Insights into Domain−Domain Motions in Proteins and RNA from Solution NMR

Enrico Ravera,[†,‡] Loïc Salmon,[§] Marco Fragai,[†,‡] Giacomo Parigi,[†,‡] Hashim Al-Hashimi,*[,‖] and Claudio Luchinat*[,†,‡]

[†]CERM, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Italy

[‡]Department of Chemistry "U. Schiff", University of Florence, via della Lastruccia 3, 50019, Sesto Fiorentino, Italy

[§]Department of Biophysics, University of Michigan, 830 N. University, Ann Arbor, Michigan 48109, United States

[‖]Department of Biochemistry and Department of Chemistry, Duke University School of Medicine, 307 Research Drive, Durham, North Carolina 27710, United States

CONSPECTUS: Many multidomain proteins and ribonucleic acids consist of domains that autonomously fold and that are linked together by flexible junctions. This architectural design allows domains to sample a wide range of positions with respect to one another, yet do so in a way that retains structural specificity, since the number of sampled conformations remains extremely small compared to the total conformations that would be sampled if the domains were connected by an infinitely long linker. This "tuned" flexibility in interdomain conformation is in turn used in many biochemical processes.

There is great interest in characterizing the dynamic properties of multidomain systems, and moving beyond conventional descriptions in terms of static structures, toward the characterization of population-weighted ensembles describing a distribution of many conformations sampled in solution. There is also great interest in understanding the design principles and underlying physical and chemical interactions that specify the nature of interdomain flexibility. NMR spectroscopy is one of the most powerful techniques for characterizing motions in complex biomolecules and has contributed greatly toward our basic understanding of dynamics in proteins and nucleic acids and its role in folding, recognition, and signaling.

Here, we review methods that have been developed in our laboratories to address these challenges. Our approaches are based on the ability of one domain of the molecule to self-align in a magnetic field, or to dominate the overall orientation of the molecule, so that the conformational freedom of other domains can be assessed by their degree of alignment induced by the aligned part. In turn, this self-alignment ability can be intrinsic or can be caused by tagging appropriate constructs to the molecule of interest. In general, self-alignment is due to magnetic susceptibility anisotropy. Nucleic acids with elongated helices have this feature, as well as several paramagnetic metal centers that can be found in, or attached to, a protein domain.
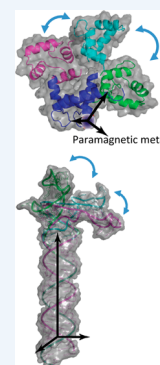
INTRODUCTION

Many biochemical processes are based on the possibility for one or more of the participating molecules to adopt different conformations, while retaining some structural specificity.[1,2] This is the case for systems consisting of independent domains linked by flexible junctions, such as multidomain proteins and ribonucleic acids (RNA). For example, changes in the relative orientation of protein domains make it possible to create distinct binding surfaces for intermolecular interactions with many different binding partners.[3] Likewise, multistep changes in the orientation of RNA A-form helices carrying catalytic residues make it possible for one ribozyme to adopt the conformations that are required in multistep catalytic cycles.[4,5] NMR can uniquely provide site-specific information on interdomain motions over a broad range of biologically relevant time scales, from picoseconds to milliseconds.[6−9]

Sampling multiple conformational states leads to the averaging of the experimental observables, and, while it is possible to calculate the average observables given any structural ensemble, there is an infinite number of ensembles that might equally account for the average experimental observables,[10] even in the absence of experimental errors. Several approaches based on the creation of "optimized" conformational ensembles have been proposed.[2,11−15]

Here, we review the methods that have been applied in our laboratories to address these challenges. The approaches are based on the ability of one domain of the molecule to self-align in a magnetic field. Different approaches, still based on the use of residual dipolar couplings (RDCs), have been developed in other laboratories to obtain information on the conformational variability of the investigated systems. For these, readers can refer, for instance, to the review by Tolman et al.[16] and references therein.

RDC ANALYSIS OF DOMAIN MOTIONS BY ANCHORING OVERALL ALIGNMENT FRAMES ONTO INDIVIDUAL DOMAINS

In the presence of a magnetic field, the nuclear spin energy levels are mainly determined by the interaction between the

nuclear magnetic moments and the external magnetic field (Zeeman effect), and modulated by interactions of the nuclear magnetic moments with additional, molecule-specific magnetic and/or electric fields. Fast isotropic reorientation of molecules in solution cancels the anisotropic part of these interactions (Table 1), simplifying the NMR spectrum to sharp lines

Table 1. Summary of Phenomena Involving Nuclear Spins and Their Manifestations in Isotropic and Anisotropic Solutions

effect	due to interaction of the nuclear spin with	isotropic average	result in the presence of partial alignment
chemical shielding	electron currents in orbitals	isotropic chemical shift	residual chemical shift anisotropy[59]
dipolar shielding	average electron magnetic moment	pseudocontact shift	residual dipolar shift[60]
quadrupolar coupling	electric field gradients	(average to zero)	residual quadrupolar coupling[61]
dipole–dipole coupling	magnetic moments of the neighboring nuclei	(average to zero)	residual dipolar coupling[62]

centered at the average value of the chemical shielding interactions and split by the scalar coupling with covalently bound spins. The relevant structural and dynamical information encoded in the anisotropic interactions is lost, but it can be partially recovered by making anisotropic the distribution of molecular orientations. This can be achieved either by dissolving biomolecules in ordering media[17,18] or when the molecules themselves have some preferred orientations in the presence of a high magnetic field, due to intrinsic magnetic susceptibility anisotropy.[19] Many structured nucleic acid fragments spontaneously align in magnetic fields,[20] as well as several paramagnetic metal centers that can be found in, or attached to, a protein.[21−26]

In diamagnetic systems, the anisotropy of the magnetic susceptibility tensor χ is due to the interaction of the magnetic field with the motion of the electrons in their orbitals. It is usually small in biomolecules, except, for instance, in heme-containing proteins or when multiple aromatic planes are stacked together, as in double stranded nucleic acids.[27]

Inducing partial alignment in a molecule reintroduces anisotropic interactions, including RDCs:[8,16] when not all orientations have the same probability, the dipole–dipole interactions between nuclear magnetic moments does not average to zero anymore. RDCs measured between two nuclei i and j carry information regarding the orientation distribution of their internuclear vector relative to the applied magnetic field, averaged over all orientations sampled at a rate roughly faster than milliseconds.[8,16]

The RDCs measured in a partially aligned rigid-body can be fully accounted for by describing separately the property of alignment and the local geometry. The degree, the asymmetry and the direction of alignment of the molecule can be described using a traceless symmetric tensor, called alignment tensor. Information about structure or internal motions can then be recovered.[8,16] While this "decoupling approximation" holds true for many globular proteins, it breaks down in multidomain proteins and RNA:

(a) Changes in the arrangement of domains can result in significant changes in the shape of a molecule, coupling the interdomain conformational freedom to the diffusion property of the molecule[20,28] (Figure 1).

(b) Measurement of RDCs generally requires dissolution into an ordering medium with which the domains may have differential interactions, and the electric potential of a charged alignment medium may perturb differently the different conformations of the biomolecule.

(c) Different alignment media, to access multiple independent sets of RDCs, can result in differential perturbations on the conformational freedom. Furthermore, in the case of RNA, changing the ordering medium generally does not yield the desired independent RDCs.[29]

These challenges can be overcome using approaches based on anchoring the overall alignment tensor frame on specific domains through paramagnetic alignment or helix elongation (Figure 2).

Paramagnetic Alignment

In this approach, partial self-orientation is induced by a paramagnetic metal ion in a protein domain (Figure 2A), either naturally present or introduced by substituting a diamagnetic metal, or by attachment of a paramagnetic tag. The orientation induced by the presence of an anisotropic paramagnetic center can be modulated by introducing different ions in the molecule.[25,30,31] In these cases, the magnetic susceptibility anisotropy is caused by the anisotropy of the average electron magnetic moment induced by a magnetic field.

Such anisotropy also causes a shift in the NMR signals, called pseudocontact shift (PCS).[30,31] PCSs originate from the nonzero averaging upon rotation of the dipolar interaction between the nuclear magnetic moment and the average induced electron magnetic moment, and they provide an independent measure of the susceptibility anisotropy tensor. Furthermore, they depend on the position of each detected nucleus with respect to the paramagnetic center and its anisotropy frame, thus providing additional data related to the relative position of the protein domains.

The first paramagnetic RDCs were measured on the protein cyanometmyoglobin.[19] Paramagnetic RDCs were first used as structural restraints for cytochrome b_5[32] and, since then, have been largely used for the calculation of protein structures.[33−35]

For multidomain proteins, partial self-orientation is induced on the domain to which the metal is attached, and RDCs reflect the orientation distribution of domains relative to the reference metal-containing domain.

Helix Elongation

In this approach, applied to RNA, a domain is engineered such that its shape dominates the overall orientation of the molecule (Figure 2B):[11,36] a helix is elongated to create an extended shape. The overall alignment of such a RNA, either dissolved in an ordering medium, or self-aligned by its diamagnetic susceptibility, is dictated by the elongated helix. As a result, RDCs can be interpreted in terms of motions of the other helices relative to the elongated one. Furthermore, introducing kinks along various positions of the elongated helix modulates the alignment of the helix itself, regaining access to multiple independent sets of RDCs. Finally, by elongating different helices, one can anchor the NMR frame along different domains and thereby measure changes in orientation of domains relative to all domains.

Another similar approach for anchoring frames of reference onto specific sites of RNA involves installing protein binding sites and then adding a protein to modulate its alignment.[37]
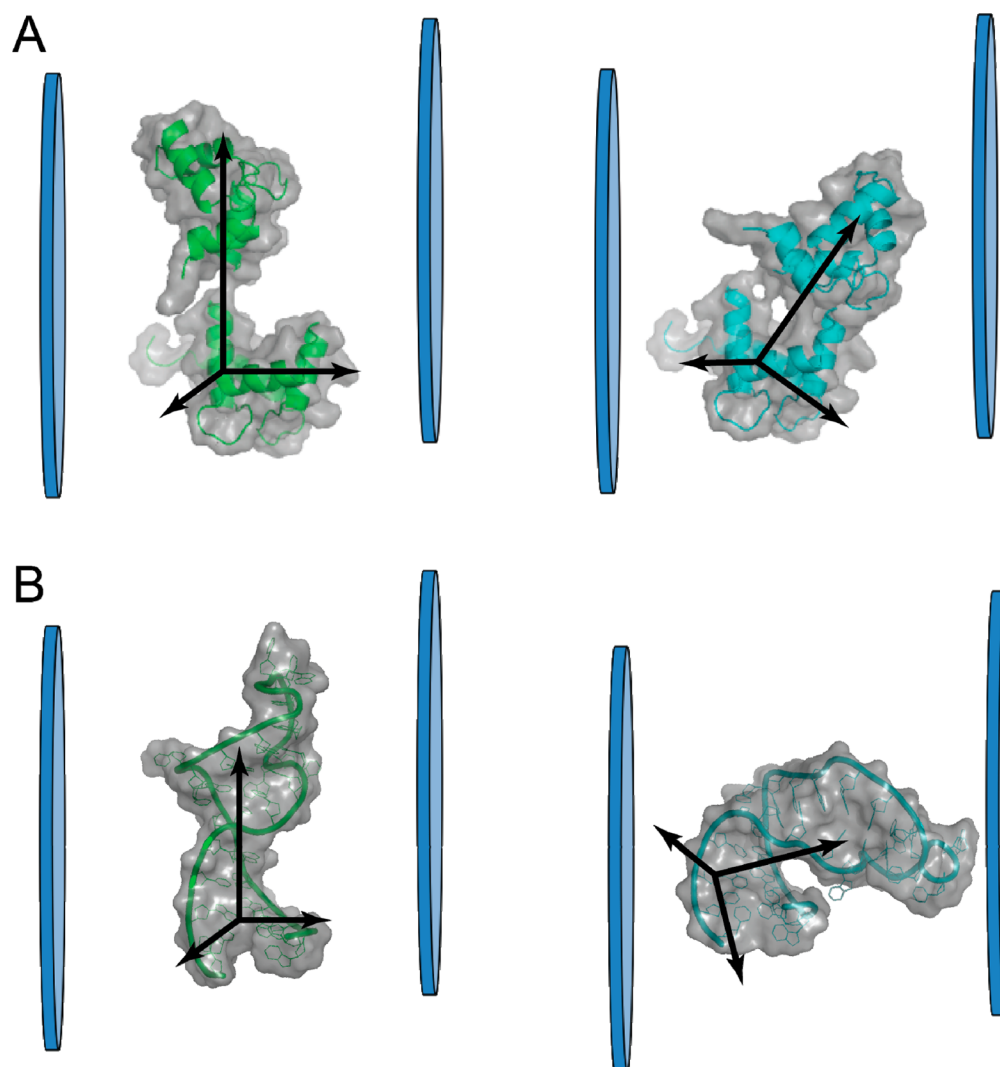
Figure 1. Coupling between internal motion and alignment properties in multidomain systems. Proteins (A) and RNA (B) exhibit a correlated change in shape and alignment properties.

Elongation or added protein should not perturb the structure and dynamics of the RNA, as can be monitored by the comparison of the NMR chemical shifts for the modified and unmodified RNA.

Although the degree of self-alignment $(10^{-4}-10^{-5})$ and magnitude of RDCs measured in magnetically aligned RNA is 1−2 orders of magnitude smaller than optimum (10^{-3}), self-alignment has a simple dependence on nucleic acid structure, which could be exploited to understand the effects of motional couplings. For a magnetically aligned nucleic acid, the alignment tensor is given by the overall diamagnetic susceptibility (χ) tensor that, to a good approximation, is given by a tensor summation over all χ-tensors associated with individual nucleobases.[20,38] This property makes it possible to relate the alignment tensor to the nucleic acid structure, and specifically to the orientation of nucleobases.

An alternative approach for treating correlations between conformations and overall alignment in systems dissolved in ordering media involves using programs such as PALES[39] or PATI[2] that compute the alignment based on molecular shape (and electrostatic properties). By treating correlations between internal and overall alignment and obviating the need to elongate helices extensively, this approach made it possible to

use variable degrees of linear and kinked elongations in the ensemble determination of two helices connected by a trinucleotide bulge in HIV-1 TAR.[40]

■ RDCs ANALYSIS IN TERMS OF INTERDOMAIN CONFORMATIONAL DISTRIBUTIONS

We have used two distinct approaches for extracting information from RDCs. Both methods rely on the experimental determination of the alignment tensor for all domains in the multidomain system. For a rigid system, all domains sense the same alignment tensor of the reference domain, that is, of the domain that dominates alignment or bears the alignment device; however, in the presence of interdomain conformational freedom, the alignment tensor sensed by the other domains will be averaged over the various relative arrangements of the two domains. Therefore, the extent of the conformational variability can be evaluated by comparing the alignment tensor of the reporter domain with that of the reference domain (Figure 3).[8,30,41] Notably, complete independence of the second domain would reduce its alignment tensor to zero. This does not hold true in the case of external alignment, especially for domains of similar size: it can be shown that the tensor after averaging can have the same
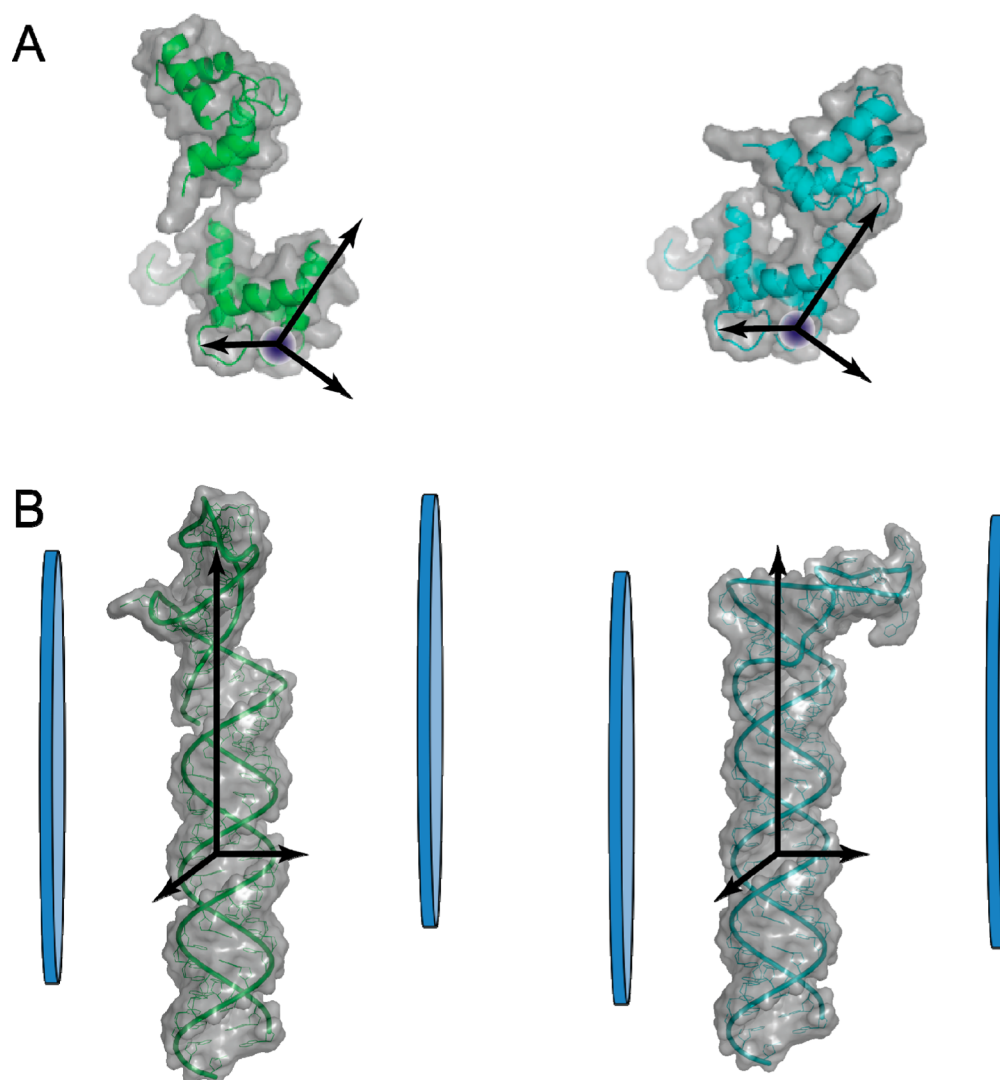
Figure 2. Decoupling between internal motion and alignment properties in multidomain systems (A) using a paramagnetic ion or (B) by elongating the terminal helix of a RNA.

magnitude of the alignment tensor determined for a single rigid conformation.[28]

Data Guided Selection of Conformers from a Pool

RDCs and other data are used to select conformations from a pool generated using computational methods (molecular dynamics,[14,40,42] enhanced sampling,[13] or Monte Carlo models[11,12]). This approach involves two steps: (i) generation of a pool of conformations that broadly sample the interdomain free energy landscape and (ii) use of experimental data to select a subensemble from the conformational pool.[14,43] This approach is sometimes referred to as "sample and select" (SAS, Figure 4A).[14]

The success of SAS-based approaches critically depends on the sampling of all allowed conformations in the starting pool.[15] This condition can generally be met for two-domain systems, for which all conformations can be generated from 3 rotational and 3 translational degrees of freedom and those which are chemically impossible to achieve (either because the linker is too short to maintain connectivity or because of severe steric clashes) are subsequently removed. In a second step, structures are selected from the conformational pool in order to reproduce the experimental data. The selection procedure can

be accomplished using a variety of search algorithms including simulated annealing[11,14,40] and genetic algorithms.[12,15] To construct the subensembles, N conformers are selected from the conformational pool to maximize the agreement between measured and predicted data. The ensemble size is then incrementally increased from $N = 1$ until either the experimental data are reproduced within experimental error or the agreement is not improved by adding more conformers (Figure 4A). This approach has to be followed by rigorous analysis and cross-validation.[15,44] The procedure can be repeated hundreds of times, with the family of conformations selected over all runs pooled together to obtain a final ensemble. Recent studies suggest that the SAS approach employing RDCs can be used to capture the statistical weights of dominant conformers in the ensemble.[45]

Maximum Occurrence Calculations

A second complementary approach, called "Maximum Occurrence" (MaxOcc), at variance with methods based on ensemble reconstruction, aims at finding the maximum percent of time that the system can spend in one given conformation and still be compatible with the experimental observations,
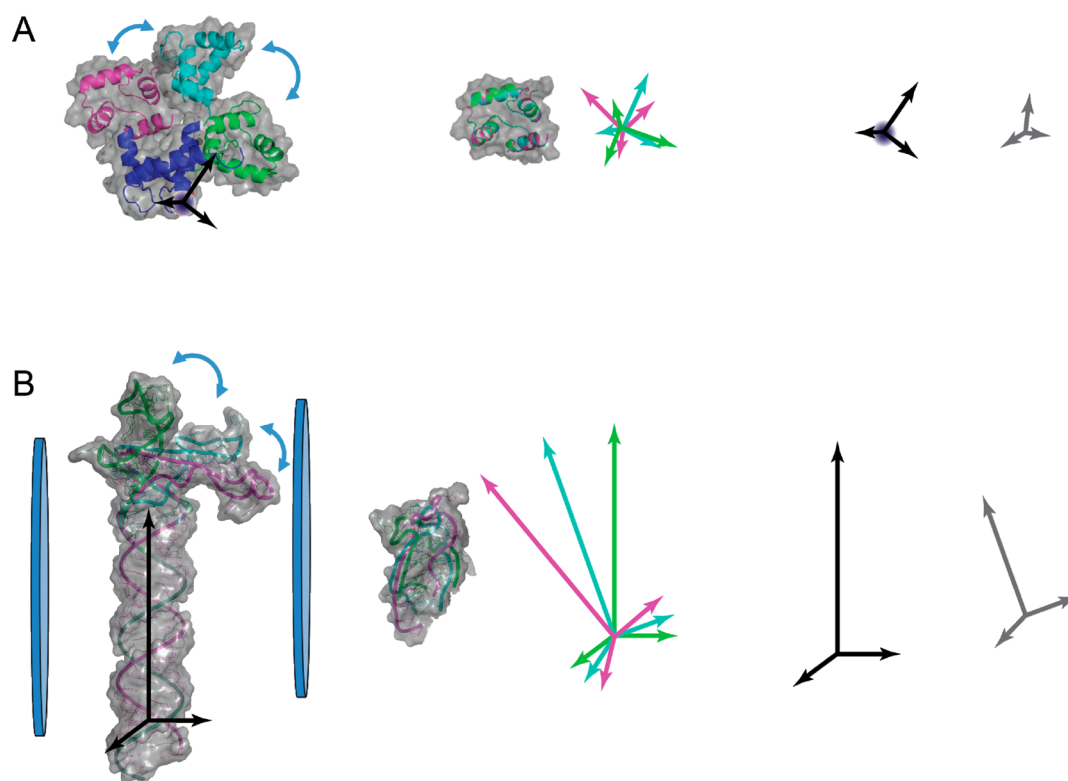
Figure 3. Averaged tensors resulting from the fit of the RDCs of domains mobile with respect to the reference domain: (A) A two-domain protein is shown with the first domain bearing a paramagnetic ion, depicted in blue, and a second domain in three different positions (in magenta, cyan, and green). The nuclei of the second domain see the magnetic susceptibility anisotropy tensor in three different orientations. Therefore, the RDCs depend on an averaged tensor (gray), resulting as the average of the three tensors and different from the real magnetic susceptibility anisotropy tensor (black). (B) Similar effect, presented for an elongated RNA.
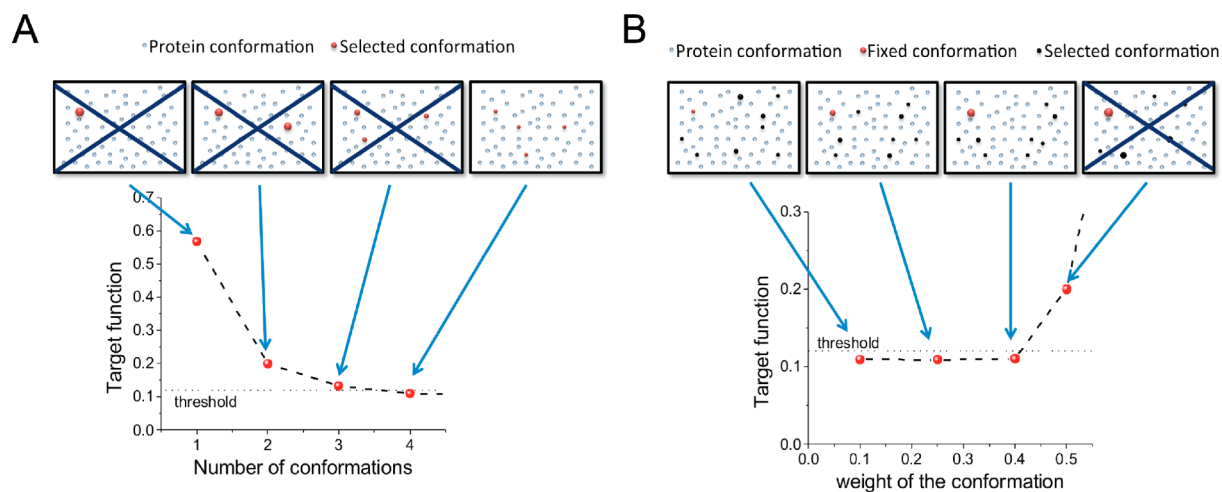

Figure 4. In the presence of conformational variability, ensembles of conformations must be considered for reproducing averaged data. (A) These ensembles are built by selecting protein conformations from a pregenerated pool of structures. The agreement between backcalculated and experimental data increases (lower target function) by increasing the number of conformations included in the ensemble, until a lower threshold is reached. In this example, ensembles of four structures are needed to fit the experimental data. (B) The Maximum Occurrence (MaxOcc) of a chosen conformation is obtained by searching for ensembles of structures which include the chosen (fixed) conformation, with different weights, together with other freely selected conformations. In this example, the MaxOcc is between 0.4 and 0.5.

when taken together with any optimized combination of other conformations.[30,46]

Also this method relies on a very broad pool of conformations, in order to map the whole conformational space that the system can sample.[47] MaxOcc calculations are performed separately for any conformation of interest. The calculations are done by selecting an ensemble which includes such conformation with a fixed weight, and tens of other conformations providing averaged data in best agreement with the experimental data. These calculations are repeated for increasing weights of the selected conformation, until it becomes impossible to find an ensemble in agreement with
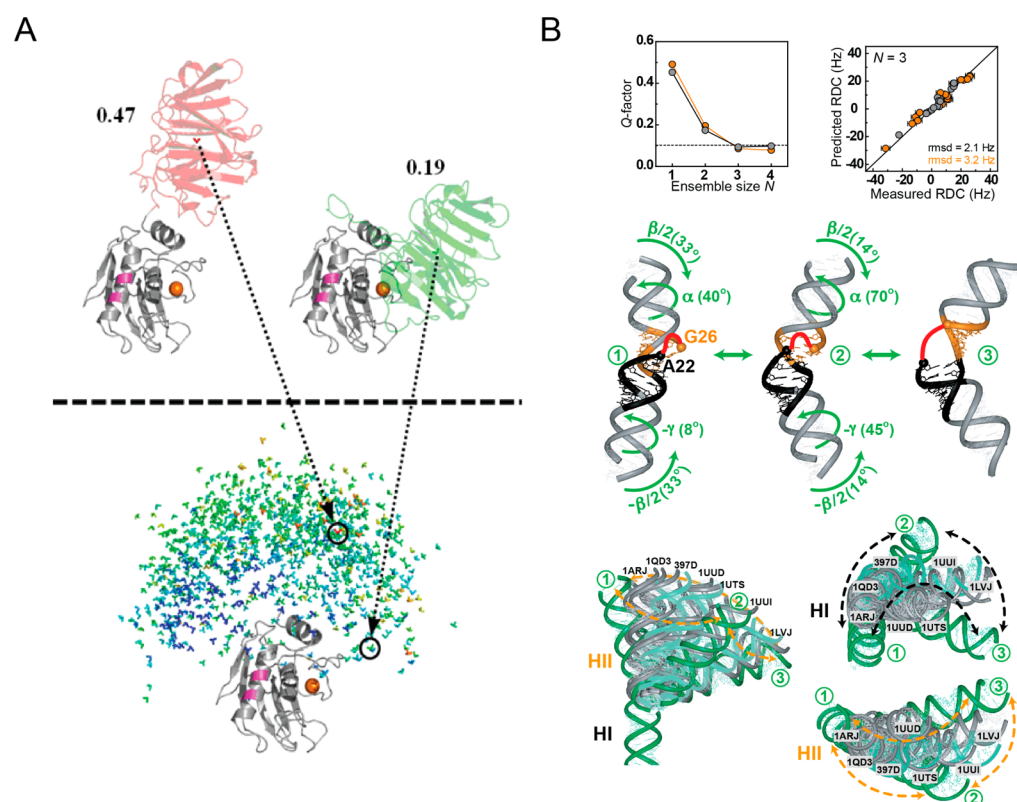
Figure 5. (A) Two conformations of MMP-1 with MaxOcc equal to 0.19 and 0.47. In the bottom, the conformations are displayed superimposed on the catalytic domain (in gray), which bears the paramagnetic (Ln)CLaNP-5 bound to the residues in pink, by representing the hemopexin domain with a 3-axes frame, positioned in its center of mass and color-coded according to the corresponding MaxOcc, from blue (0.05) to red (0.47). (B) Three-state description of HIV-1 TAR dynamics: (top) Capability to reproduce the experimental RDCs for different ensemble size (left) and for the final ensemble (right); (middle) the three conformers constituting the ensemble; (bottom) comparison of the ensemble (green) with known ligand bound TAR conformations (gray).

the experimental data, thus determining the MaxOcc value of that conformation (Figure 4B). Once the MaxOcc values, that is, the largest possible weights, are determined for a large number of conformations, it is possible to identify the conformations which must necessarily have a negligibly small weight and those which may have a large weight.

APPLICATIONS TO PROTEINS

A widely studied example of a flexible two domain protein is calmodulin, a protein composed of two domains connected by a flexible linker.[48] The extensive conformational variability of free calmodulin is testified by the sizable reduction of the RDC-derived anisotropy tensor of the C-terminal domain with respect to the anisotropy tensor of the N-terminal domain, where a paramagnetic lanthanide ion is introduced.[30,46]

The MaxOcc approach was applied to characterize the conformational variability experienced by calmodulin both when free in solution[22,46] or bound to intrinsically disordered proteins.[49] Three sets of PCSs and RDCs were used, by substituting one of the calcium(II) ions in the N-terminal domain with lanthanides. In these cases, the protein samples a large ensemble of conformations, as no single structure, or ensemble of structurally similar conformations, agrees with the experimental data. Other restraints like paramagnetic relaxation enhancements[50] or SAXS data[46] were also included for better discrimination of the MaxOcc values. This approach provided a picture of the regions in the conformational space which can be mostly sampled by the protein and of those that can only be sampled to a limited extent.

Another remarkable two-domain protein example is matrix metalloproteinase 1 (MMP-1), an enzyme that can cleave collagen despite the fact that collagen's quaternary and superquaternary organization conceals the cleavage site.[51] The key resides in the relative motions of the two domains of the protein, that open to accommodate the substrate, then come closer again to unwind the triple-helix, and finally accomplish the cleavage.[51] The MaxOcc approach was used to characterize the conformational variability of MMP-1 (Figure 5A) by rigidly attaching a paramagnetic tag[21] to the catalytic domain.[10] Again, the mean tensors determined for the hemopexin domain are significantly smaller than the susceptibility anisotropy tensors determined from the catalytic domain, pointing to some conformational averaging. The reduction was less dramatic than for calmodulin, and indeed the MMP-1 conformations with the highest MaxOcc are clustered in a relatively restricted region.[10] This corresponds to protein structures very different from the crystal structure and much more extended but, strikingly, not distant from the conformation that MMP-1 was proposed to adopt when binding the collagen substrate in the first step of the collagenolytic mechanism.[51]

APPLICATIONS TO RNAs

By independently elongating two helices in HIV-1 TAR, it was possible to anchor the NMR frame for RDC to each of the two helices HI and HII (Figure 5B).[11] The RDCs carried the

required sensitivity to all three Euler angles defining interhelical orientation. Using these RDCs, an ensemble was determined using the SAS approach and a grid search was performed over sterically allowed conformations. A static representation of the two helices is incompatible with the RDCs, and an ensemble consisting of a minimum of three equally populated states is required. A striking feature of the RDC-derived ensemble was that the three conformations fell nearly along a straight line in the 3D interhelix Euler space defining twisting around each helix and interhelical bending. Thus, although the helices HI and HII undergo large amplitude collective motions (>90°) relative to one another, and they appear to move in a very specific and directional manner. This was a clear sign of "directional flexibility" in RNA, and interestingly, the three-state ensemble enveloped many of the known ligand-bound TAR conformations, indicating that, on its own, RNA is capable of sampling a variety of conformations that are stabilized on ligand binding (Figure 5B). Subsequent works[52,53] showed that the molecular basis for these large and directional interhelical motions consists of topological constraints (steric and connectivity) that play essential roles in RNA folding and conformational adaptation.[54−56]

The same two sets of RDCs measured in HIV-1 TAR were used to determine atomic-resolution ensembles using the SAS approach and a conformational pool derived from a 80 ns MD trajectory of HIV-1 TAR[42] computed using CHARMM.[57] Although some correlation was observed between the measured and predicted RDCs for both EI-TAR and EII-TAR, the deviations were substantially larger than the estimated uncertainty. However, the simulation time was not long enough to match the RDC time scale (milliseconds), and this failure to predict the RDCs could not be considered an evidence for a poor force field. Using the SAS approach, an ensemble of $N = 20$ conformations was constructed that satisfies the measured RDCs. The RDC-derived TAR ensemble was qualitatively cross-validated using independent NMR measurements that were not included in the ensemble determination including NOEs and trans-hydrogen bond scalar couplings. It featured very similar correlated variations in the interhelical bend angle as observed with the three-state ensemble of TAR but, importantly, it also allowed the visualization of local motions in and around the bulge. More recently,[40] a SAS approach was used in which PALES is used to back-predict RDCs to construct an ensemble for TAR using four independent sets of RDCs measured in four differentially elongated TAR samples, and a broad pool of conformations derived from a much longer 8.2 μs MD trajectory. The approach allowed to directly treat the coupling between internal motion and alignment, and to use a construct in which the alignment is not dominated by a given domain. The ensemble showed similar interhelical distributions as determined previously, but also showed that large transitions in interhelical orientation are coupled to local melting of base-pairs near the junction. The RDC-selected ensemble included conformations that bear strong resemblance to the ligand bound conformations of TAR, including with regards to the details of binding pocket near the bulge, again indicating that intrinsic motions specify the TAR ligand bound conformations.[42] In later studies, the dynamic ensemble was targeted using virtual screening yielding new compounds that bind TAR and inhibit HIV replication,[58] illustrating one example of a biomedical application involving conformational ensembles.

CONCLUSIONS AND PERSPECTIVES

The modular design of biomolecules as beads on a string is widely used by nature to create robust biomolecular systems that are endowed with specific conformational flexibility. The latter can be tuned to carry out biochemical processes that require not one but a range of conformations. These systems present unique challenges to NMR structural and dynamic characterization that we have sought to address in proteins and RNA by anchoring frames of reference onto individual domains through paramagnetic tagging and elongation of helical domains. Such approaches have made it possible to disentangle contributions to common NMR parameters such as RDCs due to internal and overall motions, and thereby to quantitatively characterize interdomain motions in terms of some form of a probability distribution. Finally, while we have focused on two domain systems, there is a need to address the behavior of systems containing a larger number of domains, in both proteins and RNA, where we expect to see new behaviors and complexities.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: claudioluchinat@cerm.unifi.it.
*E-mail: hashim.al.hashimi@duke.edu.

Notes

Biographies

Enrico Ravera obtained a B.Sc. in Chemistry in 2008, a M.Sc. in Chemistry cum laude in 2009, and a Ph.D. in Chemistry in 2013 at the University of Florence, Italy. He is currently a postdoctoral fellow under the supervision of Prof. Claudio Luchinat.

Loïc Salmon obtained his Ph.D. in 2010 in Grenoble, under the supervision of Dr. Martin Blackledge. He then moved to the University of Michigan as a postdoctoral fellow in the lab of Prof. Hashim Al-Hashimi. He is currently postdoctoring in the lab of Prof. James Bardwell, where he is investigating the impact of conformational flexibility in chaperone protein biological function.

Marco Fragai graduated in medicinal chemistry cum laude and obtained his Ph.D. in chemistry at the University of Florence, Italy. He was a postdoctoral associate until 2005 and then a researcher at the University of Florence. His research interests include drug design and applications of NMR techniques in drug discovery.

Giacomo Parigi graduated in physics and obtained his Ph.D. in chemistry at the University of Florence, Italy. He was a postdoctoral associate and researcher and is Associate Professor of Chemistry since 2006 at the University of Florence. His research interests include NMR effects related to paramagnetism for structural and dynamic characterization of biomolecules and to nuclear and electron relaxation.

Hashim Al-Hashimi received his doctorate in Biophysical Chemistry from Yale University in 2000. He was a postdoctoral fellow at the Memorial Sloane-Kettering Cancer Center in NYC (2000−2002). He was assistant (2002−2008), associate (2008−2009), and full Professor (2009−2014) at the University of Michigan, and is Professor at Duke University since January 2014. He is the recipient of the National Science Foundation Career Award, LSA Excellence in Teaching Award, Ralph E. Powe Junior Faculty Enhancement Award, Founder's Medal, the Vilcek Prize for Creative Promise in Biomedical Science, and the Agilent Thought Leader Award.

Claudio Luchinat obtained his doctorate in chemistry cum laude at the University of Florence, was a researcher at the University of Florence and Full Professor of Chemistry at the University of Bologna (1986−1996) and is now at the University of Florence (1996 to the present). He is the recipient of the "Raffaello Nasini" gold medal award of the Italian Chemical Society; Federchimica Award "For an Intelligent Future"; European Medal for Biological Inorganic Chemistry; and "GDRM gold medal for magnetic resonance".

■ REFERENCES

(1) Schroeder, R.; Barta, A.; Semrad, K. Strategies for RNA Folding and Assembly. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 908−919.

(2) Berlin, K.; Castañeda, C. A.; Schneidman-Dohovny, D.; Sali, A.; Nava-Tudela, A.; Fushman, D. Recovering a Representative Conformational Ensemble From Underdetermined Macromolecular Structural Data. *J. Am. Chem. Soc.* **2013**, *135*, 16595−16609.

(3) Bornberg-Bauer, E.; Beaussart, F.; Kummerfeld, S. K.; Teichmann, S. A.; Weiner, J., III. The Evolution of Domain Arrangements in Proteins and Interaction Networks. *Cell. Mol. Life Sci.* **2005**, *62*, 435−445.

(4) Williamson, J. R. Induced Fit in RNA-Protein Recognition. *Nat. Struct. Biol.* **2000**, *7*, 834−837.

(5) Dethoff, E. A.; Chugh, J.; Mustoe, A. M.; Al Hashimi, H. M. Functional Complexity and Regulation Through RNA Dynamics. *Nature* **2012**, *482*, 322−330.

(6) Fragai, M.; Luchinat, C.; Parigi, G. "Four-Dimensional" Protein Structures: Examples From Metalloproteins. *Acc. Chem. Res.* **2006**, *39*, 909−917.

(7) Bothe, J. R.; Nikolova, E. N.; Eichhorn, C. D.; Chugh, J.; Hansen, A. L.; Al Hashimi, H. M. Characterizing RNA Dynamics at Atomic Resolution Using Solution-State NMR Spectroscopy. *Nat. Methods* **2011**, *8*, 919−931.

(8) Blackledge, M. Recent Progress in the Study of Biomolecular Structure and Dynamics in Solution From Residual Dipolar Couplings. *Prog. NMR Spectrosc.* **2005**, *46*, 23−61.

(9) Rinnenthal, J.; Buck, J.; Ferner, J.; Wacker, A.; Furtig, B.; Schwalbe, H. Mapping the Landscape of RNA Dynamics With NMR Spectroscopy. *Acc. Chem. Res.* **2011**, *44*, 1292−1301.

(10) Cerofolini, L.; Fields, G. B.; Fragai, M.; Geraldes, C. F. G. C.; Luchinat, C.; Parigi, G.; Ravera, E.; Svergun, D. I.; Teixeira, J. M. C. Examination of Matrix Metalloproteinase-1 (MMP-1) in Solution: a Preference for the Pre-Collagenolysis State. *J. Biol. Chem.* **2013**, *288*, 30659−30671.

(11) Zhang, Q.; Stelzer, A. C.; Fisher, C. K.; Al-Hashimi, H. M. Visualizing Spatially Correlated Dynamics That Directs RNA Conformational Transitions. *Nature* **2007**, *450*, 1263−1267.

(12) Nodet, L.; Salmon, L.; Ozenne, V.; Meier, S.; Jensen, M. R.; Blackledge, M. Quantitative Description of Backbone Conformational Sampling of Unfolded Proteins at Amino Acid Resolution From NMR Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2009**, *131*, 17908−17918.

(13) Guerry, P.; Salmon, L.; Mollica, L.; Ortega Roldan, J. L.; Markwick, P.; van Nuland, N. A.; McCammon, J. A.; Blackledge, M. Mapping the Population of Protein Conformational Energy Sub-States From NMR Dipolar Couplings. *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 3181−3185.

(14) Chen, Y.; Campbell, S. L.; Dokholyan, N. V. Deciphering Protein Dynamics From NMR Data Using Explicit Structure Sampling and Selection. *Biophys. J.* **2007**, *93*, 2300−2306.

(15) Salmon, L.; Yang, S.; Al Hashimi, H. M. Advances in the Determination of Nucleic Acid Conformational Ensembles. *Annu. Rev. Phys. Chem.* **2014**, *65*, 293−316.

(16) Tolman, J. R.; Ruan, K. NMR Residual Dipolar Couplings As Probes of Biomolecular Dynamics. *Chem. Rev.* **2006**, *106*, 1720−1736.

(17) Tjandra, N.; Bax, A. Direct Measurement of Distances and Angles in Biomolecules by NMR in a Diluite Liquid Crystalline Medium. *Science* **1997**, *278*, 1111−1114.

(18) Hansen, M. R.; Mueller, L.; Pardi, A. Tunable Alignment of Macromolecules by Filamentous Phage Yields Dipolar Coupling Interactions. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.

(19) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. Nuclear Magnetic Dipole Interactions in Field-Oriented Proteins: Information for Structure Determination in Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.

(20) Zhang, Q.; Throolin, R.; Pitt, S. W.; Serganov, A.; Al Hashimi, H. M. Probing Motions Between Equivalent RNA Domains Using Magnetic Field Induced Residual Dipolar Couplings: Accounting for Correlations Between Motions and Alignment. *J. Am. Chem. Soc.* **2003**, *125*, 10530−10531.

(21) Keizers, P. H. J.; Saragliadis, A.; Hiruma, Y.; Overhand, M.; Ubbink, M. Design, Synthesis, and Evaluation of a Lanthanide Chelating Protein Probe: CLaNP-5 Yields Predictable Paramagnetic Effects Independent of Environment. *J. Am. Chem. Soc.* **2008**, *130*, 14802−14812.

(22) Das Gupta, S.; Hu, X.; Keizers, P. H. J.; Liu, W.-M.; Luchinat, C.; Nagulapalli, M.; Overhand, M.; Parigi, G.; Sgheri, L.; Ubbink, M. Narrowing the Conformational Space Sampled by Two-Domain Proteins With Paramagnetic Probes in Both Domains. *J. Biomol. NMR* **2011**, *51*, 253−263.

(23) Bertini, I.; Calderone, V.; Cerofolini, L.; Fragai, M.; Geraldes, C. F. G. C.; Hermann, P.; Luchinat, C.; Parigi, G.; Teixeira, J. M. C. The Catalytic Domain of MMP-1 Studied Through Tagged Lanthanides. Dedicated to Prof. A.V. Xavier. *FEBS Lett.* **2012**, *586*, 557−567.

(24) Ma, C.; Opella, S. J. Lanthanide Ions Bind Specifically to an Added "EF-Hand" and Orient a Membrane Protein in Micelles for Solution NMR Spectroscopy. *J. Magn. Reson.* **2000**, *146*, 381−384.

(25) Pintacuda, G.; John, M.; Su, X. C.; Otting, G. NMR Structure Determination of Protein-Ligand Complexes by Lanthanide Labeling. *Acc. Chem. Res.* **2007**, *40*, 206−212.

(26) Wöhnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. Protein Alignment by a Coexpressed Lanthanide-Binding Tag for the Measurement of Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2003**, *125*, 13338−13339.

(27) Clore, G. M.; Gronenborn, A. M. NMR Structure Determination of Proteins and Protein Complexes Larger Than 20 KDa. *Curr. Opin. Chem. Biol.* **1998**, *2*, 564−570.

(28) Yuwen, T.; Post, C. B.; Skrynnikov, N. R. Domain Cooperativity in Multidomain Proteins: What Can We Learn From Molecular Alignment in Anisotropic Media? *J. Biomol. NMR* **2011**, *51*, 131−150.

(29) Latham, M. P.; Hanson, P.; Brown, D. J.; Pardi, A. Comparison of Alignment Tensors Generated for Native TRNA(Val) Using Magnetic Fields and Liquid Crystalline Media. *J. Biomol. NMR* **2008**, *40*, 83−94.

(30) Fragai, M.; Luchinat, C.; Parigi, G.; Ravera, E. Conformational Freedom of Metalloproteins Revealed by Paramagnetism-Assisted NMR. *Coord. Chem. Rev.* **2013**, *257*, 2652−2667.

(31) Bertini, I.; Luchinat, C.; Parigi, G. Moving the Frontiers in Solution Solid State BioNMR. A Celebration of Harry Gray's 75th Birthday. *Coord. Chem. Rev.* **2011**, *255*, 649−663.

(32) Banci, L.; Bertini, I.; Huber, J. G.; Luchinat, C.; Rosato, A. Partial Orientation of Oxidized and Reduced Cytochrome B_5 at High Magnetic Fields: Magnetic Susceptibility Anisotropy Contributions and Consequences for Protein Solution Structure Determination. *J. Am. Chem. Soc.* **1998**, *120*, 12903−12909.

(33) Hus, J. C.; Marion, D.; Blackledge, M. De Novo Determination of Protein Structure by NMR Using Orientational and Long-Range Order Restraints. *J. Mol. Biol.* **2000**, *298*, 927−936.

3125

dx.doi.org/10.1021/ar5002318 | *Acc. Chem. Res.* 2014, 47, 3118−3126

(34) Balayssac, S.; Bertini, I.; Luchinat, C.; Parigi, G.; Piccioli, M. ^{13}C Direct Detected NMR Increases the Detectability of Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2006**, *128*, 15042−15043.

(35) Bertini, I.; Kursula, P.; Luchinat, C.; Parigi, G.; Vahokoski, J.; Willmans, M.; Yuan, J. Accurate Solution Structures of Proteins From X-Ray Data and Minimal Set of NMR Data: Calmodulin Peptide Complexes As Examples. *J. Am. Chem. Soc.* **2009**, *131*, 5134−5144.

(36) Zhang, Q.; Sun, X.; Watt, E. D.; Al Hashimi, H. M. Resolving the Motional Modes That Code for RNA Adaptation. *Science* **2006**, *311*, 653−656.

(37) Bardaro, M. F., Jr.; Varani, G. Independent Alignment of RNA for Dynamic Studies Using Residual Dipolar Couplings. *J. Biomol. NMR* **2012**, *54*, 69−80.

(38) Tjandra, N.; Garrett, D. S.; Gronenborn, A. M.; Bax, A.; Clore, G. M. Defining Long Range Order in NMR Structure Determination From the Dependence of Heteronuclear Relaxation Times on Rotational Diffusion Anisotropy. *Nat. Struct. Biol.* **1997**, *4*, 443−449.

(39) Zweckstetter, M.; Bax, A. Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.

(40) Salmon, L.; Bascom, G.; Andricioaei, I.; Al Hashimi, H. M. A General Method for Constructing Atomic-Resolution RNA Ensembles Using NMR Residual Dipolar Couplings: the Basis for Interhelical Motions Revealed. *J. Am. Chem. Soc.* **2013**, *135*, 5457−5466.

(41) Fischer, M. W.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. Domain Orientation and Dynamics in Multidomain Proteins From Residual Dipolar Couplings. *Biochemistry* **1999**, *38*, 9013−9022.

(42) Frank, A. T.; Stelzer, A. C.; Al-Hashimi, H. M.; Andricioaei, I. Constructing RNA Dynamical Ensembles by Combining MD and Motionally Decoupled NMR RDCs: New Insights into RNA Dynamics and Adaptive Ligand Recognition. *Nucleic Acids Res.* **2009**, *37*, 3670−3679.

(43) Blackledge, M. J.; Bruschweiler, R.; Griesinger, C.; Schmidt, J. M.; Xu, P.; Ernst, R. R. Conformational Backbone Dynamics of the Cyclic Decapeptide Antamanide. Application of a New Multi-conformational Search Algorithm Based on NMR Data. *Biochemistry* **1993**, *32*, 10960−10974.

(44) Clore, G. M.; Garrett, D. S. R-Factor, Free R, and Complete Cross-Validation for Dipolar Coupling Refinement of NMR Structures. *J. Am. Chem. Soc.* **1999**, *121*, 9008−9012.

(45) Yang, S.; Salmon, L.; Al Hashimi, H. M. Measuring Similarity Between Dynamic Ensembles of Biomolecules. *Nat. Methods* **2014**, *11*, 552−554.

(46) Bertini, I.; Giachetti, A.; Luchinat, C.; Parigi, G.; Petoukhov, M. V.; Pierattelli, R.; Ravera, E.; Svergun, D. I. Conformational Space of Flexible Biological Macromolecules From Average Data. *J. Am. Chem. Soc.* **2010**, *132*, 13553−13558.

(47) Bertini, I.; Ferella, L.; Luchinat, C.; Parigi, G.; Petoukhov, M. V.; Ravera, E.; Rosato, A.; Svergun, D. I. MaxOcc: A Web Portal for Maximum Occurence Analysis. *J. Biomol. NMR* **2012**, *53*, 271−280.

(48) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. Backbone Dynamics of Calmodulin Studied by [15]N Relaxation Using Inverse Detected Two-Dimensional NMR Spectroscopy; the Central Helix Is Flexible. *Biochemistry* **1992**, *31*, 5269−5278.

(49) Nagulapalli, M.; Parigi, G.; Yuan, J.; Gsponer, J.; Deraos, S.; Bamm, V. V.; Harauz, G.; Matsoukas, J.; de Planque, M.; Gerothanassis, I. P.; Babu, M. M.; Luchinat, C.; Tzakos, A. G. Recognition Pliability Is Coupled to Structural Heterogeneity: A Calmodulin-Intrinsically Disordered Binding Region Complex. *Structure* **2012**, *20*, 522−533.

(50) Bertini, I.; Luchinat, C.; Nagulapalli, M.; Parigi, G.; Ravera, E. Paramagnetic Relaxation Enhancements for the Characterization of the Conformational Heterogeneity in Two-Domain Proteins. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9149−9156.

(51) Bertini, I.; Fragai, M.; Luchinat, C.; Melikian, M.; Toccafondi, M.; Lauer, J. L.; Fields, G. B. Structural Basis for Matrix Metalloproteinase 1-Catalyzed Collagenolysis. *J. Am. Chem. Soc.* **2012**, *134*, 2100−2110.

(52) Bailor, M. H.; Sun, X.; Al Hashimi, H. M. Topology Links RNA Secondary Structure With Global Conformation, Dynamics, and Adaptation. *Science* **2010**, *327*, 202−206.

(53) Chu, V. B.; Lipfert, J.; Bai, Y.; Pande, V. S.; Doniach, S.; Herschlag, D. Do Conformational Biases of Simple Helical Junctions Influence RNA Folding Stability and Specificity? *RNA* **2009**, *15*, 2195−2205.

(54) Mustoe, A. M.; Bailor, M. H.; Teixeira, R. M.; Brooks, C. L., III; Al Hashimi, H. M. New Insights into the Fundamental Role of Topological Constraints As a Determinant of Two-Way Junction Conformation. *Nucleic Acids Res.* **2012**, *40*, 892−904.

(55) Bailor, M. H.; Mustoe, A. M.; Brooks, C. L., III; Al Hashimi, H. M. 3D Maps of RNA Interhelical Junctions. *Nat. Protoc.* **2011**, *6*, 1536−1545.

(56) Mustoe, A. M.; Al Hashimi, H. M.; Brooks, C. L., III. Coarse Grained Models Reveal Essential Contributions of Topological Constraints to the Conformational Free Energy of RNA Bulges. *J. Phys. Chem. B* **2014**, *118*, 2615−2627.

(57) Mackerell, A. D., Jr.; Banavali, N.; Foloppe, N. Development and Current Status of the CHARMM Force Field for Nucleic Acids. *Biopolymers* **2000**, *56*, 257−265.

(58) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nat. Chem. Biol.* **2011**, *7*, 553−559.

(59) Boyd, J.; Redfield, C. Characterization of N-15 Chemical Shift Anisotropy From Orientation-Dependent Changes to N-15 Chemical Shifts in Dilute Bicelle Solutions. *J. Am. Chem. Soc.* **1999**, *121*, 7441−7442.

(60) Bertini, I.; Luchinat, C.; Parigi, G. Magnetic Susceptibility in Paramagnetic NMR. *Prog. NMR Spectrosc.* **2002**, *40*, 249−273.

(61) Lohman, J. A. B.; Maclean, C. Magnetic Field Induced Alignment Effects in ^{2}H NMR Spectra. *Chem. Phys. Lett.* **1978**, *58*, 483−486.

(62) Lohman, J. A. B.; Maclean, C. Alignment Effects on High Resolution NMR Spectra Induced by the Magnetic Field. *Chem. Phys.* **1978**, *35*, 269−274.