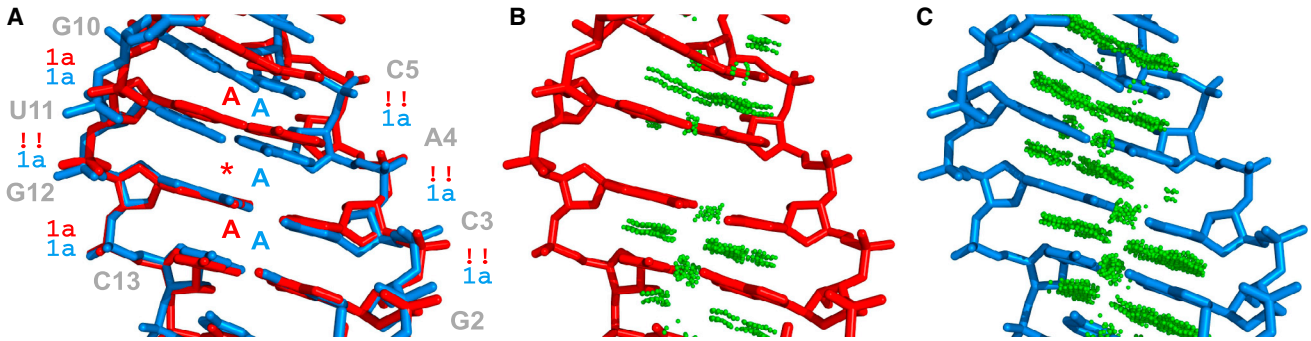


Figure 6. Improvement of A-Form in PDB Structure 2KOC

Heavy-atom "stick" molecular representation of the stem from the representative structure of the original 2KOC bundle (model 1; red) and that of the structure with the lowest experimental energy calculated with the RNA-ff1 force field (blue).

(A) Structure superposition. The base-pair step conformational classification by DSSR is indicated between the corresponding steps ("A," A-form; "*," unclassified). The suite backbone conformation determined with Suitename is indicated next to the corresponding suites ("**1a**," A-form; "**!!**," outlier). The coloring scheme of conformational labels corresponds to that of the molecular representations. Residue names and numbers are indicated (gray).

(B and C) Small-probe contact dots (green spheres) outline favorable van der Waals interaction surfaces between bases.

The RNA-ff1 Force Field

The new force field RNA-ff1 for Xplor-NIH (Schwieters et al., 2003, 2006) specifies all covalent and nonbonded parameters in topology/parameter files nucleic-3.1.top/nucleic-3.1.par, with the exception of torsion angle information, which is handled separately by the new statistical potential torsionDB$_{RNA}$ (see below). Files nucleic-3.1.top/nucleic-3.1.par implement the covalent parameters of Parkinson et al. (1996) as in CNS topology/parameter files dna-rna-allatom.top/dna-rna-allatom.param (version 1.2) (Brünger et al., 1998), except that uniform force constants are given to bond lengths (1,000 kcal mol^{-1} Å^{-2}), bond angles (500 kcal mol^{-1} rad^{-2}), and improper dihedral angles (500 kcal mol^{-1} rad^{-2}). (Note that there is no proper dihedral term in these files.) Nonbonded interactions rely on atomic radii reported elsewhere (Word et al., 1999a), used in a repulsive-only energy term, E_{repel}, given by Nilges et al. (1988):

$$E_{repel} = \begin{cases} k_{vdw}\left[(s_{vdw}\cdot r_{min})^2 - r^2\right]^2 & \text{if } r < s_{vdw}\cdot r_{min} \\ 0 & \text{if } r \geq s_{vdw}\cdot r_{min} \end{cases}, \qquad \text{(Equation 1)}$$

where k_{vdw} is a force constant, r is the interatomic distance, and r_{min} is the sum of the atomic radii, which are scaled by s_{vdw} to account for the absence of an attractive component in the potential. Use of a small s_{vdw} yields unacceptable numbers of steric clashes when judged with realistic atomic radii. Conversely, a large s_{vdw} produces expanded structures that do not satisfactorily agree with the experimental restraints. The optimal s_{vdw} was found here to be 0.9 via a grid search calculation (Figure S7). The atom nomenclature adopted by RNA-ff1 is the same as that in Xplor-NIH topology/parameter files nucleic-1.1.top/nucleic-1.1.par.

The new statistical torsional potential for RNA, torsionDB$_{RNA}$, contributes the torsion angle component of RNA-ff1, and was developed with an overall strategy and tools recently introduced for the analogous protein potential in Xplor-NIH (Bermejo et al., 2012). In brief, a database of suite fragments was used to estimate probability density functions of torsion angles of interest $\boldsymbol{\theta}$. Each function, $p(\boldsymbol{\theta})$, was converted into a potential energy term, $E(\boldsymbol{\theta}) = -\ln p(\boldsymbol{\theta})$, represented by cubic interpolation during the course of structure calculations.

With 287 X-ray PDB/NDB (Nucleic Acid Database) entries of 3.0-Å resolution or better, the RNA09 database from the Richardson Lab (Duke University; downloaded from http://kinemage.biochem.duke.edu/downloads/datasets/RNA09PDBs.tgz) was the starting point for the generation of the custom database used here. The latter consists of 9,195 RNA09 suites, the atoms of which have (1) B-factors $\leq$60 Å^2, (2) no alternative conformations, (3) no serious steric clashes determined by MolProbity, and (4) no sugar-ring pucker conformations deemed suspicious by MolProbity. It is noteworthy that the consensus rotamer library (Richardson et al., 2008) was not explicitly used to bias the compilation of the database. Database torsion angles were measured with the program Dangle (Word et al., 1999a).

The torsion angle conformational space of a suite can be described by an eight-dimensional probability density, which, in turn, can be expressed in terms of densities of dimensionality $\leq$3:

$$p(\delta_{-1},\varepsilon,\zeta,\alpha,\beta,\delta,\chi) = \frac{p(\delta,\chi)}{p(\delta)}\frac{p(\beta,\gamma,\delta)}{p(\beta,\gamma)}\frac{p(\alpha,\beta,\gamma)}{p(\alpha,\beta)}\frac{p(\zeta,\alpha,\beta)}{p(\zeta,\alpha)}\frac{p(\varepsilon,\zeta,\alpha)}{p(\varepsilon,\zeta)}p(\delta_{-1},\varepsilon,\zeta)\,, \qquad \text{(Equation 2)}$$

where δ_{-1}, ε, and ζ belong to the first residue of the suite, and the remaining torsions to the second residue (Figure 2). Equation 2 assumes conditional independence relationships similar to those used elsewhere (Bermejo et al., 2012), and that χ is conditionally independent of all torsion angles but δ. These approximations are needed because cubic interpolation (see below) becomes too expensive computationally beyond three dimensions. Each density function on the right-hand side of Equation 2 was obtained from the suite database via adaptive KDE, using the densityEstimation module of Xplor-NIH, as described previously (Bermejo et al., 2012). The overall window widths of the two- and three-dimensional Gaussian kernels used were 2° and 4°, respectively. All suites in the database were considered in this procedure, regardless of the identity of the bases (i.e., torsionDB$_{RNA}$ is sequence independent).

Application of the negative logarithm to the right-hand side of Equation 2 results in a sum of different energy terms, each evaluated on a uniform grid (spacing: 10°), used for cubic interpolation with periodic boundary conditions. The interpolated terms make up torsionDB$_{RNA}$. Although, by default, torsionDB$_{RNA}$ is applied to all suites of the molecule of interest (as in all computations performed here), arbitrary suites can be excluded using Xplor-NIH's powerful atom selection language. It is noteworthy that the same principles and tools used here to generate torsionDB$_{RNA}$ can be readily applied to create an analogous DNA potential from an appropriate database.

Finally, RNA-ff1 employs a statistical potential that accounts for base–base interaction preferences (Clore and Kuszewski, 2003); details of its implementation are provided below and in Supplemental Experimental Procedures.

Structure Calculations

Experimental NMR restraints for structure calculations were obtained from the PDB, except for those of PDB: 1O15, taken from the Xplor-NIH distribution package. These restraints were implemented as quadratic square-well potentials for interatomic distances and torsion angles, and harmonic potentials for RDCs and base-pair planarity restraints. Although all original structural studies claimed the use of planarity restraints to prevent undue buckling across Watson–Crick pairs (see references in Table 1), several did not publish them along with the rest of the NMR data. In such cases, base pairs were identified by their hydrogen bond distance restraints, and assigned planarity restraints as previously described (Kuszewski et al., 2001), with a uniform force constant of 30 kcal mol^{-1} Å^{-2}. Original planarity restraints, when available, were used as is.

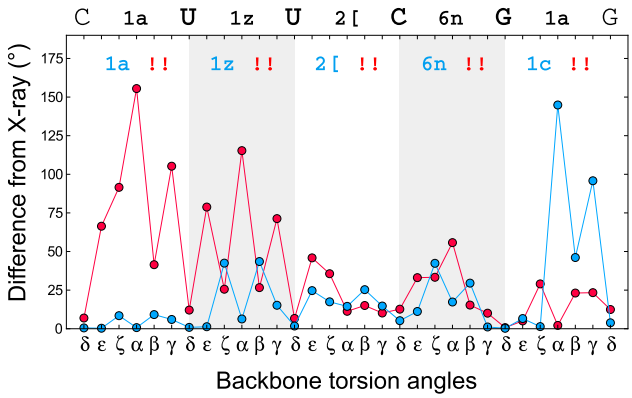


Figure 7. Improvement of the Backbone Conformation of the UUCG Tetraloop in PDB Structure 2LU0

The sequence of the tetraloop, including two flanking residues, is indicated on top. The conformation assigned by the program Suitename to each suite between the corresponding bases is indicated for the TL1 loop of PDB X-ray structure 1F7Y (black), the model 1 of the original NMR bundle 2LU0 (red), and the structure with the lowest experimental energy calculated with the RNA-ff1 force field (blue). "**!!**" denotes an outlier suite; all other labels represent known rotamers (Richardson et al., 2008). For each backbone torsion angle, the plot shows the difference from the X-ray structure for the 2LU0 structure (red) and the RNA-ff1 structure (blue). Background shading delineates the different suites. See also Figures S2–S4.

All structure calculations relied on the Internal Variable Module of Xplor-NIH (Schwieters and Clore, 2001), where the molecules evolved exclusively in torsion angle space, with the exceptions of the sugar-ring C4′–O4′ bond and its associated bond angles, which were restrained by the force field. Calculations with the RNA-ff1 force field additionally allowed angular degrees of freedom to all sugar-ring endocyclic bond angles; the extra flexibility improved structure agreement with the force field parameters (not shown).

Computations based on Xplor-NIH's "old force field" were intended to reflect the state of the art in conventional RNA structure determination by NMR. They relied on the covalent and nonbonded parameters in topology/parameter files nucleic-1.1.top/nucleic-1.1.par, with the torsion angle information omitted, as was provided by the preexisting statistical torsional potential in Xplor-NIH (Clore and Kuszewski, 2003). In addition, the statistical base–base positional potential (Clore and Kuszewski, 2003) was used. The molecular dynamics/simulated annealing procedure used with both the old force field and RNA-ff1 are virtually identical (see Supplemental Experimental Procedures for details), and similar to that previously implemented with proteins (Bermejo et al., 2012). The procedure consists of two sequential protocols for (1) folding an initial extended conformation with satisfied covalent geometry, subject to the torsion angle, distance, and base-pair planarity restraints, and (2) refinement of a selected folded model, with the addition of RDC restraints.

Two hundred structures were computed when using the full RDC dataset, ranked according to increasing experimental energy, keeping the top 20 for further analysis. Under the complete RDC cross-validation scheme (see Results), a total of ten different random RDC test sets were generated; for each case, 100 structures were computed, and ranked according to increasing experimental energy, keeping the top ten for further analysis. Statistics on such cross-validated structures are the average over those obtained from the different RDC sets.

Availability

The force field RNA-ff1, including the statistical torsional term torsionDB$_{RNA}$, is available in Xplor-NIH version 2.41, downloadable from http://nmr.cit.nih.gov/xplor-nih/; an example of its use will be provided in the eginput/rna directory of the distribution package. The tools needed to generate a statistical potential from a torsion angle database will also be made available.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.str.2016.03.007.

AUTHOR CONTRIBUTIONS

Conceptualization: G.A.B., G.M.C., and C.D.S.; Methodology/Software: G.A.B. and C.D.S.; Investigation: G.A.B.; Writing – Original Draft: G.A.B.; Writing – Review & Editing: G.A.B., G.M.C., and C.D.S.

ACKNOWLEDGMENTS

We thank Mijeong Kang and Juli Feigon for kindly providing the Xplor-NIH topology and parameter files associated with the PreQ$_1$ small-molecule ligand in PDB: 2L1V, and Nico Tjandra's group for useful discussions. This work was supported by the NIH Intramural Research Programs of CIT (to G.A.B and C.D.S.), NIDDK (to G.M.C.), NHLBI (through Nico Tjandra) and NCI (through R. Andrew Byrd), and by the AIDS Targeted Antiviral Program of the Office of the Director of the NIH (to G.M.C.).

Received: December 23, 2015
Revised: February 26, 2016
Accepted: March 4, 2016
Published: April 7, 2016

REFERENCES

Bermejo, G.A., Clore, G.M., and Schwieters, C.D. (2012). Smooth statistical torsion angle potential derived from a large conformational database via adaptive kernel density estimation improves the quality of NMR protein structures. Protein Sci. *21*, 1824–1836.

Bertini, I., Cavallaro, G., Luchinat, C., and Poli, I. (2003). A use of Ramachandran potentials in protein solution structure determinations. J. Biomol. NMR *26*, 355–366.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Clore, G.M., and Garrett, D.S. (1999). R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures. J. Am. Chem. Soc. *121*, 9008–9012.

Clore, G.M., and Kuszewski, J. (2002). Chi(1) Rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force. J. Am. Chem. Soc. *124*, 2866–2867.

Clore, G.M., and Kuszewski, J. (2003). Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. J. Am. Chem. Soc. *125*, 1518–1525.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., et al. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. *35*, W375–W383.

Desjardins, G., Bonneau, E., Girard, N., Boisbouvier, J., and Legault, P. (2011). NMR structure of the A730 loop of the Neurospora VS ribozyme: insights into the formation of the active site. Nucleic Acids Res. *39*, 4427–4437.

Donghi, D., Pechlaner, M., Finazzo, C., Knobloch, B., and Sigel, R.K.O. (2013). The structural stabilization of the kappa three-way junction by Mg(II)

represents the first step in the folding of a group II intron. Nucleic Acids Res. *41*, 2489–2504.

Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. (2000). The crystal structure of UUCG tetraloop. J. Mol. Biol. *304*, 35–42.

Gesteland, R.F., Cech, T.R., and Atkins, J.F. (2006). The RNA World, 3rd Edition: The Nature of Modern RNA Suggests a Prebiotic RNA World, *vol. 43* (Cold Spring Harbor Laboratory Press).

Hendrix, D.K., Brenner, S.E., and Holbrook, S.R. (2005). RNA structural motifs: building blocks of a modular biomolecule. Q. Rev. Biophys. *38*, 221–243.

Kang, M., Peterson, R., and Feigon, J. (2009). Structural insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. Mol. Cell *33*, 784–790.

Keating, K.S., and Pyle, A.M. (2010). Semiautomated model building for RNA crystallography using a directed rotameric approach. Proc. Natl. Acad. Sci. USA *107*, 8177–8182.

Kruschel, D., Skilandat, M., and Sigel, R.K.O. (2014). NMR structure of the 5' splice site in the group IIB intron *Sc*.ai5γ—conformational requirements for exon-intron recognition. RNA *20*, 295–307.

Kuszewski, J., and Clore, G.M. (2000). Sources of and solutions to problems in the refinement of protein NMR structures against torsion angle potentials of mean force. J. Magn. Reson. *146*, 249–254.

Kuszewski, J., Gronenborn, A.M., and Clore, G.M. (1996). Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. Protein Sci. *5*, 1067–1080.

Kuszewski, J., Gronenborn, A.M., and Clore, G.M. (1997). Improvements and extensions in the conformational database potential for the refinement of NMR and X-ray structures of proteins and nucleic acids. J. Magn. Reson. *125*, 171–177.

Kuszewski, J., Schwieters, C., and Clore, G.M. (2001). Improving the accuracy of NMR structures of DNA by means of a database potential of mean force describing base-base positional interactions. J. Am. Chem. Soc. *123*, 3903–3918.

Lu, X.J., Bussemaker, H.J., and Olson, W.K. (2015). DSSR: an integrated software tool for dissecting the spatial structure of RNA. Nucleic Acids Res. *43*, e142.

Lu, X.J., and Olson, W.K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res. *31*, 5108–5121.

Lu, X.J., and Olson, W.K. (2008). 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc. *3*, 1213–1227.

Marchanka, A., Simon, B., Althoff-Ospelt, G., and Carlomagno, T. (2015). RNA structure determination by solid-state NMR spectroscopy. Nat. Commun. *6*, 7024.

Mertens, H.D., and Gooley, P.R. (2005). Validating the use of database potentials in protein structure determination by NMR. FEBS Lett. *579*, 5542–5548.

Montelione, G.T., Nilges, M., Bax, A., Güntert, P., Herrmann, T., Richardson, J.S., Schwieters, C.D., Vranken, W.F., Vuister, G.W., Wishart, D.S., et al. (2013). Recommendations of the wwPDB NMR validation task force. Structure *21*, 1563–1570.

Murray, L.J.W., Arendall, W.B., Richardson, D.C., and Richardson, J.S. (2003). RNA backbone is rotameric. Proc. Natl. Acad. Sci. USA *100*, 13904–13909.

Nilges, M., Clore, G.M., and Gronenborn, A.M. (1988). Determination of 3-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. FEBS Lett. *229*, 317–324.

Nozinovic, S., Furtig, B., Jonker, H.R.A., Richter, C., and Schwalbe, H. (2010). High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. Nucleic Acids Res. *38*, 683–694.

Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., Heinemann, U., Lu, X.J., Neidle, S., Shakked, Z., et al. (2001). A standard reference frame for the description of nucleic acid base-pair geometry. J. Mol. Biol. *313*, 229–237.

Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A.T., and Berman, H.M. (1996). New parameters for the refinement of nucleic acid-containing structures. Acta Crystallogr. D Biol. Crystallogr. *52*, 57–64.

Pechlaner, M., Donghi, D., Zelenay, V., and Sigel, R.K.O. (2015). Protonation-dependent base flipping at neutral pH in the catalytic triad of a self-splicing bacterial groupII intron. Angew. Chem. Int. Ed. Engl. *54*, 9687–9690.

Read, R.J., Adams, P.D., Arendall, W.B., Brünger, A.T., Emsley, P., Joosten, R.P., Kleywegt, G.J., Krissinel, E.B., Lutteke, T., Otwinowski, Z., et al. (2011). A new generation of crystallographic validation tools for the protein data bank. Structure *19*, 1395–1412.

Richardson, J.S., Schneider, B., Murray, L.W., Kapral, G.J., Immormino, R.M., Headd, J.J., Richardson, D.C., Ham, D., Hershkovits, E., Williams, L.D., et al. (2008). RNA backbone: consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). RNA *14*, 465–481.

Schwieters, C.D., and Clore, G.M. (2001). Internal coordinates for molecular dynamics and minimization in structure determination and refinement. J. Magn. Reson. *152*, 288–302.

Schwieters, C.D., and Clore, G.M. (2007). A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. Biochemistry *46*, 1152–1166.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. J. Magn. Reson. *160*, 65–73.

Schwieters, C.D., Kuszewski, J.J., and Clore, G.M. (2006). Using Xplor-NIH for NMR molecular structure determination. Prog. Nucl. Magn. Reson. Spectrosc. *48*, 47–62.

Sibille, N., Pardi, A., Simorre, J.P., and Blackledge, M. (2001). Refinement of local and long-range structural order in theophylline-binding RNA using ^{13}C-^{1}H residual dipolar couplings and restrained molecular dynamics. J. Am. Chem. Soc. *123*, 12135–12146.

Word, J.M., Lovell, S.C., LaBean, T.H., Taylor, H.C., Zalis, M.E., Presley, B.K., Richardson, J.S., and Richardson, D.C. (1999a). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J. Mol. Biol. *285*, 1711–1733.

Word, J.M., Lovell, S.C., Richardson, J.S., and Richardson, D.C. (1999b). Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. *285*, 1735–1747.

Yang, J.S., Kim, J.H., Oh, S., Han, G., Lee, S., and Lee, J. (2012). STAP Refinement of the NMR database: a database of 2405 refined solution NMR structures. Nucleic Acids Res. *40*, D525–D530.

Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P., and Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. Nat. Struct. Biol. *4*, 644–649.