JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

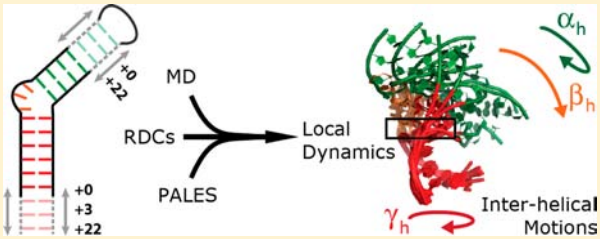
A General Method for Constructing Atomic-Resolution RNA Ensembles using NMR Residual Dipolar Couplings: The Basis for Interhelical Motions Revealed

Loïc Salmon,[†] Gavin Bascom,[‡] Ioan Andricioaei,[‡] and Hashim M. Al-Hashimi*[,†]

[†]Department of Chemistry and Biophysics, University of Michigan, Ann Arbor, Michigan 48109, United States
[‡]Department of Chemistry, University of California, Irvine, California 92697, United States

Supporting Information

ABSTRACT: The ability to modulate alignment and measure multiple independent sets of NMR residual dipolar couplings (RDCs) has made it possible to characterize internal motions in proteins at atomic resolution and with time scale sensitivity ranging from picoseconds up to milliseconds. The application of such methods to the study of RNA dynamics, however, remains fundamentally limited by the inability to modulate alignment and by strong couplings between internal and overall motions that complicate the quantitative interpretation of RDCs. Here, we address this problem by showing that RNA alignment can be generally modulated, in a controlled manner, by variable elongation of A-form helices and that the information contained within the measured RDCs can be extracted even in the presence of strong couplings between motions and overall alignment via structure-based prediction of alignment. Using this approach, four RDC data sets, and a broad conformational pool obtained from a 8.2 μs molecular dynamics simulation, we successfully construct and validate an atomic resolution ensemble of human immunodeficiency virus type I transactivation response element RNA. This ensemble reveals local motions in and around the bulge involving changes in stacking and hydrogen-bonding interactions, which are undetectable by traditional spin relaxation and drive global changes in interhelical orientation. This new approach broadens the scope of using RDCs in characterizing the dynamics of nucleic acids.



INTRODUCTION

The growing importance of RNA conformational changes in gene expression and regulation has spurred great interest in moving beyond static structures of RNA and toward dynamic ensembles describing RNA flexibility at atomic resolution.[1−5] A major challenge in determining structural ensembles is that the number of experimental measurements that can be made typically pales in comparison to the number of parameters needed to specify the structure and population weights of all conformers populating the free-energy landscape. The ability to modulate partial alignment of proteins[6,7] in solution nuclear magnetic resonance (NMR) spectroscopy studies through dissolution into different ordering media[7−11] has made it possible to measure up to five independent sets of residual dipolar couplings (RDCs), providing a rich source of information for guiding the construction and validation of atomic resolution ensembles.[10,12−16] RDCs provide information regarding the orientational distribution of bond vectors relative to the overall alignment frame and are particularly attractive measurements for constructing dynamic ensembles, given their broad sensitivity to motions occurring on submillisecond time scales.[17,18] In contrast to proteins, the uniform charge distribution in nucleic acids results in similar electrostatic and steric alignment forces, making it difficult if not impossible to modulate alignment by simply changing the ordering medium.[19−21] This, combined with the coupling between an RNA's interhelical motions and its overall alignment,[22−24] limits the applicability of RDCs in studies of nucleic acid dynamics.

We previously introduced an approach for both decoupling internal and overall motions in RNA and modulating alignment that relies on extensive elongation of target helices by ∼22 base pairs.[23,25] The resulting overall alignment of elongated RNAs is nearly axially symmetric and to a good approximation, independent of other motions, takes place in other parts of the molecule. Using two sets of RDCs measured in such elongated RNA samples, we previously reported an ensemble of the transactivation response element (TAR) RNA (Figure 1A) from the human immunodeficiency virus type I (HIV-1).[26,27] Internal motions in TAR have previously been shown to play important roles in adaptive protein recognition.[28,29] TAR is also an excellent model system for exploring the basic dynamic properties of RNA and has been extensively characterized using a wide variety of techniques.[23,25,30−33] The TAR ensemble was constructed by using the RDCs to guide selection[34] of conformers from a pool generated using an 80 ns molecular dynamics (MD) trajectory of TAR. Although this provides a

Received: January 30, 2013
Published: March 8, 2013

   dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

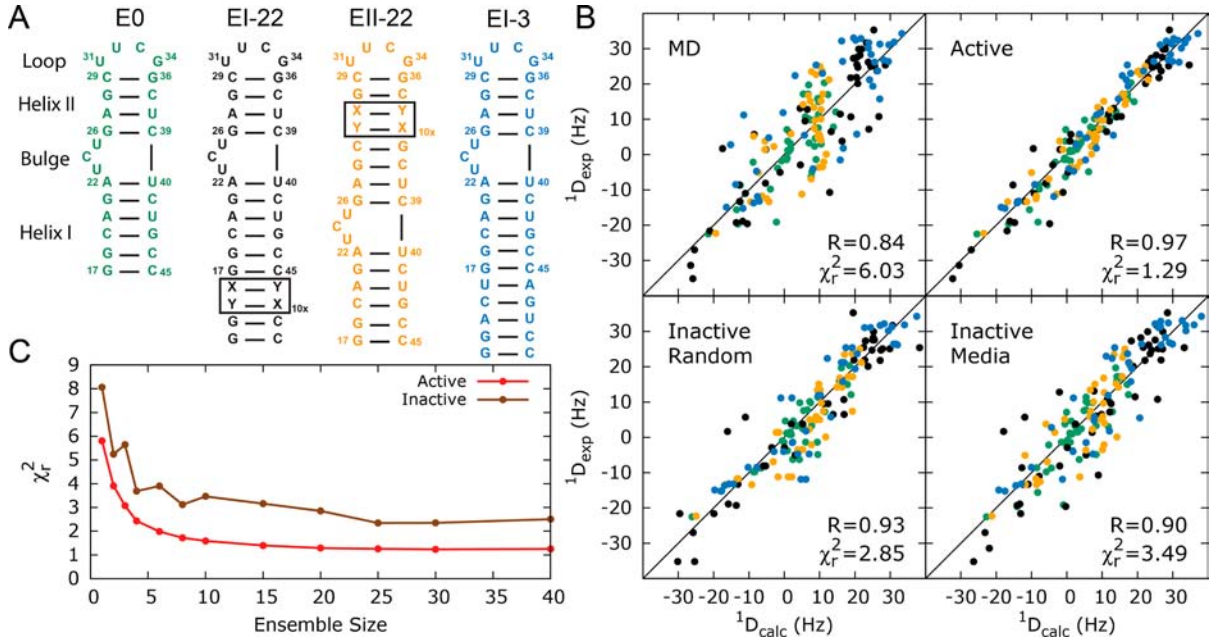


Figure 1. Constructing HIV-1 TAR ensemble using variable elongation RDCs and structure-based prediction of alignment. (A) Secondary structures of the four differentially elongated TAR constructs used to measure multiple sets of RDCs. (B) Comparison of RDCs measured in the four TAR constructs [color coded according to (A)] and those predicted for the MD trajectory (MD) and an ensemble of 20 conformers in which all RDCs are used in the selection process (active). Also shown is how well the RDC-selected ensembles reproduce a subset of RDCs that were not used in the selection process including randomly omitted RDCs (inactive random) and when omitting each of the four RDC data sets (inactive media). (C) Comparison of measured and predicted RDCs in terms of reduced χ^2 as a function of the size of the selected ensembles (N) when using all the RDCs in the ensemble selection (red) and for a subset of randomly chosen RDCs that are excluded from the selection process (brown).

powerful approach for constructing RNA ensembles, the requirement for extensive elongation in order to decouple internal and overall motions generally limits the number of independent alignments that can be attained typically to no more than two. This not only severely limits the achievable spatial resolution with which the ensemble can be constructed but also makes it impossible to rigorously assess the validity of the ensemble using cross-validation. Moreover, many RNA systems cannot tolerate extensive elongation because either it can affect structural properties or it can lead to prohibitively poor relaxation properties.

We previously showed that variable elongation of a RNA terminal helix by as little as three base pairs provides a convenient approach for controllably modulating overall alignment of the molecule without affecting the intrinsic structural and dynamic properties of the RNA target[23,35] and that incremental elongation of a target helix leads to an overall alignment that progressively approaches axial symmetry with its principal direction oriented along the elongated helix.[23,35] While variable helix elongation provides a practical, general, and robust route for modulating RNA alignment, the quantitative interpretation of RDCs measured in partially elongated RNAs proves difficult owing to couplings between internal and overall motions.[23,35] Here, we introduce a new strategy that enables the use of RDCs measured in variably elongated RNAs for constructing ensembles. In this approach, we predict RDCs for a given RNA conformer based on its overall shape using the program PALES.[36,37] Prior studies have shown that PALES[23,35,38−40] and other structure-based approaches[41] can accurately reproduce the overall orientation and asymmetry of the alignment of nucleic acids dissolved in Pf1 phage when assuming a simple steric model. With this approach, we have successfully constructed and validated an atomic-resolution ensemble of HIV-1 TAR, based on four sets of RDCs.

RESULTS AND DISCUSSION

Constructing HIV-1 TAR Ensemble. We analyzed four different sets of directly bounded C−H, C−C, and N−H RDCs spanning bulge and helical residues measured on four variably elongated HIV-1 TAR constructs: a native nonelongated construct[24] (E0), two constructs in which either helix I (EI-22) or helix II (EII-22) is elongated by 22 base pairs,[23] and one construct (EI-3) in which helix I is elongated by only three base pairs[35] (Figure 1A). Previous analysis[35] of the RDC data sets measured on these four constructs using Pf1 phage[42,43] strongly suggests that they carry independent information (Figure S1). To more quantitatively evaluate the independence of RDC data sets, we determined the alignment tensor for helix I by fitting RDCs measured in nonterminal base pairs to a canonical idealized A-form helix, as described previously[35] (Table S1). We then computed the normalized scalar product between pairs of alignment tensors,[44] which magnitude varies between 0 and 1 from orthogonal to coaxial alignment tensors. The six scalar products (0.092, 0.351, 0.586, 0.795, 0.823, and 0.840, Table S2) suggest the presence of independent information in the different data sets, with the short constructs E0 and EI-3 providing less independent information than the elongated constructs EI-22 and EII-22.

To construct an HIV-1 TAR ensemble, we obtained a broad conformational pool for nonelongated HIV-1 TAR by computing a 8.2 μs MD trajectory computed on the Anton supercomputer[45] using the CHARMM36 force field.[46−48] We first examined the ability of the MD trajectory to reproduce the RDC data. To compute RDCs for each of the four TAR constructs, each of the 10 000 nonelongated TAR snapshots in

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

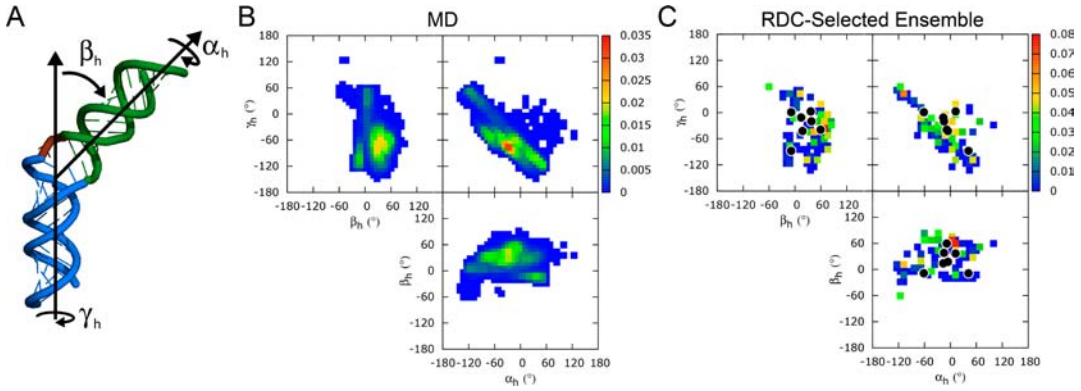


Figure 2. Interhelical dynamics. (A) Definition of the interhelical Euler angles (α_h, β_h, γ_h) used to specify the relative orientation of two A-form RNA helices. The interhelical Euler angle distribution for HIV-1 TAR observed in (B) 8.2 μs MD trajectory and (C) RDC-selected ensemble obtained from combining 100 rounds of $N = 20$ selections. The population of a given interhelical orientation is color coded, increasing from blue to red. Black circles correspond to seven distinct ligand-bound HIV-1 TAR structures (pdb codes: 1QD3, 1UUI, 1UTS, 1UUD, 1ARJ, 1LVJ, and 397D).

the MD pool were variably elongated *in silico* as needed (Methods) and submitted to PALES for structure-based predication of alignment and RDCs. The predicted RDCs were then averaged over all snapshots, and their value uniformly scaled for a given type of TAR construct to optimize the agreement with the measured RDCs[49] (Methods). Note that the MD simulation was carried out on a TAR construct containing a CUGGGA hexanucleotide apical loop that differs form the UUCG tetra-loop used to measure the RDCs (Figure 1A). However, we previously showed that this apical loop substitution has little to no effect on the structural and dynamic properties of the TAR helices and the bulge.[50] Furthemore, simulations show that the two different loops affected the PALES predicted RDCs by an amount smaller than the measurement uncertainty (<4 Hz), as expected given their similar size and overall shape. Interestingly, the agreement between the measured RDCs and values predicted for the long MD trajectory (χ^2_r = 6.03, RMSD = 8.6 Hz, Pearson's correlation coefficient R = 0.84) has improved significantly as compared to the agreement observed for a shorter 80 ns trajectory[51] (χ^2_r = 10.11, RMSD = 11.3 Hzk, R = 0.73) employing the CHARMM27 force field. However the RMSD remains significantly larger than the RDC measurement uncertainty, estimated to range between 2 and 4 Hz for the different TAR constructs[51] (Figure 1B).

The inability of the MD simulation to accurately reproduce the measured RDCs may arise from improper weighting of the different conformers rather than incomplete sampling of the conformational space. We therefore examined if the RDCs could be used to guide selection of subensembles from the MD conformational pool. Here, the sample and select (SAS) Monte Carlo selection scheme was used to minimize a reduced χ^2 function assessing the agreement between measured and predicted RDCs for a given selected ensemble.[23,51] Sub-ensembles with increasing size (N) were then constructed to find the smallest ensemble satisfying all of the measured RDCs.

As shown in Figure 1C, the agreement between measured and predicted RDCs improves significantly with increasing ensemble size and reaches a plateau at $N \sim 20$ (χ^2_r = 1.29, RMSD = 4.0 Hz, R = 0.97). A similar trend is observed when comparing the agreement with RDCs that are left out and not actively used in the ensemble selection (Figure 1C). Importantly, the quality of the RDC fits deteriorated considerably (χ^2_r = 2.55, RMSD = 5.8 Hz, R = 0.93) when constructing ensembles using a conformational pool obtained from a shorter 80 ns MD trajectory (data not shown), underscoring the importance of having broad conformational sampling for highly flexible RNAs.

Using this SAS approach, we were able to construct ensembles that simultaneously satisfy RDCs measured in all four TAR constructs close to experimental precision (Figure 1C), as evaluated by direct comparison of the measured and predicted RDCs and by cross-validation and comparison of how well the selected ensembles predict RDC data that was not used in the selection, including very demanding validations in which each of the four RDC data sets (~25% of the data) was removed entirely. The agreement observed for the cross-validations (χ^2_r = 2.85, RMSD = 5.9 Hz, R = 0.93 for randomly removed RDCs and χ^2_r = 3.49, RMSD = 6.7 Hz, R = 0.90 for the successive removal of each data set) represents a significant improvement compared to a randomly selected ensemble of the same size (χ^2_r = 6.47, RMSD = 8.8 Hz, R = 0.83). This self-consistency suggests that PALES accurately reproduces the overall alignment of the RNA and provides additional support that variable elongation does not significantly affect the properties of TAR, as also inferred independently based on comparison of chemical shifts.[25,50]

While the ensemble is able to back-predict correctly unused RDC data, this could arise in part due to partial correlations between experimental data set (Supporting Information, SI). Therefore, we further quantitatively tested the accuracy of the ensemble by examining its ability to reproduce data of other physical nature that were completely ignored in the ensemble construction, namely magnetic field-induced RDCs and ^{1}H chemical shifts. First, the ensemble reproduces small magnetic field-induced RDCs (ranging in size between −2.1 to 2.6 Hz) measured in a nonelongated TAR at 18.8T (800 MHz ^{1}H Larmor frequency) in the absence of ordering media to within almost experimental precision (reproduction of the experimental data with RMSD = 0.59 Hz compared to experimental error ~0.50 Hz)[22,52] (Figure S2). Much poorer agreement is observed for the MD starting pool (RMSD = 0.90 Hz). As a second independent test, we examined how well the ensemble reproduces ^{1}H experimental chemical shifts measured at the bulge using the program NUCHEMICS[53,54] to compute ^{1}H chemical shifts based on structure. The RDC ensemble reproduces the ^{1}H chemical shifts with RMSD of 0.17 ppm as compared to 0.20 ppm for the entire MD trajectory

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

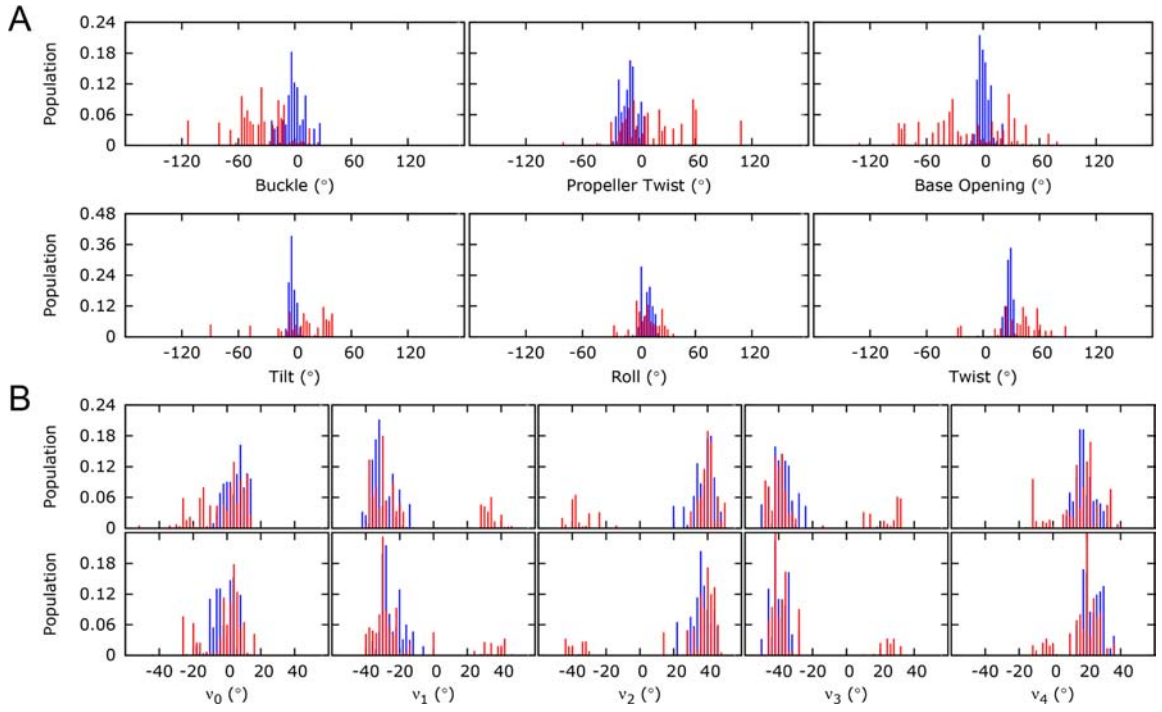


Figure 3. Local dynamics within A-form helices. For the RDC-selected ensemble: (A) distribution of intra- and interbase pair (buckle, propeller twist, and opening and tilt, roll, and twist) angular parameters for the junction A22-U40 (red) and central A20-U42 (blue) base pairs and (B) five sugar torsion angles (ν_0–ν_4) for the same nucleotides (top) A22 (red) and A20 (blue) and (bottom) U40 (red) and U42 (blue).

(Methods). While this improvement is small, it does bring the agreement closer to the reported NUCHEMICS ^{1}H chemical shift error prediction (~0.16 ppm). As we discuss in subsequent sections, the RDC TAR ensemble agrees with other independent information available on this RNA.

To further evaluate how well the RDC constructed ensemble actually reproduces key conformational features of HIV-1 TAR, we preformed extensive simulations employing synthetic RDCs and a variety of target ensembles (SI). These simulations establish the ability to extract accurate information regarding the orientation distribution of base pair and interhelical parameters using the four RDC data sets measured in TAR (Table S3, Figures S3–S5). These well-defined angular distributions will be the focus of further analysis in the following sections.

Interhelical Motions. A general and functionally important motion in RNA involves collective changes in the orientation of A-form helices across bulges and junctions. Such motions have been shown to play important roles in adaptive protein and ligand recognition, in the catalytic cycles of ribozymes, and in the assembly of ribonucleoprotein machines.[1,55] TAR provides an excellent model system to explore interhelical motional modes across its three-residue bulge. The relative reorientation of two helices (Figure 2A) can be defined using three interhelical Euler angles (α_h, β_h, γ_h)[56,57] that specify the twist angles about the two helices (α_h and γ_h) and an interhelical bend angle (β_h).

Figure 2 compares the RDC-selected interhelical ensemble with that of the MD generated pool. Interestingly, we find that many of the RDC-selected conformers fall in low populated regions of the MD trajectory. This includes conformers with large bend angles ($\beta_h > 50°$) and relatively small twist angles ($|\alpha_h| < 30°$ and $|\gamma_h| < 30°$). These conformers also tend to have a distorted geometry for the junction A22-U40 base pair, which is known to be locally flexible.[58] Overall the RDC-selected ensemble features both a larger mean bend angle and larger variations about this mean ($|\beta_h| = 52 \pm 27°$ for the selected ensemble compared to $32 \pm 18°$ in the MD simulation). The twist angles around the two helices for the RDC-selected ensemble are more similar to the starting MD pool but feature on average ~30° shift in γ_h (α_h and $\gamma_h = -15 \pm 47°$ and $-28 \pm 44°$, respectively, for the selected ensemble compared to $-22 \pm 42°$ and $-57 \pm 40°$ for MD), leading to a smaller degree of interhelical overtwisting as compared to the MD pool[56] (interhelical twist $\zeta_h = \alpha_h + \gamma_h$ being $-43 \pm 36°$ and $-79 \pm 23°$ for the selected ensemble and the starting pool, respectively). Thus improper weighting of the MD trajectory in the (α_h, β_h, γ_h) can partly explain its inability to reproduce experimental RDCs. This could be due to a lack of convergence, given that the MD trajectory remains much shorter than the RDC time scale sensitivity (<milliseconds), improper parametrization of the force fields, or inadequate modeling of interactions with monovalent ions since it is known that increasing the concentration of monovalent or divalent ions simultaneously decreases the average bend angle and the amplitude of bending motions in TAR.[59] In the MD simulation, ions were simply placed to ensure electroneutrality, which may not necessarily reflect experimental conditions.

The RDC-selected ensemble reproduces many salient features of TAR interhelical dynamics that have been characterized previously[23,51] (Figures S6 and S7) including the presence of very large amplitude bending and twisting motions ($|\beta_h|$ varies over the range of 3° to 91°, and the twist angles α_h and γ_h vary between −123° and 101° and −127° and 64°, respectively). The ensemble also shows strong correlations between the interhelical twist angles α_h and γ_h which have been attributed to steric and connectivity constraints imposed by the TAR bulge.[57,60] Interestingly, the selected ensemble comes

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457–5466

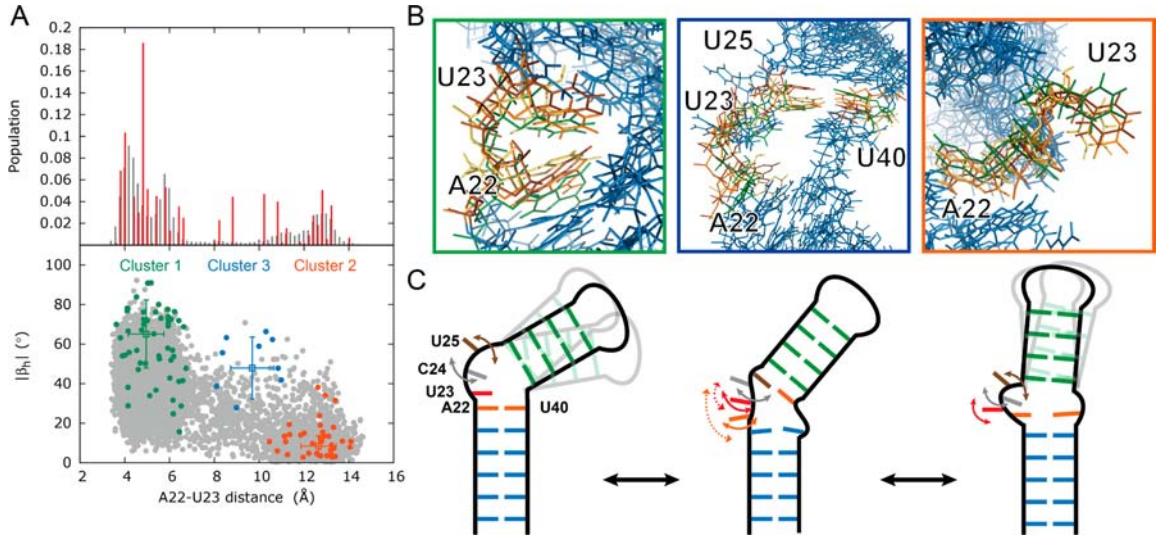


Figure 4. Relationship between local and interhelical dynamics. (A) Clustering of the selected ensemble. Distribution of the A22-U23 distance (top) and its correlation between the $|\beta_h|$ interhelical angle (bottom). Results are shown for the MD trajectory (gray), full selected ensemble, obtained from combining 100 rounds of $N = 20$ selections (red). Shown are clusters 1 (green), 2 (orange) and 3 (blue) along with their population-weighted averages and standard deviations. (B) Conformational properties of the three clusters presented using a subensemble of the selected conformers: clusters 1 (left), 2 (right), and 3 (middle). (C) Schematic representation of the three clusters and a proposed ordering for transitioning between conformations with different bend angles. Curved arrows indicate local dynamics. Interactions with helix II are indicated with a dashed line.

closer to the sampling of known ligand bound structures of TAR as compared to a previous ensemble constructed using only RDCs measured in EI-22 and EII-22 TAR (Figure S8), reinforcing the idea that adaptive recognition could in principle occur via conformational selection.[23,51]

Local Motions. The validated multialignment RDC ensemble of TAR afforded a unique opportunity to examine with exquisite detail local motions that are more challenging to characterize given that they involve many more degrees of freedom. Specifically, we investigated the angles defining the intra- and interbase pair geometry[61] (buckle, propeller twist, and base opening and tilt, roll, and twist) and sugar torsion angles (ν_0, ν_1, ν_2, ν_3, and ν_4). The distribution of those quantities for the base pairs G18-C44, C19-G43, A20-U42, G21-C41, A27-U38, and G28-C37 indicates that the central base pairs of the two helices remain statistically very close to an idealized A-form helix and exhibit variations usually comparable or slightly bigger to those expected based on a statistical survey of X-ray structures (Tables S4−S6). This is consistent with other studies showing excellent agreement between measured RDCs and those predicted using an idealized A-form geometry containing Watson−Crick base pairs[19] and recent chemical shift-based analysis of A-form helices.[62] This also validates *a posteriori* the use of idealized A-form helices to elongate conformers during the PALES analysis (Methods).

Interestingly, the two base pairs around the bulge show asymmetric behavior with G26-C39 adopting an A-from like conformation and A22-U40 adopting a much broader conformational distribution, deviating from a classical Watson−Crick base pairing (Figures 3 and S9). This is consistent with prior NMR studies employing *trans*-hydrogen-bonding scalar coupling, ^{13}C and ^{15}N spin relaxation, and RDC measurements.[51,58,63] The observed conformational flexibility in the A22-U40 base pair is not only limited to the two bases but also includes the sugars, which sample both C2′- and C3′-endo conformations. An high degree of flexibility is also observed for bulge residues, particularly C24 and U25 and to a lesser extent U23, which frequently stacks on A22, consistent with previous nuclear Overhauser effect (NOE), spin relaxation, RDC, and MD studies of TAR.[51,63,64]

To gain further insights into the local dynamics at the bulge linker, we examined the stacking interactions between bulge and neighboring residues. Previous studies showed that the interhelical hinge is primarily defined by bulge residue U23 and the lower residue A22, which forms a flexible base pair with U40.[51,63] In free TAR, U23 stacks on A22 and promotes interhelical bending. However, in coaxial TAR structures observed under high ionic strength conditions, or when bound to ligands, U23 is flipped out.[59,65,66] We observe a unique multimodal A22-U23 distance distribution in the RDC-selected ensemble (Figure 4A), whereas other distances exhibit significantly flatter distributions (Figure S10), a trend clearly emphasized by the selection procedure. Relative to MD, the RDC-selected ensemble significantly increases the population of unstacked C24-U25 and U25-G26 conformations, indicating a potential bias in the force field toward stacked conformations at these sites. The observation of an higher population of nearby conformations for A22-U23 is consistent with the observation of NOEs between these two bases.[64]

Interplay between Local and Collective Motions. Despite many studies on interhelical motions in RNA, the local hinge motions at interhelical junction that activate these dynamics remain poorly understood. To gain insights into the interplay between the TAR bulge local conformation and interhelical orientation, we examined how the A22-U23 distance distribution correlates with the interhelical bend angle $|\beta_h|$ (Figure 4A). This analysis revealed three distinct clusters: cluster 1 (population 66%) features highly bent conformations in which A22 and U23 are in close proximity and adopt a looped in stacked conformation; cluster 2 (population 19%) features smaller bend angles and conformers in which A22 and U23 are far apart with U23 adopting a flipped out conformational; and finally cluster 3 (population 15%) features intermediate bend angles and distorted bulge

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

conformations, where A22 may interact with the upper helix or where the deformation may propagate down to the G21-C41 base pair (Figure 4A,B). In this cluster U40 tends to form an unusual base pair with U25 rather than A22, whereas U23 and C24 remain unpaired and flexible. Interestingly, this alternative secondary structure is predicted to be the second most energetically favorable bulge secondary structure for HIV-1 TAR using the structure prediction program MC-fold.[67,68]

Together, the above observations suggest a mechanism for interhelical bending that is tightly coupled to the local conformation of the bulge and specifically the conformation of U23, which may serve to modulate the steric constraints imposed on certain interhelical orientations, with linear coaxial structures becoming sterically more accessible when U23 is flipped out (Figure 4C). Cluster 3 suggests that the transition between those two substates might proceed through the breaking of the flexible A22-U40 base pair, possibly accompanied by local changes in secondary structure around the bulge, including formation of a noncanonical U25-U40 base pair and possibly other interactions between A22 or U23 and the upper helix, as previously documented in ligand bound states of TAR, e.g., in the TAR argininamide complex, where U23 forms a base triple with A27-U38 in the upper helix.[58] Strong support for such a pathway for interhelical motions comes from previous studies showing that both the average interhelical bend angle and amplitude of interhelical motions decrease dramatically when replacing A22-U40 with a stronger G22-C40 base pair.[69]

Time Scale of Motions: Reconciling RDC and Spin Relaxation Data. Previous studies showed that the amplitudes of interhelical motions determined by RDCs[23] (ϑ_{int} = 0.45 ± 0.05, where ϑ_{int}, ranges between 0 and 1 for maximum to minimum amplitude interhelical motions) greatly exceed those derived by spin relaxation[25] (S_s = 0.86 ± 0.02). This was attributed to slower micro-to-millisecond motions occurring at time scales longer than overall rotational diffusion of the molecule[25] (~19 ns for EI-22-TAR) which are sensed by RDCs but not spin relaxation. Though RDCs provide valuable information about the amplitude and direction of motions, they do not provide direct information regarding motional time scales. Therefore the comparison with spin relaxation data provides a simple way to classify dynamics into fast and slow motion compared to the molecular rotational diffusion time, as documented previously in studies of protein dynamics.[6,10,13−15,70−74] However, the nature of the slower motions sensed by RDCs and not spin relaxation has remained elusive for nucleic acids. To gain further insights into the nature of these slower motions, we used the RDC ensemble to compute site-specific order parameters[75−77] describing the amplitude of motions at C−H and N−H bonds and compared values with order parameters determined by ^{13}C and ^{15}N spin relaxation data[63,78] (Figure 5 and Methods).

Similar RDC and spin relaxation order parameters are observed for residues in the reference elongated helix I as well as for the highly flexible bulge residues C24 and U25. Thus, these sites appear to experience insignificant motions at micro-to-millisecond time scales. By contrast, the RDC-derived order parameters are significantly smaller than their spin relaxation counterparts in the upper helix as well as for bulge residue U23 and the junction A22-U40 base pair, suggesting the presence of significant micro-to-millisecond dynamics at these sites. The lower RDC-derived order parameters for the upper helix can be attributed to excess micro-to-millisecond interhelical motions

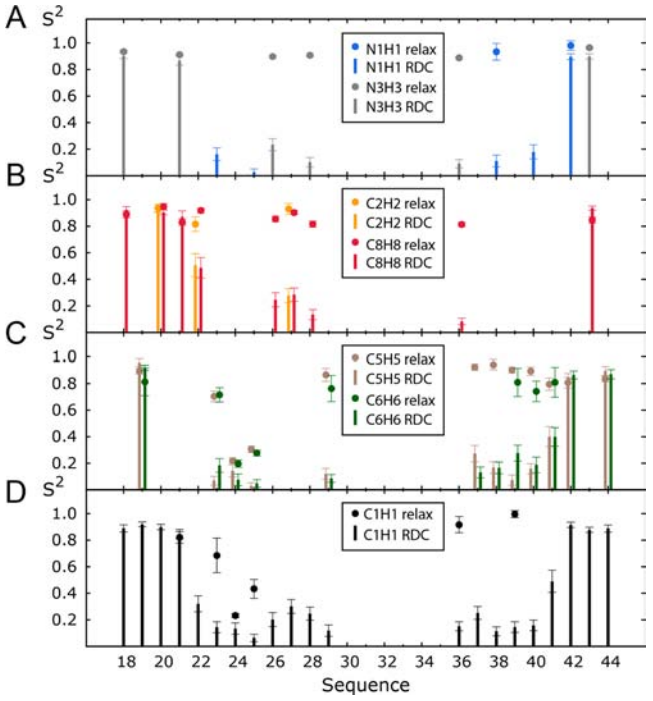


Figure 5. Resolving motional modes occurring at different time scales. Comparison of order parameters (S^2) describing motions of individual bond vectors in TAR obtained from the RDC-selected ensemble (bars) and based on analysis of ^{13}C and ^{15}N spin relaxation data (circles) measured in EI-22 HIV1-TAR. RDC-based order parameters (bars) are obtained by aligning all the conformers of the selected ensemble on an idealized helix I. Error bars represent experimental uncertainty. Color coding: (A) N1H1 (blue) and N3H3 (gray), (B) C2H2 (orange) and C8H8 (red), (C) C5H5 (brown) and C6H6 (green), and (D) C1H1 (black).

that yield a net ϑ_{int} = 0.46 ± 0.04 (Methods), in good agreement with values obtained previously using an order tensor analysis of RDCs.[23] Interestingly, the much lower RDC-derived order parameters for U23 and the adjacent A22-U40 base pair, coincides with lower order parameters for interhelical motions and may reflect slower local motions at the junction that lead to the breaking of the A22-U40 base pair or loss of A22-U23 stacking interactions, which results in transitions between bent and coaxial interhelical conformations. The need to break favorable base paring, and stacking interactions can explain why such motions occur at slower time scales inaccessible by spin relaxation. The absence of micro-to-millisecond exchange broadening at these sites using $R_{1\rho}$ relaxation−dispersion experiments[50] suggests that the motions occur on the nano-to-microsecond time scales, a window which is invisible to spin relaxation methods.

The picture that emerges is one in which the helices 'rattle' about a predefined bulge conformation at pico-to-nanosecond time scales but that transitions between different conformational substates involving highly bent or linear interhelical conformations occurs at slower nano-to-microsecond time scales through pathways that require disruption of favorable stacking and hydrogen-bonding interactions in and around the bulge. This is consistent with recent NMR relaxation dispersion data showing that even stable Watson−Crick base pairs in and around bulges, internal loops, and apical loops can undergo slow changes in stacking and hydrogen-bonding interactions.[50,68]

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

CONCLUSION

In conclusion we have developed a new method for constructing, at high resolution, dynamic ensembles of nucleic acids. The combination of long MD simulations with elongation modulated RDCs allowed us to probe elusive nucleic acid motions at atomic resolution and to explore the interplay between local and global motions. Our results expose complex local motions occurring at different time scales, including pico-to-nanosecond motions of bulged-out residues and slower nano-to-microsecond motions involving the disruption of stacking and hydrogen bonding and the flipping in and out of residues, all concentrated within the bulge two way junction and the immediately neighboring base pairs. These local motions provide the molecular basis for larger-amplitude collective motions of juxtaposed helices. The ability of the approach to bypass the coupling between internal dynamics and global reorientation and to accommodate any level of elongation makes it readily applicable to a large variety of nucleic acid systems, providing a basis for exploring the dynamic properties of diverse structural motifs in nucleic acids at atomic resolution. The approach can easily be extended to accommodate other sources of experimental information regarding the ensemble,[55] including small angle X-ray scattering (SAXS),[79,80] which can provide complementary distance-based information and chemical shifts[53,54,81] reporting on local conformational environment.

METHODS

RDC Data. The TAR ensemble was constructed using four phage-induced RDC data sets previously measured on four variably elongated HIV-1 TAR constructs in which the wild-type apical loop is replaced by a UUCG apical loop: E0,[24] EI-22,[23] EII-22,[23] and EI-3[35] (Figure 1A). All RDCs were measured at 298 K using identical buffer conditions (15 mM sodium phosphate, 25 mM sodium chloride, 0.1 mM EDTA, and pH ∼ 6.4). The TAR concentration ranged between ∼0.6−1.2 mM and that of Pf1 phage between 6 and 22 mg/mL depending on the level of elongation of the construct. Previous studies have shown that this apical loop mutation has little to no effect on the bulge and interhelical dynamics.[50] The data included sugar (C1′−H1′, C2′−H2′, C3′−H3′, and C4′−H4′) and base (C2−H2, C5−H5, C6−H6, C8−H8, C5−C6, N1−H1, and N3−H3) one-bond RDCs measured in the two helices and the bulge. A total of 44, 45, 36, and 38 RDCs were measured in E0, EI-22, EII-22, and EI-3 TAR, respectively.

Molecular Dynamics Simulations. The starting pool for selection consisted of 10 000 snapshots obtained by regular sampling of a 8.2 μs MD simulation run on the supercomputer Anton; a special-purpose machine built specifically for highly efficient computation of accurate molecular dynamics trajectories.[45] Computational hours were obtained on Anton through the National Resource for Biomedical Supercomputing (NRBSC) and the Pittsburgh Supercomputing Center donated by D.E. Shaw Research. Initial coordinates for HIV-1 TAR were obtained from the Protein Data Bank (access code 1ANR). The system was solvated in VMD[82] with TIP3 water and 27 Na^+ ions to neutralize the overall electrical charge in a 64 × 64 × 64 Å cube for initial heating, which was carried out using the CHARMM molecular dynamics package.[83,84] CHARMM36 force field parameters for ribonucleic acids were used, which included recent changes made in 2011 by Mackerell et al. to the 2′ hydroxyl dihedral parameters.[46−48] The system was heated to 300 K with harmonic constraints on backbone atoms for 100 ps, at which point restraints were gradually released over another 100 ps, and the system was equilibrated for 5 ns.

Velocities, coordinates, system, and force field parameters were then all transferred from the initial heating run to Anton style formats, and the simulation was extended on Anton for 8.2 μs. The Nose-Hoover NVT integrator with a time step of 2 fs was used, and coordinates were saved every 820 ps, yielding 10 000 conformational snapshots. According to the standard Reference System Propagator Algorithm (RESPA) near bonded and nonbonded forces were computed every time step, while far nonbonded forces were computed every third time step. Standard periodic boundary conditions were applied, with long-range interactions calculated according to the particle mesh Ewald summation[85] with cutoff parameter of 12.99 Å. In addition to convergence of RMSD and energy, the trajectory was found to be predominantly in A-form by pucker angles and interhelical distances (data not shown).

RDC Calculation. Since the MD simulation was preformed on nonelongated E0-TAR, an elongation procedure was used to elongate, as needed, each snapshot prior to PALES structure-based calculation of alignment and RDCs. For each snapshot, the elongated helix was aligned onto an idealized elongated A-form helix that contains the sequences of the TAR helices by superimposing heavy atoms of the RNA backbone[78] (base pairs G18-C44 to G21-C41 for helix I and base pairs G26-C39 to G28-C37 for helix II). The required number of base pairs for elongation were then added to the MD snapshot. In the case of EII-22, the apical loop from the MD snapshot was translated to the end of helix II to conserve its global structural features. As the experimental RDCs were measured on a UUCG mutant of HIV-1 TAR, the wild-type loop of the snapshots could have been substitute by a structural model of the UUCG loop.[86] However, due to the absence of good dynamical description of this loop, the small effect expected from previous studies[50] and the absence of improvement in the RDC reproduction after such substitution, the wild type loop was kept to avoid unnecessary complication of the protocol. RDCs were then calculated for each conformer using PALES,[36,37] using a pure steric description according to the cylindrical wall model with a low effective concentration (0.022g/mL).

The averaging over an ensemble assumes equiprobable conformations and a given conformation cannot be selected more than once. As the magnitude of alignment is dependent on experimental conditions, including the concentration of Pf1 phage and nature of elongation, an overall scaling factor is allowed to float for each of the four RDC sets as described previously.[49] Therefore the RDCs are expressed as

$$D_{i,j}^{\mathrm{calc}} = \frac{\lambda_j}{N}\sum_{k=1}^{N} D_{i,j}^{k} \tag{1}$$

where k runs over the N conformers in the selected ensemble, λ_j represents an overall scaling factor for j^{th} TAR construct, and $D_{i,j}$ is the i^{th} coupling in the j^{th} construct.

Sample and Select. TAR conformations were selected in generating ensembles using the standard Monte Carlo based selection approach that minimizes a χ^2 function representing the quality of the data reproduction:[34]

$$\chi^2 = \sum_{i,j} \frac{(D_{i,j}^{\mathrm{calc}} - D_{i,j}^{\mathrm{exp}})^2}{\delta_{i,j}^2} \tag{2}$$

where i runs over all the RDCs measured for the different constructs j and δ is the weight used to normalize different RDCs (C−H, C−C, and N−H), and was fixed for each construct at one-tenth of the range of RDCs clustered in $^1D_{NH}$, $^1D_{CC}$, aromatic and nonaromatic $^1D_{CH}$ RDCs.

The process of selection starts with the generation of a random ensemble of size N from the conformational pool of 10 000 snapshots. At each evolution step a new ensemble is created by randomly replacing one conformer by another in the pool. If the fitness of the new ensemble is lower than the old ensemble, then the newest ensemble is kept, otherwise it is accepted only with a probability:

$$P = \exp\left(\frac{\chi_{\mathrm{old}}^2 - \chi_{\mathrm{new}}^2}{T}\right) \tag{3}$$

where T is an effective temperature that starts at 100 and decreasing by a factor of 0.9 every 5×10^5 steps. Each selection is composed in total of 5×10^7 steps.

dx.doi.org/10.1021/ja400920w | J. Am. Chem. Soc. 2013, 135, 5457−5466

The selection procedure was repeated 10 times, and the ensemble that best fits the experimental data among the 10 trials is kept and presented in terms of reduced χ^2:

$$\chi_r^2 = \frac{1}{K}\chi^2 \quad (4)$$

where K is the total number of RDCs involved in the selection process.

Cross-Validation. New RDC data sets were generated by randomly removing four RDCs per TAR construct. The reduced RDC data set was then used in the Monte Carlo selection, and the determined optimized ensemble was used to predict RDCs that are left out of the selection. As for direct analysis, this procedure was repeated 10 times, and the ensemble with the best active χ^2 is kept and used to back-calculate experimental data. A similar procedure was used for the cross-validation where each of the four data sets measured on each TAR construct was independently removed. The calculations of field-induced RDCs were carried out using a previously described protocol.[22,52]

Chemical Shift Calculation. Chemical shifts for H1′, H2, H5, H6, or H8 (total of 48) were computed using NUCHEMICS[53,54] for the TAR helices and bulge (excluding terminal base pairs G17-C45 and C29-G36 which are near sites of elongation and apical loop, respectively).

Ensemble Analysis. Interhelical angles (α_h, β_h, γ_h) describing the relative orientation of two A-form helices were computed using in-house software (http://hashimi.biop.lsa.umich.edu/resources) following a previously described procedure.[56] Owing to the high flexibility of the A22-U40 base pair, the lower helix is defined using central base pairs C19-G43, A20-U42, and G21-C41, and the interhelical Euler angles relate the upper helix defined by base pairs G26-C39, A27-U38, and G28-C37 to the lower reference helix. Previous studies used an identical procedure, except the three base pairs used for the lower helix were A20-U42, G21-C41, and A22-U40. Note that because the Euler angles used here are defined relative to a reference A-form helix lacking the locally flexible A22-U40 base pair; the α_h and γ_h interhelical Euler angles reported here are each systematically shifted by ∼17° as compared to previously reported TAR interhelical Euler angles.[23,51,56] This systematic shift of a constant value δ in Euler angles ($\alpha_h + \delta$, β_h, $\gamma_h - \delta$) is due to the fact that identical reference helices were used compared to previous studies but that the set of base pairs used for the calculation was shifted by one. This induces a shift in the interhelical twist of $2\delta = 34°$ simply corresponding to the twist between two successive base pairs in an idealized A-form helix (twist between G21-C41 and A22-U40). This referencing does not affect any local parameters and is used to ensure maximal accuracy in the description of the relative orientation of the two helices. Parameters defining local geometry of the bases were determined using Curves+,[61] and the calculation of sugar dihedral angles was preformed using in-house programs available from the authors upon request. The distance between two bases was calculated by measuring the distance between the center of mass of each base. All order parameters were calculated using the following equation:[76,77]

$$S^2 = \frac{1}{2}\left(3\sum_{i=1}^{3}\sum_{j=1}^{3} <\mu_i\mu_j>^2 - 1\right) \quad (5)$$

where μ_i represent the Cartesian coordinates of the normalized internuclear vector, after alignment of the conformers by superimposing heavy of the RNA backbone in the central base pairs C19-G43, A20-U42, and G21-C41 to an idealized A-form helix containing the same sequence. Error bars were obtained through Monte Carlo analysis (*vide infra*). The internal degree of order, ϑ_{int}, was calculated using helix I as reference according to a previously described procedure,[23] and the associated uncertainties were derived from Monte Carlo analysis (*vide infra*).

Monte Carlo Analysis. A Monte Carlo based approach was used to obtain error bars in the presented parameters, and 250 independently noise corrupted pseudoexperimental RDC data sets were generated using the RDCs predicted from the ensemble selected during the direct analysis of the experimental data (with 20 conformers). For each set of data, a selection of an ensemble of 20 conformers is done, and the presented errors bars correspond to the standard deviation obtained from this distribution of results.

Testing the Approach on Simulated Data. Target conformational ensembles were generated by selecting 2000 conformers from the conformational pool using, as probability of acceptance, a Gaussian distribution with predefined width and center for each of the three Euler angles. The set of 2000 conformers were then used to generate noise corrupted synthetic RDCs for the four different TAR constructs. The noise level used for the calculations was set to the weights used in the selection procedure. The synthetic RDCs were then inputted in the selection procedure. For each synthetic data set, 100 ensembles of 20 conformers were selected and combined for increased resolution and the ability of the procedure to reproduce an introduced sampling is characterized for both (α_h, β_h, γ_h) angles and local geometry.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1, Tables S1 and S2: comparisons of the four sets of RDCs. Figure S2: reproduction of magnetic field-induced RDCs. Figures S3−S5 and Table S3: capability of the approach to reproduce the sampling of synthetic target ensembles. Figures S6−S8: comparison with previous approaches for constructing ensembles. Tables S4−S6 and Figures S9−S10: additional information on local conformation dynamics of the constructed ensemble. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

hashimi@umich.edu

Notes

The authors declare the following competing financial interest(s): H.M.A.-H. is an advisor to and holds an ownership interest in Nymirum, an RNA-based drug-discovery company. The research reported in this article was performed by the University of Michigan faculty and students and was funded by a U.S. National Institutes of Health contract to H.M.A.-H.

■ ACKNOWLEDGMENTS

We thank Shan Yang, Dr. Aaron T. Frank, Anthony Mustoe, and Joseph Yesselman for their input and stimulating discussions. We acknowledge the National Resource for Biomedical Supercomputing (NRBSC) and the Pittsburgh Supercomputing Center (PSC) that provided Anton computer time through the grant RC2GM093307 from the National Institutes of Health. The Anton machine at NRBSC/PSC was generously made available by D.E. Shaw Research. This work was supported by the U.S. National Institutes of Health (R01 AI066975).

■ REFERENCES

(1) Dethoff, E. A.; Chugh, J.; Mustoe, A. M.; Al-Hashimi, H. M. *Nature* **2012**, *482*, 322−330.
(2) Rinnenthal, J.; Buck, J.; Ferner, J.; Wacker, A.; Fürtig, B.; Schwalbe, H. *Acc. Chem. Res.* **2011**, *44*, 1292−1301.
(3) Bardaro, M. F.; Varani, G. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 122−132.
(4) Schroeder, R.; Barta, A.; Semrad, K. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 908−919.
(5) Doudna, J. A.; Lorsch, J. R. *Nat. Struct. Mol. Biol.* **2005**, *12*, 395−402.
(6) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292−297.
(7) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.

(8) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106−9107.
(9) Ulmer, T. S.; Ramirez, B. E.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179−9191.
(10) Salmon, L.; Pierce, L.; Grimm, A.; Ortega-Roldan, J.-L.; Mollica, L.; Jensen, M. R.; van Nuland, N.; Markwick, P. R. L.; McCammon, J. A.; Blackledge, M. *Angew Chem., Int. Ed.* **2012**, *51*, 6103−6106.
(11) Lakomek, N.-A.; Walter, K. F. A.; Fares, C.; Lange, O. F.; Groot, B. L.; Grubmüller, H.; Brüschweiler, R.; Munk, A.; Becker, S.; Meiler, J.; Griesinger, C. *J. Biomol. NMR* **2008**, *41*, 139−155.
(12) Salmon, L.; Bouvignies, G.; Markwick, P.; Blackledge, M. *Biochemistry* **2011**, *50*, 2735−2747.
(13) Markwick, P. R. L.; Bouvignies, G.; Salmon, L.; McCammon, J. A.; Nilges, M.; Blackledge, M. *J. Am. Chem. Soc.* **2009**, *131*, 16968−16975.
(14) Showalter, S. A.; Brüschweiler, R., *J. Am. Chem. Soc.* **2007**, *129*, 4158−4159.
(15) Lange, O. F.; Lakomek, N.-A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmüller, H.; Griesinger, C.; De Groot, B. L. *Science* **2008**, *320*, 1471−1475.
(16) Clore, G. M.; Schwieters, C. D. *Biochemistry* **2004**, *43*, 10678−10691.
(17) Tolman, J. R.; Ruan, K. *Chem Rev* **2006**, *106*, 1720−1736.
(18) Blackledge, M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 23−61.
(19) Getz, M.; Sun, X.; Casiano-Negroni, A.; Zhang, Q.; Al-Hashimi, H. M. *Biopolymers* **2007**, *86*, 384−402.
(20) Latham, M. P.; Hanson, P.; Brown, D. J.; Pardi, A. *J. Biomo.l NMR* **2007**, *40*, 83−94.
(21) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. *ChemBioChem* **2005**, *6*, 1492−1505.
(22) Zhang, Q.; Al-Hashimi, H. M. *Nat. Methods* **2008**, *5*, 243−245.
(23) Zhang, Q.; Stelzer, A. C.; Fisher, C. K.; Al-Hashimi, H. M. *Nature* **2007**, *450*, 1263−1267.
(24) Al-Hashimi, H. M.; Gosser, Y.; Gorin, A.; Hu, W.; Majumdar, A.; Patel, D. J. *J. Mol. Biol.* **2002**, *315*, 95−102.
(25) Zhang, Q.; Sun, X.; Watt, E. D.; Al-Hashimi, H. M. *Science* **2006**, *311*, 653−656.
(26) Dingwall, C.; Ernberg, I.; Gait, M. J.; Green, S. M.; Heaphy, S.; Karn, J.; Lowe, A. D.; Singh, M.; Skinner, M. A.; Valerio, R. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6925−6929.
(27) Weeks, K. M.; Ampe, C.; Schultz, S. C.; Steitz, T. A.; Crothers, D. M. *Science* **1990**, *249*, 1281−1285.
(28) Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. *Science* **1995**, *270*, 1200−1203.
(29) Puglisi, J.; Chen, L.; Frankel, A.; Williamson, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3680−3684.
(30) Olsen, G. L.; Bardaro, M. F.; Echodu, D. C.; Drobny, G. P.; Varani, G. *J. Am. Chem. Soc.* **2010**, *132*, 303−308.
(31) Olsen, G. L.; Echodu, D. C.; Shajani, Z.; Bardaro, M. F.; Varani, G.; Drobny, G. P. *J. Am. Chem. Soc.* **2008**, *130*, 2896−2897.
(32) Lu, J.; Kadakkuzha, B. M.; Zhao, L.; Fan, M.; Qi, X.; Xia, T. *Biochemistry* **2011**, *50*, 5042−5057.
(33) Emani, P. S.; Olsen, G. L.; Echodu, D. C.; Varani, G.; Drobny, G. P. *J. Phys. Chem. B* **2010**, *114*, 15991−16002.
(34) Chen, Y.; Campbell, S. L.; Dokholyan, N. V. *Biophys. J.* **2007**, *93*, 2300−2306.
(35) Dethoff, E. A.; Hansen, A. L.; Zhang, Q.; Al-Hashimi, H. M. *J. Magn. Reson.* **2010**, *202*, 117−121.
(36) Zweckstetter, M. *Nat. Protoc.* **2008**, *3*, 679−690.
(37) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(38) Sigel, R. K. O.; Sashital, D. G.; Abramovitz, D. L.; Palmer, I.; Arthur, G.; Butcher, S. E.; Pyle, A. M. *Nat. Struct. Mol. Biol.* **2004**, *11*, 187−192.
(39) Zweckstetter, M.; Hummer, G.; Bax, A. *Biophys. J.* **2004**, *86*, 3444−3460.
(40) Eichhorn, C. D.; Feng, J.; Suddala, K. C.; Walter, N. G.; Brooks, C. L. I.; Al-Hashimi, H. M. *Nucleic Acids Res.* **2012**, *40*, 1345−1355.
(41) Wu, B.; Petersen, M.; Girard, F.; Tessari, M.; Wijmenga, S. S. *J. Biomol. NMR* **2006**, *35*, 103−115.
(42) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(43) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(44) Sass, J.; Cordier, F.; Hoffmann, A.; Rogowski, M.; Cousin, A.; Omichinski, J. G.; Löwen, H.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 2047−2055.
(45) Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C. *Commun ACM* **2008**, *51*, 91−97.
(46) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D., Jr. *J. Comput. Chem.* **2011**, *32*, 1929−1943.
(47) Foloppe, N.; MacKerell, A. D., Jr *J. Comput. Chem.* **2000**, *21*, 86−104.
(48) MacKerell, A. D.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21*, 105−120.
(49) Nodet, G.; Salmon, L.; Ozenne, V.; Meier, S.; Jensen, M. R.; Blackledge, M. *J. Am. Chem. Soc.* **2009**, *131*, 17908−17918.
(50) Dethoff, E. A.; Hansen, A. L.; Musselman, C.; Watt, E. D.; Andricioaei, I.; Al-Hashimi, H. M. *Biophys. J.* **2008**, *95*, 3906−3915.
(51) Frank, A. T.; Stelzer, A. C.; Al-Hashimi, H. M.; Andricioaei, I. *Nucleic Acids Res.* **2009**, *37*, 3670−3679.
(52) Zhang, Q.; Throolin, R.; Pitt, S. W.; Serganov, A.; Al-Hashimi, H. M. *J. Am. Chem. Soc.* **2003**, *125*, 10530−10531.
(53) Wijmenga, S. S.; Kruithof, M.; Hilbers, C. W. *J. Biomol. NMR* **1997**, *10*, 337−350.
(54) Cromsigt, J. A.; Hilbers, C. W.; Wijmenga, S. S. *J. Biomol. NMR* **2001**, *21*, 11−29.
(55) Al-Hashimi, H. M.; Walter, N. G. *Curr. Opin. Struct. Biol.* **2008**, *18*, 321−329.
(56) Bailor, M. H.; Mustoe, A. M.; Brooks, C. L.; Al-Hashimi, H. M. *Nat. Protoc.* **2011**, *6*, 1536−1545.
(57) Bailor, M. H.; Sun, X.; Al-Hashimi, H. M. *Science* **2010**, *327*, 202−206.
(58) Pitt, S. W.; Majumdar, A.; Serganov, A.; Patel, D. J.; Al-Hashimi, H. M. *J. Mol. Biol.* **2004**, *338*, 7−16.
(59) Casiano-Negroni, A.; Sun, X.; Al-Hashimi, H. M. *Biochemistry* **2007**, *46*, 6525−6535.
(60) Mustoe, A. M.; Bailor, M. H.; Teixeira, R. M.; Brooks, C. L. I.; Al-Hashimi, H. M. *Nucleic Acids Res.* **2012**, *40*, 892−904.
(61) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1988**, *6*, 63−91.
(62) Barton, S.; Heng, X.; Johnson, B. A.; Summers, M. F. *J. Biomol. NMR* **2012**, *55*, 33−46.
(63) Hansen, A. L.; Al-Hashimi, H. M. *J. Am. Chem. Soc.* **2007**, *129*, 16072−16082.
(64) Aboul-ela, F.; Karn, J.; Varani, G. *Nucleic Acids Res.* **1996**, *24*, 3974−3981.
(65) Puglisi, J. D.; Tan, R.; Calnan, B. J.; Frankel, A. D.; Williamson, J. R. *Science* **1992**, *257*, 76−80.
(66) Ippolito, J. A.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9819−9824.
(67) Parisien, M.; Major, F. *Nature* **2008**, *452*, 51−55.
(68) Dethoff, E. A.; Petzold, K.; Chugh, J.; Casiano-Negroni, A.; Al-Hashimi, H. M. *Nature* **2012**, *491*, 724−728.
(69) Stelzer, A. C.; Kratz, J. D.; Zhang, Q.; Al-Hashimi, H. M. *Angew. Chem., Int. Ed.* **2010**, *49*, 5731−5733.
(70) Tolman, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 12020−12030.
(71) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Brüschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885−13890.
(72) Salmon, L.; Bouvignies, G.; Markwick, P.; Lakomek, N.; Showalter, S.; Li, D.-W.; Walter, K.; Griesinger, C.; Brueschweiler, R.; Blackledge, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 4154−4157.
(73) Meiler, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098−6107.
(74) Peti, W.; Meiler, J.; Brüschweiler, R.; Griesinger, C., *J. Am. Chem. Soc.* **2002**, *124*, 5822−5833.

(75) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(76) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239−250.
(77) Markwick, P. R. L.; Bouvignies, G.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 4724−4730.
(78) Musselman, C.; Zhang, Q.; Al-Hashimi, H. M.; Andricioaei, I. *J. Phys. Chem. B* **2010**, *114*, 929−939.
(79) Zuo, X.; Wang, J.; Foster, T. R.; Schwieters, C. D.; Tiede, D. M.; Butcher, S. E.; Wang, Y.-X. *J. Am. Chem. Soc.* **2008**, *130*, 3292−3293.
(80) Grishaev, A.; Ying, J.; Canny, M. D.; Pardi, A.; Bax, A. *J. Biomol. NMR* **2008**, *42*, 99−109.
(81) Frank, A. T.; Horowitz, S.; Andricioaei, I.; Al-Hashimi, H. M. *J. Phys. Chem. B* **2013**, *117*, 2045−2052.
(82) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−8- 27−8.
(83) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D., Jr; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(84) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(85) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(86) Ennifar, E.; Nikulin, A.; Tishchenko, S.; Serganov, A.; Nevskaya, N.; Garber, M.; Ehresmann, B.; Ehresmann, C.; Nikonov, S.; Dumas, P. *J. Mol. Biol.* **2000**, *304*, 35−42.