A method of determining RNA conformational ensembles using structure-based calculations of residual dipolar couplings

Aditi N. Borkar,[1] Alfonso De Simone,[1,2] Rinaldo W. Montalvao,[1,3] and Michele Vendruscolo[1,a)]
[1]*Department of Chemistry, University of Cambridge, Cambridge, United Kingdom*
[2]*Division of Molecular Biosciences, Imperial College, London, United Kingdom*
[3]*Institute of Physics, University of São Paulo, São Carlos, Brazil*

(Received 25 February 2013; accepted 24 April 2013; published online 7 June 2013)

We describe a method of determining the conformational fluctuations of RNA based on the incorporation of nuclear magnetic resonance (NMR) residual dipolar couplings (RDCs) as replica-averaged structural restraints in molecular dynamics simulations. In this approach, the alignment tensor required to calculate the RDCs corresponding to a given conformation is estimated from its shape, and multiple replicas of the RNA molecule are simulated simultaneously to reproduce *in silico* the ensemble-averaging procedure performed in the NMR measurements. We provide initial evidence that with this approach it is possible to determine accurately structural ensembles representing the conformational fluctuations of RNA by applying the reference ensemble test to the *trans*-activation response element of the human immunodeficiency virus type 1. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4804301]

INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy offers a variety of techniques that can provide information on the motions of biological macromolecules.[1–13] Since NMR observables are obtained from time and ensemble averaged measurements, it is possible to use them for determining conformational ensembles of proteins and nucleic acids.[14–26] Achieving this result for systems that are conformationally highly heterogeneous, however, is still particularly challenging because of the great technical difficulties in the incorporation of the experimental measurements in the structure determination procedures.[27,28] The use of residual dipolar couplings (RDCs),[29,30] offers particularly good opportunities in this respect.[15,17,18,31–33]

In the approach that we discuss here, RDCs are incorporated as structural restraints in molecular dynamics simulations. This strategy is convenient to describe the dynamics of proteins and nucleic acids, as it can provide atomic resolution information about their structural fluctuations by integrating numerically the equations of motions with a bias constructed in order to maximise the agreement with experimental observations, in a manner consistent with the maximum entropy principle.[34–36] More generally, by using RDC measurements as structural restraints it has been shown that accurate structural ensembles of proteins and nucleic acids can be obtained.[15–18] In this work we extend to RNA a structure-based approach for the calculation of the alignment tensor that has been recently introduced for proteins[37,38] and that is complementary to a series of related methods.[15,18,31,32]

We illustrate this approach by considering the case of the *trans*-activation response (TAR) element of the human immunodeficiency virus type 1 (HIV-1), a system that is well characterised experimentally[39–43] (Figure 1). Several structures of TAR have been deposited in the Protein Data Bank (PDB), including for the free state,[39] and for the states bound to Tat[40] and other small molecules.[39–43] A comprehensive study of these structures has also revealed that TAR exhibits significant dynamics in the global orientation of the two helices and local structure of the binding pocket (Figure 1). In a series of recent studies, Al-Hashimi and co-workers have shown that TAR binds to its partners by a conformational selection mechanism in which bound-like conformations are sampled during the structural fluctuations in the free state.[18,44]

METHODS

Assessment of the similarity between two structural ensembles: The S score

In order to compare the similarity between two structural ensembles, A and B, in this work we used the S score.[21,26,37] For a RNA molecule of N nucleotides, the S score is calculated from a $N \times N$ matrix, in which each element S_{ij} represents the difference between the distributions P^A and P^B of the distances between atoms i and j in the two ensembles to be compared:

$$S_{ij} = \frac{1}{2} \sum_k \left| P^A_{ij,k} - P^B_{ij,k} \right|, \quad (1)$$

where k runs over the bins used to characterise the distributions. The values of S_{ij} range from 0 for identical distributions to 1 for non-overlapping ones.[26] The S matrix, thus, effectively compares the similarity of the reference ensemble with the simulated ensemble. From the S matrix, an overall S score is calculated by averaging the S_{ij} elements.

a)Author to whom correspondence should be addressed. Electronic mail: mv245@cam.ac.uk

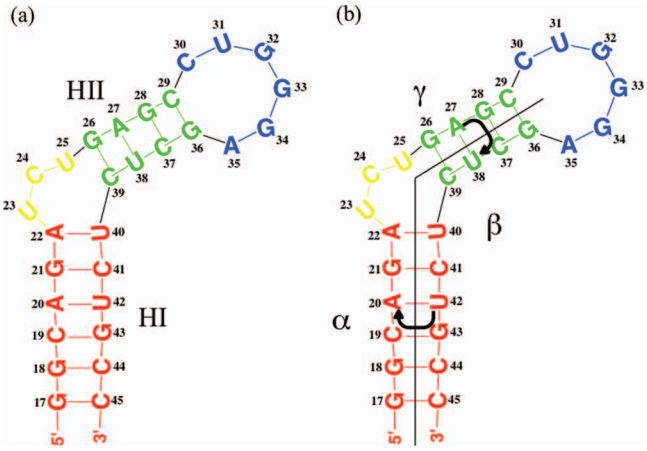

FIG. 1. Schematic representation of the structure of the HIV-1 TAR element. (a) The TAR structure comprises two helices, HI and HII, joined by a bulge (yellow). HII is capped by 6-residue loop (blue). The bulge and the loop residues form the binding pocket of TAR. (b) Illustration of the three Euler angles[73] for describing the conformation of TAR: rotation of HI with respect to HII (α), rotation of HII with respect to HI (γ) and interhelical bend angle (β); the interhelical twist (ζ) is defined as $\alpha - \gamma$.

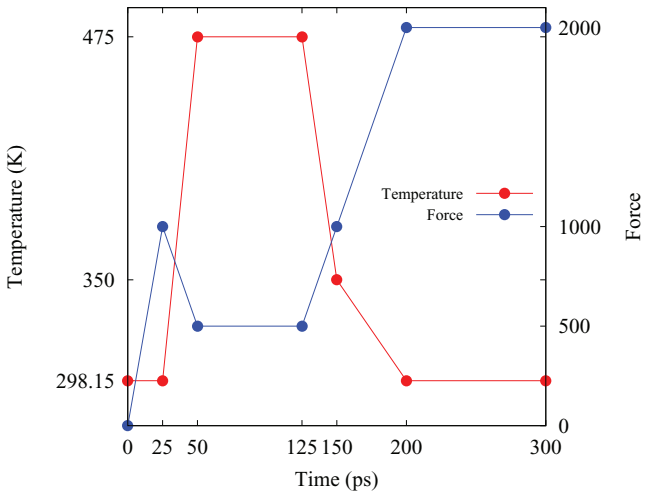

FIG. 2. Simulated annealing schedule used in the simulations. The variations in the temperature (red) and relative restraint force (blue, see Eq. (3)) during a cycle of simulated annealing are shown.

Structure-based calculation of the RDCs

For each given conformation of TAR, we calculated the alignment tensor to obtain the RDCs using a structure-based method[37,38] that considers the orientations of the RNA molecule in a given conformation that are allowed by the alignment medium.[45–52]

Molecular dynamics simulations and generation of the starting conformations

All molecular dynamics simulations were performed with GROMACS 4.5.[53,54] We chose eight structures from the 1ANR file of free TAR[39] as the starting set of conformations for the eight replicas required for the simulations. Each of these eight structures was placed in an individual octahedron box with sides 12 Å and solvated with TIP3P water.[55] In addition to K^+ neutralising ions, a 100 mM concentration of $MgCl_2$ was used. After an initial energy minimization, first with steepest descent and then with conjugate gradient methods, the system was simulated for 50 ps with positional restraints on the RNA molecules. Subsequently, these position restraints were removed and the system was simulated under NVT conditions for further 20 ps and then heated to the final temperature of 298.15 K under NPT condition for ~100 ps. All further simulations were continued as production runs from these conformations.

Generation of the reference ensemble and reference RDCs

The reference ensemble was generated using the CHARMM27 force field[56–59] by running a trajectory of 12 ns on each replica, i.e., of 96 ns in total, using the setup described above. As the sampling was performed on the last 0.1 ns of each simulated annealing cycle (Figure 2), about 4000 conformations were sampled in total from the trajectory of each replica. Finally, about 15 000 structures with RMSD between 0.6 nm and 1.0 nm (Figure 3) from the equilibrated part of these sampled conformations were extracted to build up the "reference ensemble." An alignment tensor and the corresponding RDCs for 56 internuclear vectors (Figure 4) were

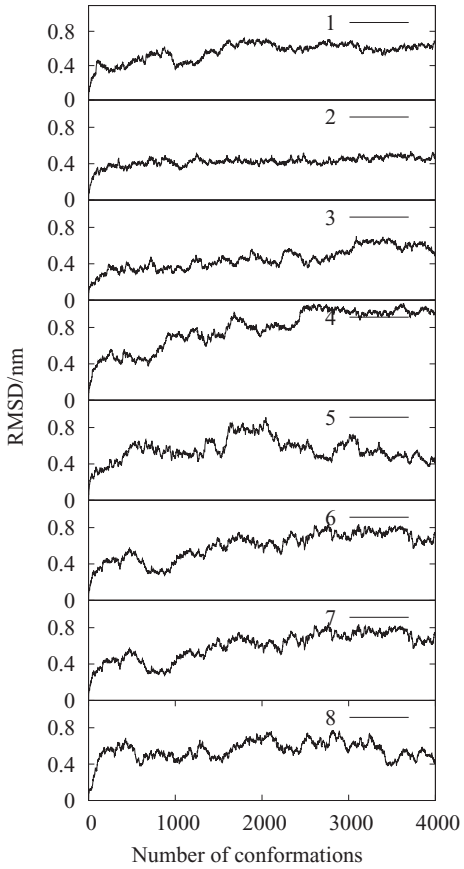

FIG. 3. RMSD of the conformations sampled by the CHARMM27 simulation. About 15 000 structures with RMSD between 0.6 nm and 1 nm from the equilibrated part of these sampled conformations were selected to build up the reference ensemble.

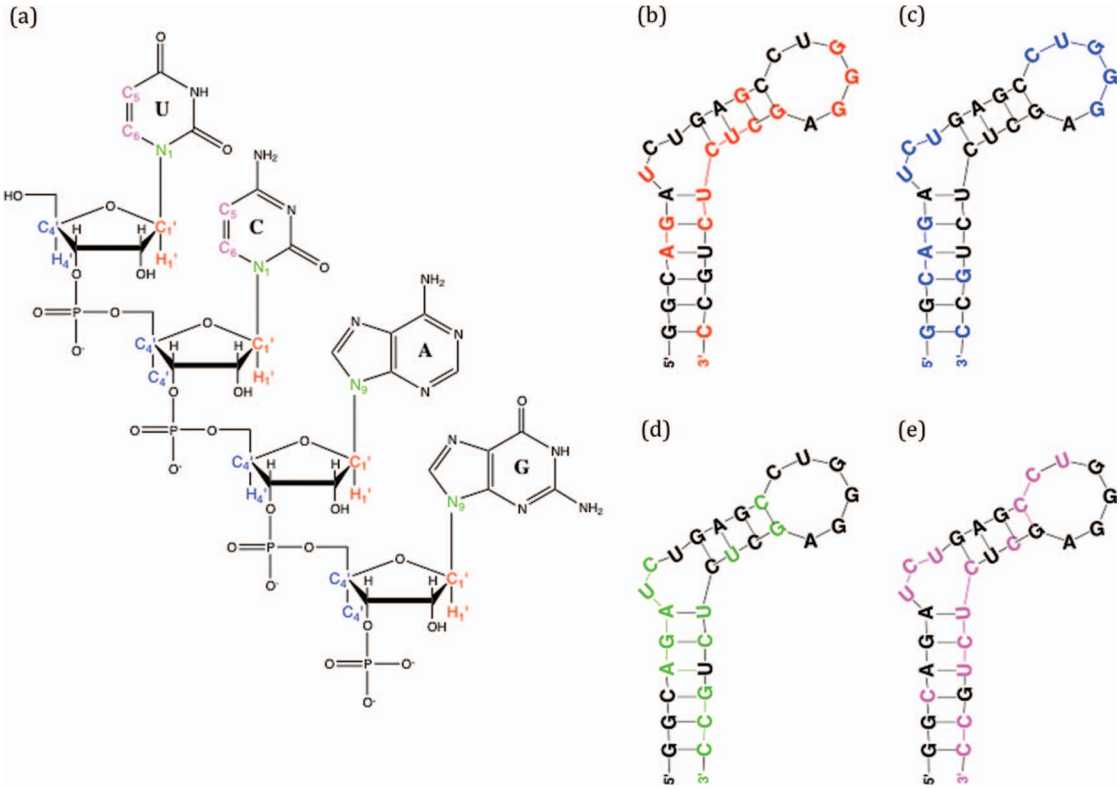

FIG. 4. Summary of the 56 RDC restraints used in the simulations. (a) Illustration of the positions of these bonds within a nucleotide. (b)-(e) Illustration of the positions of the restrained bonds. The RDCs are colour coded as C4′–H4′ (blue), C1′–H1′ (red), C1′–N1/N9 (green), and C5-C6 (pink).

individually calculated for each of these structures using the structure-based method mentioned above.[37,38] The linear average of these RDCs was then used to obtain a set of 56 "reference RDCs" (Figure 4).

Generation of the unrestrained ensemble

The starting conformations were next simulated with AMBER99bsc0 force field[60–64] using the set up described above and a simulated annealing protocol (Figure 2). To generate the "unrestrained ensemble," only the temperature was annealed. This unrestrained ensemble was used for comparison with the restrained simulation.

Generation of the reconstructed ensembles

As for the unrestrained ensemble, the starting conformations were next simulated with AMBER99bsc0 force field using the setup described above and a simulated annealing protocol (Figure 2). In these simulations, restraints were imposed by adding a pseudoenergy term (E_{RDC}) to a standard molecular mechanics force field (E_{MM}):[35,36]

$$E_{\mathrm{TOT}} = E_{\mathrm{MM}} + E_{\mathrm{RDC}}. \quad (2)$$

The resulting force field (E_{TOT}) was employed in molecular dynamics simulations. The pseudoenergy term E_{RDC} is given by[35,36]

$$E_{\mathrm{RDC}} = \alpha \sum_i (D^{ref} - D^{cal})^2, \quad (3)$$

where the sum is taken over the available reference RDCs (D^{ref}) and α is the restraint force constant. This term is used to bias the trajectory towards conformations in which the calculated RDCs (D^{cal}) match the reference ones. The calculated RDCs were obtained as averages over the eight replicas of the RNA molecule used in the simulations. As mentioned above, for each replica the alignment tensors are independently computed using a structure-based method[36,37] and are used in Eq. (3). To enhance the sampling, annealing of both temperature and force constant for the RDC restraints[37] was employed on the system. These restrained simulations generated the "reconstructed ensemble."

RESULTS

The reference ensemble method

In this study, we address the problem of assessing the performance of an approach in which the RDCs are implemented as replica-averaged restraints in molecular dynamics simulations of nucleic acids, and in which the RDCs are calculated during the simulations using a method that was recently proposed for proteins in which the alignment tensors are calculated from the conformations of the molecules.[37,38]

To perform the assessment, we carried out the test of the "reference ensemble"[21,26,37] to the HIV-1 TAR element. In this test, a reference ensemble of conformations is generated at first by unrestrained molecular dynamics simulations using a given force field (FF1, Figure 5, path A), in this case the CHARMM27 force field[56–59] (see Methods section). A set of

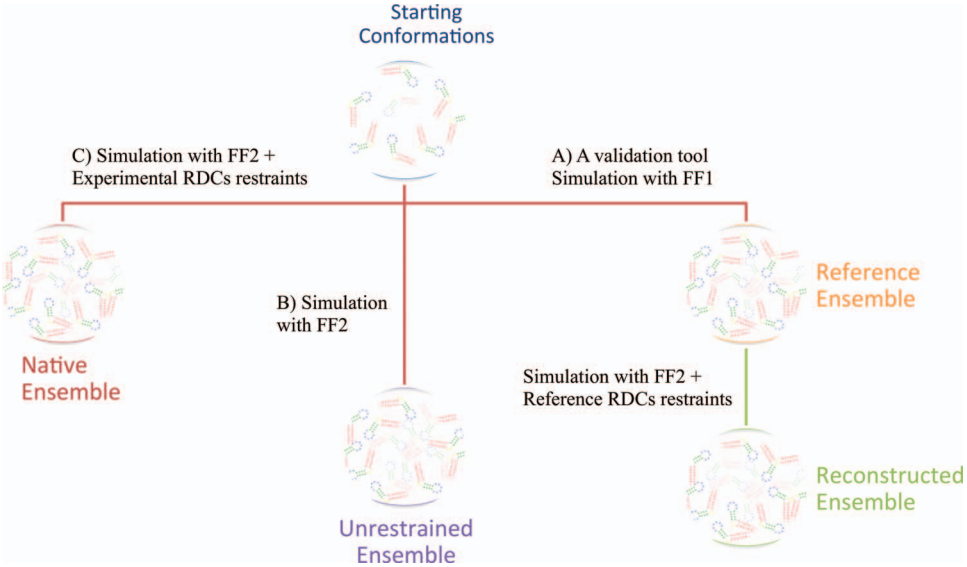

FIG. 5. Scheme for generating ensembles of RNA structures. (A) A "reference ensemble" is generated by molecular dynamics simulations using force field 1 (FF1). RDC restraints are calculated as average values over the reference ensemble. (B) An "unrestrained ensemble" is generated in a similar manner but using force field 2 (FF2). Then by applying the RDC restraints in conjunction with FF2 a "reconstructed ensemble" (or "restrained ensemble") is generated. If the restraints are applied correctly the reconstructed ensemble is closer to the reference ensemble than the unrestrained ensemble. (C) After the validation by this "reference ensemble test," the method can be applied using experimentally measured RDCs to determine an ensemble of conformations representing the structural fluctuations of RNA.

"reference RDCs" is then calculated from the structures making up this reference ensemble (see Methods section) and employed as structural restraints to generate a "reconstructed ensemble" using a second force field (FF2, Figure 5, path B), in this case the AMBER99bsc0 force field[60–64] (see Methods section). Although these two force fields are specifically parameterized for nucleic acids simulations, they generate two distinct structural ensembles (i.e., the "reference ensemble" for FF1 and the "unrestrained ensemble" for FF2, see Methods section) because of a series of small but significant differences in their functional forms and parameters. When, however, the RDC restraints are imposed through the approach that we describe in this work, they bias the sampling of conformational space carried out with FF2 towards the regions consistent with the RDC values resulting from FF1. Thus, the "reference" and the "reconstructed" ensembles should end up being very similar.

An advantage of using the reference ensemble test is that it allows for a stringent validation analysis in which the atomic coordinates of the conformations in the reference ensemble are known exactly, and therefore the accuracy of the conformations in the reconstructed ensemble can be assessed with great confidence (Figure 5, paths A and B). Once the use of RDC as replica-averaged restraints in molecular dynamics simulations has been validated, experimental RDCs can be used to drive the sampling of conformational space and generate ensembles of conformations representing the structural fluctuations of the molecules under observation (Figure 5, path C).

Conformational sampling in the simulations

The extent of conformational sampling in the simulations was initially verified by plotting the free energy landscapes corresponding to the reference, unrestrained, and reconstructed ensembles (Figure 6). Our results show that the reference ensemble, which was generated with the CHARMM27 force field, populates a region of conformational space different from that of the unrestrained simulations, which was generated with the AMBER99bsc0 force field (Figure 6). By contrast, the reconstructed ensemble, which was also obtained using the AMBER99bsc0 force field, samples a region of the conformational space similar to that of the reference ensemble, despite the fact that the latter was generated using the CHARMM27 force field, thus showing the effects of the restraints on the sampling (Figure 6).

The three ensembles were analysed in order to identify possible structural deformations resulting from the use of the restraints. We considered the distribution of parameters specific of nucleic acids, including dihedral angles (Fig. S1),[65] helical parameters[66] (Fig. S2),[65] and H-bonding patterns (Fig. S3).[65] The dihedral angles and helical parameters were calculated using a recent version of NUPARM[67] and the H-bond base pair patterns were calculated using 3DNA.[68,69] The results of these calculations suggest that the sampling in all the ensembles is carried out without creating significant structural distortions. This result is consistent with the recent demonstration that the use of replica-averaged structural restraints in molecular dynamics simulations represents an effective way to generate structural ensembles consistent with the maximum entropy principle.[34–36] In this sense, the procedure involves the minimal possible alteration of the force field that enables the experimental observations to be consistent with the generated structures.

The convergence of the simulations for each ensemble was verified by calculating the S scores between the conformations in the first and second halves of the trajectory (Figure 7). For all the atom types considered, the S scores

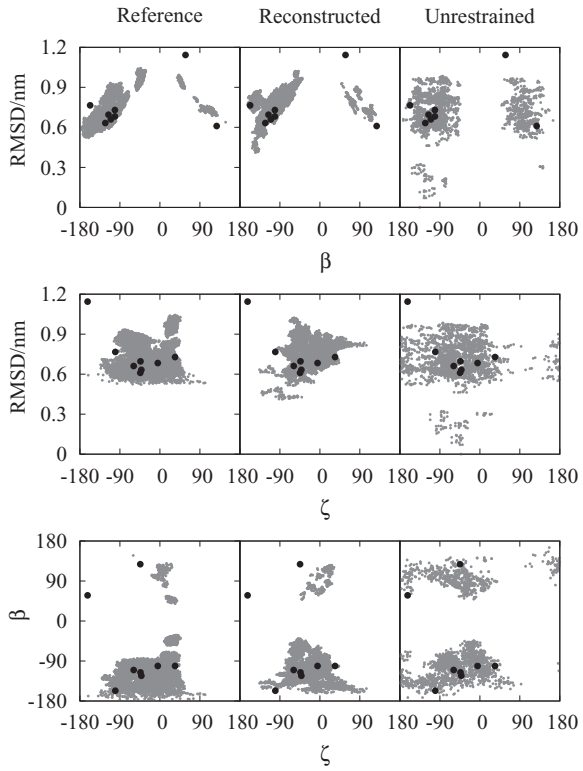

FIG. 6. Assessment of the extent of conformational sampling. Comparison of the conformational space sampled by the reference, reconstructed, and unrestrained ensembles (grey regions), as a function of the Euler angles β and ζ. The starting structures (black dots) in the simulations are also shown. The RMSD is calculated with respect to the 1ANR structure. The graphs illustrate that the CHARMM27 trajectory used to generate the reference ensemble equilibrates to a population away from the state populated by the AMBER99bsc0 force field. By contrast, the reconstructed ensemble, which was also obtained using the AMBER99bsc0 force field, samples a region of the conformational space similar to that of the reference ensemble, which was generated using the CHARMM27 force field, thus showing the effect of the restrains on the sampling.

were between 0.09 and 0.11 for the reconstructed ensemble, and between 0.10 and 0.13 for the unrestrained ensemble. These *S* scores are comparable with those found in analogous tests for proteins,[26] and correspond to the statistical fluctuations within an equilibrium ensemble.

Comparison of the Q factors of the unrestrained and reconstructed ensembles

In order to further assess whether the use of restraints helps in the reconstruction of the reference ensemble, we compared the RDCs from the conformations in the reconstructed and unrestrained ensembles using the Q factor.[70–72] The Q factor (Eq. (4)) provides a normalised metric for agreement between the reference RDCs (D^{ref}) and the RDCs calculated from the reconstructed or unrestrained ensembles (D^{cal}):

$$Q = \sqrt{\frac{\sum_{i=1}^{N} \left(D_i^{ref} - D_i^{cal}\right)^2}{\sum_{i=1}^{N} \left(D_i^{ref}\right)^2}}. \quad (4)$$

In these calculations, RDCs were back-calculated using the structure-based method[37,38] described above (see Methods section). The distribution of Q factors resulting from these calculated RDCs indicates that the reconstructed ensemble (red histograms, Figure 8) reproduces the RDCs of the reference ensemble better than the unrestrained ensemble (blue histograms, Figure 8). This conclusion holds both for the set of 56 RDCs (see Figure 4) used as restraints in the molecular dynamics simulations (Figure 8(a)) and for the set of RDCs corresponding to all the C–H, N–H, and C–N bonds in TAR not used as restraints (Figure 8(b)). These results indicate that the use of RDC restraints in the procedure that we describe in this work enables one to determine ensembles of conformations with high accuracy.

Comparison of the structures of the unrestrained and reconstructed ensembles

To assess whether the good agreement between the RDCs of the reference and reconstructed ensembles also corresponds to a better description of the conformational heterogeneity of the reference ensemble, we compared the global and local orientations of TAR in the two ensembles. The TAR global conformation can be described in terms of four Euler angles[73] (Figure 1). The distributions of these Euler angles for the reference, reconstructed, as well as unrestrained

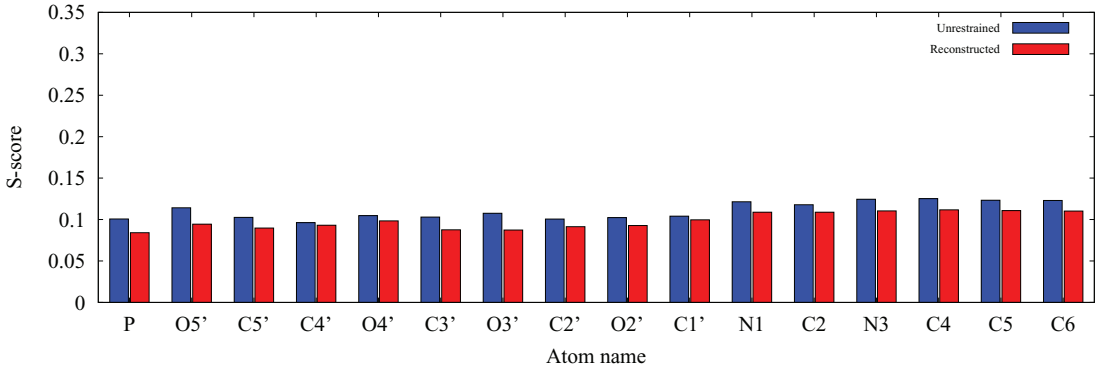

FIG. 7. Assessment of the convergence of the simulations. The S scores calculated by comparing the structures in the first and second halves of the simulations show that both the unrestrained (blue) and reconstructed (red) ensembles are converged. 16 S matrices were constructed for different atom types. These S scores are comparable with those obtained for proteins[26] and reflect the statistical differences within the conformations of an equilibrium ensemble.

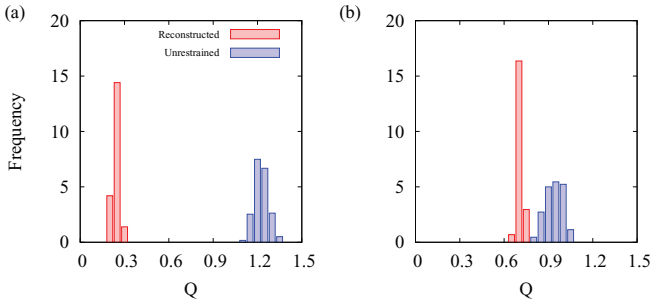

FIG. 8. Comparison of the Q factor distributions for the reconstructed and unrestrained ensembles. (a) Q factor distributions for the 56 RDCs restrained in the simulations (see Figure 2). (b) Q factor distributions for the RDCs not restrained in the simulations, which included all the C–H, N–H, and C–N bonds in TAR. These results show that both the RDCs that were restrained and those that were not restrained are better reproduced in the reconstructed ensemble than in the unrestrained ensemble.

ensembles (Figure 9) clearly show that although the reconstructed and unrestrained ensembles are generated using the same molecular dynamics protocol, the RDC restraints drive the trajectory of the reconstructed ensemble towards the reference ensemble rather than following the sampling that would result from the AMBER99bsc0 force field.

In order to further assess the similarity of the reference, unrestrained, and reconstructed ensembles, in addition to the nucleic acid statistics mentioned above (Figs. S1-S3),[65] we built a series of 29 × 29 S matrices[26] (see Methods section). The overall S scores for the S matrices corresponding to different atom type contacts indicate that the reconstructed ensemble is systematically closer to the reference ensemble (between 1% and 25% depending on the atom type) than the unrestrained ensemble (Figure 10(c)). For clarity, we show the S matrices corresponding to pairs of P atoms (Figure 10(b)) and pairs of N3 atoms (Figure 10(c)) in the three ensembles.

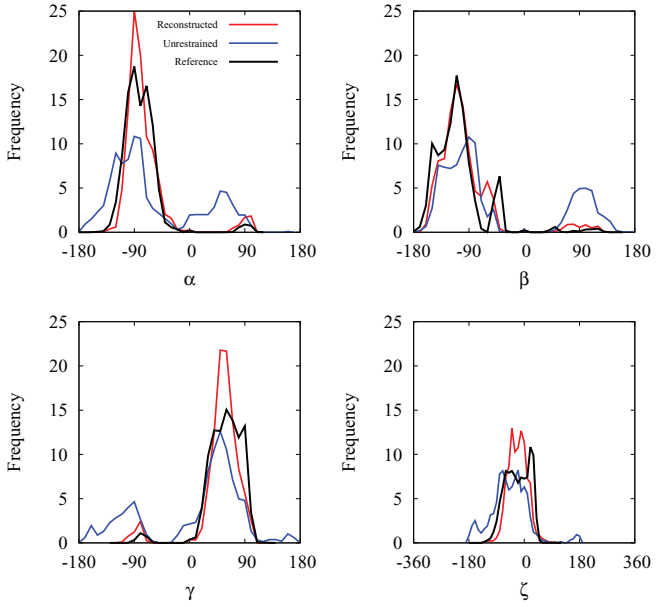

FIG. 9. Comparison of the Euler angle distributions in the reference, reconstructed, and unrestrained ensembles. The values of the Euler angles α (a), β (b), γ (c), and ζ (d) are shown for the unrestrained (green), reference (black), and reconstructed (red) ensembles.

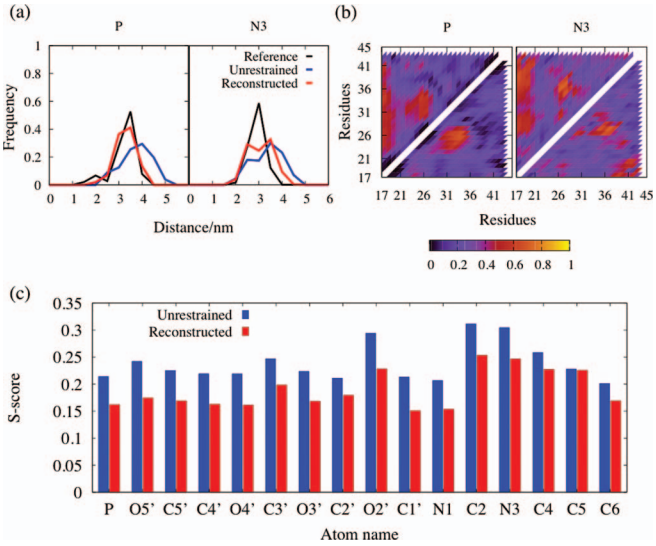

FIG. 10. Overall comparison of the interatomic distance distributions in the reconstructed, unrestrained, and reference ensembles. (a) Representative distributions of the interatomic distances of P (or N3) in C19 and P (or N3) in G32 in the three ensembles. (b) Representative S matrices of P (or N3) in all the TAR residues. The upper part shows the S matrix for the unrestrained ensemble and the lower part for the reconstructed ensemble. Yellow and red colours indicate large differences between the reference and the unrestrained and reconstructed ensembles, respectively. (c) Average S scores[26] for all the atoms in the unrestrained (blue) and reconstructed (red) ensembles.

As described in the Methods section, each element S_{ij} of a S matrix provides a measure of the distance distributions between atom pairs in two structural ensembles. We illustrate these distributions for the two pairs of P and N3 atoms (Figure 10(a)) of C19 and G32.

In summary, we have found that there are differences between the structural ensembles of TAR generated from the AMBER99bsc0 and CHARMM27 force fields. Our results show that the use of RDC restraints reduces these differences by providing a reconstructed ensemble closer to the reference (CHARMM27) ensemble than the unrestrained (AMBER99bsc0) one.

CONCLUSIONS

We have shown that the use of replica-averaged structure-based RDC restraints in molecular dynamics simulations of RNA molecules generates a bias in the sampling that directs the trajectories towards the region of conformational space compatible with the RDCs themselves. Our findings indicate that by exploiting such a bias, even relatively short simulations are sufficient to determine an ensemble of conformations that reproduces rather closely the structural features of a given reference ensemble. Taken together, the results that we have presented indicate that the use of RDCs as replica-averaged structural restraints in molecular dynamics simulations represents a promising strategy to calculate ensembles of structures representing the conformational fluctuations of RNA molecules.

ACKNOWLEDGMENTS

The authors would like to acknowledge Mike Bardaro, Gabriele Varani, and Hashim Al Hashimi for many useful discussions. A.N.B. also acknowledges the Dr. Manmohan Singh Scholarship Programme and St. John's College for funding her Ph.D. studies.

[1]H. M. Al-Hashimi, Biopolymers **86**, 345 (2007).
[2]C. G. Kalodimos, Protein Sci. **20**, 773 (2011).
[3]A. Mittermaier and L. E. Kay, Science **312**, 224 (2006).
[4]A. G. Palmer, Chem. Rev. **104**, 3623 (2004).
[5]S. E. Butcher, F. H. T. Allain, and J. Feigon, Biochemistry **39**, 2174 (2000).
[6]Z. H. Liu, I. Luyten, M. J. Bottomley, A. C. Messias, S. Houngninou-Molango, R. Sprangers, K. Zanier, A. Kramer, and M. Sattler, Science **294**, 1098 (2001).
[7]G. Varani, F. Aboulela, and F. H. T. Allain, Prog. Nucl. Magn. Reson. Spectrosc. **29**, 51 (1996).
[8]C. Dominguez, M. Schubert, O. Duss, S. Ravindranathan, and F. H. T. Allain, Prog. Nucl. Magn. Reson. Spectrosc. **58**, 1 (2011).
[9]J. Flinders and T. Dieckmann, Prog. Nucl. Magn. Reson. Spectrosc. **48**, 137 (2006).
[10]B. Fürtig, C. Richter, J. Wöhnert, and H. Schwalbe, ChemBioChem **4**, 936 (2003).
[11]K. Hall, Methods Mol. Biol. **875**, 67 (2012).
[12]M. P. Latham, D. J. Brown, S. A. McCallum, and A. Pardi, ChemBioChem **6**, 1492 (2005).
[13]P. J. Lukavsky and J. D. Puglisi, Methods Enzymol. **394**, 399 (2005).
[14]R. B. Best and M. Vendruscolo, J. Am. Chem. Soc. **126**, 8090 (2004).
[15]G. M. Clore and C. D. Schwieters, J. Am. Chem. Soc. **126**, 2923 (2004).
[16]K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson, and M. Vendruscolo, Nature (London) **433**, 128 (2005).
[17]J. R. Bothe, E. N. Nikolova, C. D. Eichhorn, J. Chugh, A. L. Hansen, and H. M. Al-Hashimi, Nat. Methods **8**, 919 (2011).
[18]A. T. Frank, A. C. Stelzer, H. M. Al-Hashimi, and I. Andricioaei, Nucl. Acids Res. **37**, 3670 (2009).
[19]B. Furtig, J. Buck, V. Manoharan, W. Bermel, A. Jaschke, P. Wenter, S. Pitsch, and H. Schwalbe, Biopolymers **86**, 360 (2007).
[20]M. Getz, X. Y. Sun, A. Casiano-Negroni, Q. Zhang, and H. M. Al-Hashimi, Biopolymers **86**, 384 (2007).
[21]B. Richter, J. Gsponer, P. Varnai, X. Salvatella, and M. Vendruscolo, J. Biomol. NMR **37**, 117 (2007).
[22]Z. Shajani and G. Varani, Biopolymers **86**, 348 (2007).
[23]O. F. Lange, N. A. Lakomek, C. Fares, G. F. Schroder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, and B. L. de Groot, Science **320**, 1471 (2008).
[24]P. R. L. Markwick, G. Bouvignies, L. Salmon, J. A. McCammon, M. Nilges, and M. Blackledge, J. Am. Chem. Soc. **131**, 16968 (2009).
[25]B. Vogeli, S. Kazemi, P. Guntert, and R. Riek, Nat. Struct. Mol. Biol. **19**, 1053 (2012).
[26]C. Camilloni, P. Robustelli, A. De Simone, A. Cavalli, and M. Vendruscolo, J. Am. Chem. Soc. **134**, 3968 (2012).
[27]T. Mittag and J. D. Forman-Kay, Curr. Opin. Struct. Biol. **17**, 3 (2007).
[28]M. Vendruscolo, Curr. Opin. Struct. Biol. **17**, 15 (2007).
[29]J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard, Proc. Natl. Acad. Sci. U.S.A. **92**, 9279 (1995).
[30]N. Tjandra and A. Bax, Science **278**, 1111 (1997).
[31]G. Bouvignies, P. Markwick, R. Bruschweiler, and M. Blackledge, J. Am. Chem. Soc. **128**, 15100 (2006).
[32]J. R. Huang and S. Grzesiek, J. Am. Chem. Soc. **132**, 694 (2010).
[33]A. C. Stelzer, A. T. Frank, M. H. Bailor, I. Andricioaei, and H. M. Al-Hashimi, Methods **49**, 167 (2009).
[34]A. Cavalli, C. Camilloni, and M. Vendruscolo, J. Chem. Phys. **138**, 094112 (2013).
[35]J. W. Pitera and J. D. Chodera, J. Chem. Theor. Comput. **8**, 3445 (2012).
[36]B. Roux and J. Weare, J. Chem. Phys. **138**, 084107 (2013).
[37]A. De Simone, R. W. Montalvao, and M. Vendruscolo, J. Chem. Theor. Comput. **7**, 4189 (2011).
[38]R. W. Montalvao, A. De Simone, and M. Vendruscolo, J. Biomol. NMR **53**, 281 (2012).
[39]F. Aboul-ela, J. Karn, and G. Varani, Nucl. Acids Res. **24**, 3974 (1996).
[40]A. Davidson, D. W. Begley, C. Lau, and G. Varani, J. Mol. Biol. **410**, 984 (2011).
[41]B. Davis, M. Afshar, G. Varani, A. I. H. Murchie, J. Karn, G. Lentzen, M. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. Swarbrick, and F. Aboul-ela, J. Mol. Biol. **336**, 343 (2004).
[42]J. A. Ippolito and T. A. Steitz, Proc. Natl. Acad. Sci. U.S.A. **95**, 9819 (1998).
[43]A. I. H. Murchie, B. Davis, C. Isel, M. Afshar, M. J. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. M. Swarbrick, S. Mirza, C. D. Prescott, P. Vaglio, F. Aboul-ela, and J. Karn, J. Mol. Biol. **336**, 625 (2004).
[44]Q. Zhang, A. C. Stelzer, C. K. Fisher, and H. M. Al-Hashimi, Nature (London) **450**, 1263 (2007).
[45]A. Almond and J. B. Axelsen, J. Am. Chem. Soc. **124**, 9986 (2002).
[46]H. F. Azurmendi and C. A. Bush, J. Am. Chem. Soc. **124**, 2426 (2002).
[47]K. Berlin, D. P. O'Leary, and D. Fushman, J. Magn. Res. **201**, 25 (2009).
[48]M. X. Fernandes, P. Bernado, M. Pons, and J. G. de la Torre, J. Am. Chem. Soc. **123**, 12037 (2001).
[49]A. Ferrarini, J. Phys. Chem. B **107**, 7923 (2003).
[50]F. van Lune, L. Manning, K. Dijkstra, H. J. C. Berendsen, and R. M. Scheek, J. Biomol. NMR **23**, 169 (2002).
[51]M. Zweckstetter, Nat. Protoc. **3**, 679 (2008).
[52]M. Zweckstetter and A. Bax, J. Am. Chem. Soc. **122**, 3791 (2000).
[53]B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theor. Comput. **4**, 435 (2008).
[54]D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).
[55]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[56]N. Foloppe and A. D. MacKerell, J. Comput. Chem. **21**, 86 (2000).
[57]A. D. MacKerell, N. Banavali, and N. Foloppe, Biopolymers **56**, 257 (2001).
[58]A. D. MacKerell and N. K. Banavali, J. Comput. Chem. **21**, 105 (2000).
[59]A. D. Mackerell, J. Comput. Chem. **25**, 1584 (2004).
[60]P. Banas, D. Hollas, M. Zgarbova, P. Jurecka, M. Orozco, T. E. Cheatham, J. Sponer, and M. Otyepka, J. Chem. Theor. Comput. **6**, 3836 (2010).
[61]D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**, 1668 (2005).
[62]V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, Proteins **65**, 712 (2006).
[63]A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, Biophys. J. **92**, 3817 (2007).
[64]E. J. Sorin and V. S. Pande, Biophys. J. **88**, 2472 (2005).
[65]See supplementary material at http://dx.doi.org/10.1063/1.4804301 for the results on the potential ligand efficiency values.
[66]R. E. Dickerson, Nucl. Acids Res. **17**, 1797 (1989).
[67]M. Bansal, D. Bhattacharyya, and B. Ravi, CABIOS, Comput. Appl. Biosci. **11**, 281 (1995).
[68]X. J. Lu and W. K. Olson, Nucl. Acids Res. **31**, 5108 (2003).
[69]X. J. Lu and W. K. Olson, Nat. Protoc. **3**, 1213 (2008).
[70]G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax, J. Am. Chem. Soc. **120**, 6836 (1998).
[71]G. M. Clore, M. R. Starich, and A. M. Gronenborn, J. Am. Chem. Soc. **120**, 10571 (1998).
[72]M. Ottiger and A. Bax, J. Biomol. NMR **13**, 187 (1999).
[73]M. H. Bailor, A. M. Mustoe, C. L. Brooks, and H. M. Al-Hashimi, Nat. Protoc. **6**, 1536 (2011).