Published in final edited form as:

J Magn Reson. 2013 December ; 237: 191–204. doi:10.1016/j.jmr.2013.08.014.

NMR studies of nucleic acid dynamics

Hashim M. Al-Hashimi*

Department of Chemistry & Biophysics, University of Michigan, 930 North University Avenue, Ann Arbor, MI 48109-1055, USA, Tel: 734-615-3361, Fax: 734-647-4865

Abstract

Nucleic acid structures have to satisfy two diametrically opposite requirements; on one hand they have to adopt well-defined 3D structures that can be specifically recognized by proteins; on the other hand, their structures must be sufficiently flexible to undergo very large conformational changes that are required during key biochemical processes, including replication, transcription, and translation. How do nucleic acids introduce flexibility into their 3D structure without losing biological specificity? Here, I describe the development and application of NMR spectroscopic techniques in my laboratory for characterizing the dynamic properties of nucleic acids that tightly integrate a broad set of NMR measurements, including residual dipolar couplings, spin relaxation, and relaxation dispersion with sample engineering and computational approaches. This approach allowed us to obtain fundamental new insights into directional flexibility in nucleic acids that enable their structures to change in a very specific functional manner.

Keywords

RDC; Spin relaxation; Relaxation dispersion; Hoogsteen; Excited states

Introduction

Biological specificity is achieved predominantly by structural complementarity between interacting biomolecules, which in turn requires the surgical positioning of functional groups and a well defined 3D structure [1]. At odds with this requirement is the fact that many biochemical processes also require, and in many cases rely on, large-scale changes in the structures of biomolecules. This is particularly the case for nucleic acids. Perhaps the most spectacular example is the unwinding of the two strands that hold the iconic DNA double helix together during replication, but equally impressive are the precise, highly coordinated, and multi-segmented conformational maneuvers executed by the ribosomes during translation.

How do nucleic acids back introduce optimum flexibility into their 3D structures to achieve functionally important conformational changes without wreaking havoc and leaving their systems exposed to promiscuous interactions with a litany of agents and cues within living cells? One approach would be to build in 'directional flexibility' that biases nucleic acid structures to specifically sample a sub-set of conformations that are deemed important for

© 2013 Published by Elsevier Inc.

*hashimi@umich.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

biological function, while strongly excluding other conformations. In the language of free energy landscapes [2, 3], the nucleic acid dynamic structure landscape would be punctuated by local minima representing those conformations required for biological function. Clearly, the specificity of the interaction between the nucleic acid and cellular co-factors will also be important, particularly in defining the extent and rate with which a given conformational change takes place in response to a given cellular cue.

Our group embarked on an effort to examine directional flexibility in nucleic acids back in 2002 [4]. Our approach was to employ solution state NMR methodology to characterize equilibrium motions in RNA and DNA at atomic resolution, and to probe whether these motions bias nucleic acids to sample specific, functionally important conformations. Nucleic acids have spectroscopic and hydrodynamic properties that are distinct from proteins, and as such, NMR methods developed to study protein dynamics could not be applied ‘as is’ to study nucleic acid dynamics. To this end, we developed NMR techniques optimized for characterizing internal motions in nucleic acids that tightly integrate a broad set of NMR measurements, including residual dipolar couplings, spin relaxation, and relaxation dispersion with sample engineering and computational approaches. This approach allowed us to visualize internal motions in nucleic acids occurring at timescales ranging between picoseconds and milliseconds and to obtain fundamental new insights into nucleic acid directional flexibility, its physical basis, and functional significance. Here, we review some of this work with an emphasis on two systems; the transactivation response elements (TAR) RNA from the human immunodeficiency virus type I (HIV-1) and a DNA oligonucleotide containing an A-tract. While this paper focuses on work from our laboratory, the reader is referred to many outstanding reviews that provide a more comprehensive account of the growing NMR studies of nucleic acid structural dynamics [5–8].

Domain-elongation

Many NMR techniques that are commonly used to study dynamics, including those that rely on measurements of spin relaxation [9, 10] and residual dipolar couplings (RDCs) [11, 12], probe the reorientation of anisotropic interaction tensors centered on nuclei of interest relative to the external applied magnetic field. Extracting information about internal motions from such measurements hinges on being able to deconvolute the smaller spectroscopic contributions due to internal motions from the much larger contributions due to overall motions. This is typically accomplished by invoking the so-called ‘decoupling approximation’ and the assumption that internal and overall motions are not correlated to one another [10]. While this approximation has been shown to hold to a high degree of validity for many globular proteins that typically undergo local excursions about a well-defined global conformation, experimental and computational studies have shown that the decoupling approximation often breaks down in nucleic acids, particularly RNA [13–17].

More than 50% of RNA secondary structure consists of A-form helices that are linked together via flexible junctions. Helices are often not pinned down in an RNA structure, but rather, are free to undergo large amplitude rigid-body collective motions. Such motions can significantly alter the RNA overall hydrodynamic shape, and therefore its overall rotational diffusion and/or molecular alignment in ordering media (Fig. 1a) [13, 14]. The resulting coupling between internal and overall motions can severely complicate analysis of spin relaxation data and residual anisotropic interactions. It can also challenge the ability to accurately predict NMR data from molecular dynamics (MD) trajectories, limiting opportunities for integrating the two approaches. In principle, bending motions in DNA could lead to similar motional couplings, though such motional modes have yet to be thoroughly investigated experimentally.

There are additional challenges in NMR studies of nucleic acids that have to do with their overall motional properties. The advent of methods to induce partial alignment of biomolecules and measurement of residual anisotropic interactions such as RDCs and residual chemical shift anisotropies (RCSAs) have revolutionized the ability to structurally characterize internal motions occurring over a broad range of timescales, spanning picoseconds to milliseconds [11, 18–22]. Extracting the rich information contained within such anisotropic interactions requires, however, the ability to modulate partial alignment of the biomolecule [23, 24]. For many proteins, this can be accomplished by changing the steric and electrostatic properties of the alignment medium used to orient samples [23, 24] or by applying site-specific mutations that alter the electrostatic properties of the solute protein without affecting its functional structure [25]. However, approaches for modulating alignment by changing the ordering medium have so far failed for nucleic acids [26, 27], likely because their uniform charge distribution closely follows that of their overall shape, resulting in similar electrostatic and steric alignment forces. Although spontaneous magnetic field alignment [28] often leads to a unique alignment [18, 27, 29], the degree of alignment is typically one to two orders of magnitudes smaller than optimum levels, resulting in RDCs that have magnitudes comparable to measurement uncertainty.

We have developed a domain-elongation strategy [13] to help address some of these limitations (Fig. 1b). Here, a terminal helix is elongated using a stretch of isotopically unlabeled Watson-Crick (WC) base-pairs designed to adopt an A-form (or B-form for DNA) helix (Fig. 1c). The resulting elongated nucleic acid has an overall shape, and consequently overall motional properties, which are far less sensitive to internal motions occurring in other parts of the molecule. By slowing down the overall rate of molecular tumbling, the elongation also serves to broaden the timescale sensitivity of spin relaxation data, which is limited to motions occurring at timescales faster than overall molecular tumbling. By elongating different helices in a nucleic acid, and introducing variable kinks within helices [30], it becomes feasible to modulate the partial alignment of the nucleic acid, and to expand the number of RDCs and other residual anisotropic interactions that can be measured and used in characterizing dynamics. These factors also make it possible to compute spin relaxation data and RDCs from MD simulations providing the basis for integrating the two approaches.

Thus far, comparison NMR spectra of elongated nucleic acids with their non-elongated counterparts (Fig. 1d) under free and protein/ligand/metal bound conditions for a variety of RNA and DNA systems suggest that elongation does not affect the structural and functional integrity of the nucleic acid and that any concomitant effects are either insignificant or primarily localized at the point of elongation [13, 31–34]. However, there will likely be cases in which elongation is not feasible without disrupting key aspects of a nucleic acid structure, and these effects should always be studied carefully. To this end, we have developed methods reviewed below that use much smaller degrees of elongation and structural perturbations to modulate alignment. It should also be noted that other interesting methods for modulating RNA alignment have been developed by Varani and co-workers that rely on the installation of a protein-binding site within RNA targets [35].

Picosecond-to-nanosecond motions in RNA using domain-elongation and $^{15}\text{N}/^{13}\text{C}$ spin relaxation

We first used domain-elongation to study the dynamic properties of the transactivation response element (TAR) RNA [36], which is a stem-loop located at the 5'-end of the HIV-1 genome that plays essential roles in various steps of the viral replication cycle (Fig. 2a). Prior NMR studies by Williamson, Puglisi, Varani, and others showed that TAR undergoes large structural rearrangements on binding to peptides, ligands [37–39], Mg^{2+} ions [40] and

small molecule inhibitors (Fig. 2a) [41–44]. In these complexes, TAR assumes inter-helical bend angles that vary between $\sim 5^\circ$ and $\sim 47^\circ$ and there are also marked alterations to its local structure in the bulge and apical loop, which form two distinct binding sites (Fig. 2a). Such adaptive conformational changes upon recognition were commonly observed in RNA [45], but it remained unclear whether ligands shaped the ligand-bound RNA conformation via ‘induced-fit’, or whether intrinsic directional flexibility predisposed RNAs to sample a subset of conformations, which are in turn captured by ligands via ‘conformational selection’ [46–48]. One of our early goals was to structurally characterize internal motions in TAR, and examine whether free TAR preferentially samples its ligand bound conformations.

Surprisingly, despite the near 3-fold increase in molecular weight accompanying domain-elongation, high-resolution 2D HSQC spectra could be recorded for elongated TAR (E-TAR) with excellent signal-to-noise-ratios (Fig. 2b) [13]. Moreover, the spectra of E-TAR were in excellent agreement with its non-elongated counterpart, both in free form and when bound to argininamide (ARG); a ligand mimic of TAR’s cognate protein target Tat (Fig. 1d). The observation of degenerate imino ^1H signals characteristic of Watson-Crick base-pairs also confirmed that the elongation residues adopted the expected helical structure [13].

We originally developed domain-elongation to study RNA dynamics using RDCs. However, the 2D HSQC spectra of elongated TAR told an unexpected story that led us to pursue spin relaxation studies first. There were marked variations in resonance intensities in spectra of E-TAR that were not observed in non-elongated TAR (Fig. 2b). These variations correlated nearly perfectly with the TAR secondary structure, with bulge residues exhibiting the most intense resonances, upper helix exhibiting medium intensity resonances, and those in the lower elongated helix exhibiting the lowest intensities (Fig. 2b). One explanation was that different degrees of internal motions were giving rise to different relaxation rates, with the elongated helix undergoing slow overall tumbling, and the bulge undergoing extensive local motions that lead to smaller amplitude collective motions of the upper helix. These internal motional contributions were diminished in non-elongated TAR likely because they occurred at timescales slower than overall tumbling of the small RNA. Prior RDC studies of non-elongated TAR had already provided evidence for collective inter-helical motions about a flexible bulge [49].

By measuring imino ^{15}N relaxation [50] data for guanine and uridine residues, we confirmed the existence of collective inter-helical motions in TAR occurring at nanosecond timescales. Uniformly smaller ratios of transverse (R_2) to longitudinal (R_1) relaxation rates (R_2/R_1) were observed for sites in the smaller helix as compared to the elongated helix (Fig. 2c). These differences could not be attributed to anisotropic diffusion in these highly elongated and anisotropic RNAs, because in an A-form helix, the imino ^{15}N CSA and N-H dipolar interaction tensors are oriented nearly perpendicular to the long axis of the molecule, and therefore the principal axis of the nearly axially symmetric diffusion tensor. Consequently the R_2/R_1 values measured in the elongated helix already correspond to a *minimum* when taking into account anisotropic overall diffusion; the observation of even lower R_2/R_1 values in the smaller helix could only be attributed to internal motions and semi-independent reorientation of the two helices at nanosecond timescales. Binding of ARG, which is known to stabilize the TAR bulge [51], led to a uniform increase in the short helix R_2/R_1 values such that they became similar to those of the elongated helix, consistent with an arrest of the inter-helical motions (Fig. 2c). In stark contrast to elongated TAR, indistinguishable R_2/R_1 values were observed for the two hydrodynamically equivalent helices in non-elongated TAR and these relative values did not change significantly on binding to ARG (Fig. 2c). Extended model free analysis [10, 52] of the ^{15}N relaxation data measured in E-TAR revealed amplitudes of inter-helical motions corresponding to spin relaxation order parameters of $S_s^2 = 0.68 - 0.83$ and time constants ($\tau_s = 1.4 - 1.9$ ns) approaching the

hydrodynamically predicted time constants for overall rotational diffusion of the smaller helix.

Interestingly, we observed a correlation between the dynamics at a given site, as measured qualitatively based on resonance intensities in 2D HSQC spectra, and the amplitude of structural rearrangement at given site due to ligand binding (Fig. 2d). This provided the first evidence for a correlation between intrinsic motions and directional flexibility in RNA and the conformational changes that take place on ligand binding. We subsequently developed TROSY-detected pulse sequences for measuring ^{13}C R_1 , $R_{1\rho}$, and heteronuclear $^{13}\text{C}\{^1\text{H}\}$ NOEs for protonated base (C2, C5, C6, and C8) and sugar (C1') carbons in large elongated RNAs and developed data analysis schemes to obtain a quantitative description of C–H motional amplitudes [53]. We developed a model free formalism which takes into account the very high anisotropy of overall rotational diffusion, asymmetry of the nucleobase CSAs and noncollinearity of C–C, C–H dipolar and CSA interactions, under the assumption that all interaction tensors for a given carbon experience identical isotropic internal motions [53]. A strong correlation was again observed between the order parameters describing the amplitudes of internal motions in TAR and amplitude of structural rearrangement at given site due to ligand binding. The ^{13}C relaxation data also revealed that while ARG binding arrests inter-helical motions, it dramatically increases the local mobility of the flipped out bulge residues linking the two helices, which undergo virtually unrestricted internal motions ($S^2 \sim 0.2$) in the ARG bound state [53].

Domain-elongation led us to uncover inter-helical and local motions occurring at nanosecond timescale in a variety of RNA systems, including the HIV-1 stem loop 1 [32], which plays essential roles in the dimerization and encapsulation of the viral genome and the P4 stem [31], which is located near the catalytic center of RNase P. From these and other studies, important trends started to emerge [5–8]. The inter-helical motional amplitudes tend to increase with the length of linker residues, and often are not fully captured by spin relaxation, because they occur at timescales slower than overall tumbling. Base-paired residues tend to experience limited motions whereas non-canonical residues undergo larger more variable amplitude motions that tend to be inversely correlated with extent of stacking. Importantly, we often observed extensive mobility at highly conserved residues that are known to undergo functionally conformational rearrangements during recognition and catalysis. In later studies, we also adapted the domain-elongation strategy to study DNA dynamics and this led us to uncover nanosecond end-fraying motions that penetrate deep into the helix interior [54].

Magnetic field induced RDCs with motional couplings

Although the spin relaxation data provided strong evidence for collective inter-helical motions in HIV-1 TAR, it remained unclear whether these motions direct TAR toward the ligand bound states. We therefore pursued measurements of RDCs, which can in principle provide a ‘structural’ description of the motions, and which enjoy much broader timescale sensitivity extending from picoseconds up to milliseconds [19, 20, 55]. Our initial efforts, prior to the development of domain-elongation, had focused on measurements of RDCs in magnetically aligned RNA [14]. Although the degree of alignment (10^{-5}) and magnitude of RDCs measured in magnetically aligned RNA is nearly two orders of magnitude smaller than optimum (10^{-3}), magnetic field alignment has a simple dependence on nucleic acid structure [14, 56, 57], which could be exploited to treat the effects of motional couplings. For a magnetically aligned nucleic acid, the overall alignment tensor is given by the overall diamagnetic susceptibility (χ) tensor, which can be computed for a given RNA structure based on a tensor summation over all χ -tensors associated with individual nucleobases [14, 58]. This makes it possible to express the overall alignment tensor in terms of the nucleic

structure, and specifically the orientation of nucleobases, and therefore, to explicitly take into account correlations between internal motions and overall magnetic alignment. This afforded us an opportunity to obtain rare insights into the effects of motional couplings on NMR observables [14, 59], which serendipitously, led us to formulate a more general conceptual framework that describes the information that can be accessed about domain motions from RDCs [60, 61]. This in turn guided our subsequent sample engineering efforts [62] for accessing this information from the more optimal RDCs measured in ordering media.

To illustrate how motional couplings influence measured magnetic field induced RDCs, let us consider an RNA composed of two locally rigid helices, HI and HII, each assumed to have axially symmetric χ -tensors (χ_I and χ_{II}) with principal axes oriented along each helical axes (Fig. 3). The overall alignment tensor experienced by HII is given by the overall χ -tensor, $\chi_I^2(II)$, which can be expressed as a tensor sum over the χ -tensors of HII and HI. In the principal axis system (PAS) of the HII χ -tensor, an expression for the overall χ -tensor experience by HII is given by (Fig. 3) [59],

$$\chi_I^2(II) = X_I^2(II) + X_0^2(I) D_{0I}^2(\alpha\beta) \quad (1)$$

The overall χ -tensor experienced by HII depends on the orientation of HI relative to HII, as specified by Wigner rotation $D_{kl}^2(\alpha\beta)$, and specifically on the bend angle (β) between helices as well as the twist angle about its own (HII) helical axis; however, it is insensitive to the angle γ , which specifies a rotation about the axially symmetric axis of HI (Fig. 3). For a dynamic ensemble consisting of many inter-helical orientations, the time-averaged overall χ -tensor is given by [59],

$$\langle \chi_I^2(II) \rangle = X_I^2(II) + X_0^2(I) \langle D_{0I}^2(\alpha\beta) \rangle$$

where the angular brackets denote a time average over all inter-helical orientations sampled at rates faster than the inverse of the dipolar interaction. Assuming the χ -tensors of individual helices is known, RDCs measured in HII can be used to determine the five elements of the overall χ -tensor, which can in turn be used to determine the five time-

averaged Wigner rotation elements, $\langle D_{0I}^2(\alpha\beta) \rangle$, which carry information about the inter-helical motions. Indeed, the $\langle D_{0I}^2(\alpha\beta) \rangle$ elements are identical in form to the widely used order tensor which provides a model-free description of the ordering of chiral molecules relative to the axially symmetric external magnetic field [63]. However, unlike the order tensor, $\langle D_{0I}^2(\alpha\beta) \rangle$ provides a model-free description [20] of the chiral helix HII ordering relative to the internal axially symmetric axis of HI. Thus, in Equation 1, an NMR frame of reference has effectively been anchored onto HI (Fig. 3).

What information can be obtained from the RDCs measured in the other HI helix? An expression for the overall χ -tensor experienced by HI in the XIPAS is given by [59],

$$\langle \chi_I^2(I) \rangle = X_I^2(I) + X_0^2(II) \langle D_{0I}^2(-\gamma - \beta) \rangle \quad (2)$$

Interestingly, the overall χ -tensor experienced by HI carries unique information about the inter-helical motions, as defined by the Wigner rotation elements $\langle D_{0I}^2(-\gamma - \beta) \rangle$. As for HI, the overall χ -tensor of HII is sensitive to β , but unlike HI, it is sensitive to the angle γ and

not α (Fig. 3). Thus, in Equation 2, the NMR frame of reference is anchored onto HII (Fig. 3). Together, the nine elements of $\langle D_{0l}^2(\alpha\beta) \rangle$ and $\langle D_{0l}^2(-\gamma - \beta) \rangle$ provide nine independent parameters that can be used to characterize inter-helix motions with complete 3D rotational sensitivity to α , β , and γ [59].

The above expressions show that the added complexity inherent to motional couplings carries additional information for dynamics characterization. Specifically, it makes it possible to anchor NMR frames of reference onto different domains in a molecule. The reference domain can then be used to probe relative motions of all other domains that impact its overall alignment. In this manner, one can observe the same inter-domain motions from different domain-centered perspectives, and thereby overcome inherent spatial limitations in NMR interactions, such as the inability to measure rotational motions about axially symmetric interactions [59]. It is worth noting that such domain-anchoring can be accomplished in proteins through binding of paramagnetic metals or attachment of paramagnetic tags to specific domains [64, 65]. Finally, the above example assumed axially symmetric χ -tensors for the individual helices. Analogous derivations for the more general cases that allows asymmetric tensors show that it is theoretically possible to access all

twenty-five Wigner elements, $\langle D_{nl}^2(\alpha\beta\gamma) \rangle$ by modulating overall alignment [61]. These 25 Wigner elements define the theoretical spatial resolution limit with which motions of chiral domains can be defined using residual anisotropic NMR interactions.

Inter-helical motions in RNA from domain-elongation and RDCs

Despite the utility of motional couplings in magnetically aligned RNA, the magnitude of field-induced RDCs remained unfavorably small. We therefore sought alternative approaches for anchoring axially symmetric frames onto helices, and obtaining the information contained within the nine elements of $\langle D_{0l}^2(\alpha\beta) \rangle$ and $\langle D_{0l}^2(\beta\gamma) \rangle$, or even better, all 25 $\langle D_{nl}^2(\alpha\beta\gamma) \rangle$ Wigner elements, under the more favorable alignment conditions obtained by dissolution of nucleic acids in ordering media. Serendipitously, one way to achieve this was to simply elongate a target domain. Indeed, when this idea came to mind, we had already measured RDCs in TAR samples in which the lower helix was elongated. What remained to be done was to prepare a second set of samples in which the upper helix is elongated – an exercise that prior to our theoretical treatment of motional couplings would have appeared to be utterly redundant (Fig. 4a). In this manner, independent elongation of HI and HII made it possible to anchor an axially symmetric alignment tensor on HI and HII respectively, and thereby to measure the nine elements of $\langle D_{0l}^2(\alpha\beta) \rangle$ and $\langle D_{0l}^2(\beta\lambda) \rangle$ (Fig. 4a) [62]. The RDCs measured in the two RNA samples under the decoupling limit could be then used to access the same information obtained in one sample under the motional coupling limit, with the added advantage of measuring the much larger RDCs measured in ordering media such as filamentous bacteriophage [66, 67].

For both elongated TAR samples, the RDC derived order parameter ($S_{\text{RDC}} \sim 0.46$) [62] indicated amplitudes for inter-helical motions that exceeded those measured by spin relaxation ($S_{\text{S}} \sim 0.86$) [13]. Given the broader timescale sensitivity of RDCs compared to spin relaxation and lack of evidence for any exchanging broadening at the TAR bulge and neighboring residues, the ‘excess’ motions detected by RDCs likely occurred at ns- μ s timescales. To visualize the inter-helical motions, we performed a grid search over the three Euler angles describing the orientation of the two helices to search for ensembles consisting of up to three ($N = 3$) equally populated inter-helical conformations that can reproduce the measured RDCs [62]. Prior studies by Clore and co-workers had demonstrated the feasibility

of using RDCs in determining ensembles of proteins [68] and DNA [69]. These discrete TAR ensembles were expected to capture key features of the inter-helical motions even though the actual motions may involve many more conformations. Loose steric constraints were implemented in the search to avoid collisions between helices and orientations that cannot be satisfactorily linked by the trinucleotide bulge. Both the $N=1$ and 2 ensemble search yielded a very poor RDC fit, implying an inter-helix motional trajectory that is more complex than a simple two-state jump (Fig. 4b). In contrast, a very good fit was obtained for $N=3$ with insignificant improvements obtained with $N=4$ (Fig. 4b).

A striking feature of the RDC-derived 3-state ensemble was that the three conformations fell nearly along a straight line in the 3D inter-helix Euler space defining twisting around each helix (α and $-\gamma$) and inter-helical bending (β) (Fig. 4c). This hinted to a motional trajectory in which HI and HII bend and twist in a correlated manner (Fig. 4d). Thus, although the helices undergo large amplitude collective motions ($> 90^\circ$), they move in a very specific and directional manner. This was a clear sign of ‘directional flexibility’ in RNA. What’s more, the 3-state ensemble enveloped many of the known ligand-bound TAR conformations, indicating that directional flexibility specifies the TAR conformations that are recognized by ligands (Fig. 4c).

Topological constraints as determinants of RNA inter-helical directional flexibility

What are the forces that underlie TAR’s directional inter-helical flexibility? In subsequent studies, we determined an all-atom ensemble of TAR by combining RDC measurements with MD simulations (see below) [70]. Using this independent approach, we found once again that the two helices sample specific regions of the Euler space in which the three Euler angles co-vary in a linear manner, as determined previously for the 3-state TAR ensemble. It was quite striking that despite the fact that the helices are free to undergo very large amplitude motions, changing inter-helical angles by more than 90° , they sampled less than 5% of the total Euler space. In other words, the helices sample wildly different orientations, yet they do specifically, and exclude the vast majority of possible orientations.

We examined whether this directional inter-helical flexibility is unique to TAR or a more general feature of RNA helices connected by bulges. To this end, we calculated the inter-helical Euler angles for every RNA two-way junction in the protein databank (PDB) [71]. This included very diverse RNA structures in complex with proteins, ligands, and metals as well as unbound RNAs. To our surprise, we found that all of trinucleotide bulges – and indeed all RNA sequences containing two-way junctions – confine the orientation of the helices to similar regions of the Euler space as observed for TAR (Fig. 5a). Thus, through the determination of a single NMR dynamics ensemble, we were able to quite literally reproduce the entire X-ray structure database of RNA two-way junctions!

The strong conservation of directional flexibility across different RNA sequence and contexts called for a simple explanation. By sketching a diagram of an RNA two-way junction geometry, and noting that A-form helices with diameter of $\sim 25 \text{ \AA}$ are connected at two positions by bulges with comparable length of $\sim 4.9 \text{ \AA}$ per nucleotide, it became apparent that the system was constrained by two simple steric and connectivity forces, which we collectively refer to as ‘topological constraints’ (i) helices cannot sterically clash with one another and (ii) the distance between points of inter-helical connectivity cannot exceed the maximum length of the bulge linker. Based on these two simple constraints, we were able to accurately predict the allowed orientations of all two-junctions in the PDB and show that they are confined on average to $<8\%$ of the allowed inter-helical Euler angles (Fig. 5a) [71]. In other words, the inter-helical structure and flexibility of RNA is already

predetermined, to a large extent, by its secondary structure and junction topology. These topological constraints are increasingly being shown to be fundamental determinants of RNA folding and conformational adaptation [72–74].

Returning to the TAR inter-helical trajectory, one can understand how topological constraints direct the inter-helical motions along specific directions and why this might not be the case, for example, for two α -helices in a protein. In the RNA case, the two A-form helices are connected at two positions around the edge of the helix and pivoted at one position. Because the diameter of the A-form helix is large compared to the inter-spacing between residues in the bulge, rotation of one helix about its axis quickly exerts a drag on the bulge linker, which can be alleviated by either rotating the other helix about its axis, or bending the two helices with respect to one another, in so doing, generating a distribution of allowed conformations in which the inter-helical Euler angles co-vary (Fig. 5b,c). In the protein case, the two α -helices are linked at the centre of the helix via a single linker (Fig. 5b,c). As a result, there is no bias in the restriction along the three Euler angles, resulting in a more isotropic distribution of orientations (Fig. 5b). Thus, by changing the topology of junctions connecting domains, and the domain shape, it is possible to graft complex 3D inter-domain ensemble distributions without the need for complex interactions. It is very likely that similar topological constraints play important roles in defining the orientation of domains in multi-domain proteins.

Constructing RNA ensembles

The visualization of the TAR inter-helical ensemble was made possible in large part because many RDCs measured in the two helices could be pooled together to characterize only three inter-helical Euler angles. Determining all-atom ensembles of nucleic acids is significantly more challenging because a much larger number of degrees of freedom have to be characterized using comparatively smaller amount of data on a per fragment basis. One strategy to address this problem was not to measure more NMR data, because it is unlikely that enough interactions could ever be measured to uniquely specify a dynamic ensemble, but rather, to rely on computational approaches such as MD simulations to bridge the gap.

Early on, in collaboration with the group of Ioan Andricioaei at UC-Irvine and more recently Charles Brooks at the University of Michigan, we began to use MD simulations to generate atomic resolution models for RNA internal motions and dynamic ensembles. The basic idea was to take MD simulations and see how well they back predicted NMR data including spin relaxation but perhaps more importantly RDCs. Domain-elongation was critical for these studies because it made it possible to compute NMR data from the MD trajectory. In particular, the elongated helix could be used as a reference to superimpose MD snap-shots, and the overall alignment (or diffusion) tensor determined for the elongated helix via order tensor analysis of RDCs could be used to compute RDCs for very MD snap-shot, and time-averaged RDCs by averaging over all snap shots.

We first examined a $50 \times 1.6 = 80$ ns MD trajectory of HIV-1 TAR [70] computed using the CHARMM package with force field parameter set 27 [75]. At the outset, it should be noted that the simulation was nowhere near long enough to match the RDC timescale sensitivity of milliseconds, and any failure to predict the measured RDCs could not be considered to be evidence for a poor force field, as predictions could always improve with longer simulation times. The need for much longer simulations continues to be a major problem in the NMR assessment of nucleic acid force fields. Although some correlation was observed between the measured and predicted RDCs for both EI-TAR and EII-TAR, the deviations (RMSD = 13–16 Hz) were substantially larger than the estimated RDC measurement uncertainty (~4 Hz). Interestingly, the MD trajectory did not systematically under- or overestimate the

RDCs, implying that it does not significantly over- or underestimate the amplitude of motions present.

Next, we examined if there might be a way to edit the MD trajectory to improve its agreement with the RDC data. Assuming that the MD simulation adequately samples the entire range conformations that are adopted by TAR in solution (but fails to predict their populations), one can use the NMR data to effectively re-weight the MD pool. To accomplish this, we implemented the “sample and select” (SAS) strategy developed by Dokholyan and his colleagues for constructing protein ensembles using NMR order parameters [76]. Here, N conformers are randomly selected from the MD-generated pool to generate trial sub-ensemble and the agreement between measured and predicted RDCs computed (Fig. 6a). Next, one of the chosen conformers is replaced randomly with another conformer from the pool, and the agreement with measured RDCs re-examined. The newly selected conformer is either accepted or rejected based on the metropolis criteria and several iterations are carried out until convergence is reached, defined as achieving agreement with the measured RDCs that exceeds the experimental error (Fig. 6a) [70, 77]. The ensemble size (N) is then incrementally increased in steps of 1 from $N=1$ until the convergence criteria is met.

With this SAS approach, we were able to construct ensembles comprising 20 conformations ($N=20$) that satisfy the measured RDCs (RMSD = 4.8 Hz) to near within experimental error (<4 Hz) [70, 77]. 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 inter-helical bend angle as observed with the 3-state ensemble of TAR but it also allowed the visualization of local motions in and around the bulge. 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 [70].

The above approach for constructing TAR ensembles suffered from some fundamental limitations. The need to elongate RNA samples and decouple internal and overall motions limited the number of independent RDCs that could be used in the ensemble construction to typically two data sets for every pair of helices. This made it impossible to acquire the full information contained within the 25 Wigner matrix elements. Moreover, many RNA systems are not amenable to elongation of more than one helix, further reducing the amount of data that can be measured. In prior studies, we had shown that partial helix elongation by as little as three base-pairs, or elongation with kinks to induce bends, allows for the modulation of alignment and the measurement of multiple independent sets of RDCs [30]. To accommodate data from such partially elongated RNAs, it was necessary to find a way to treat correlations between internal motions and overall alignment.

We therefore developed an alternative strategy that bypasses the requirement for extensive elongation [78]. Here, RDCs are predicted for a given RNA conformation based on its overall shape using the program PALES [79]. This obviates the need to elongate the RNA and makes it possible to combine RDCs measured in multiple RNA constructs containing variable and much smaller degrees of elongation. Using a total of four RDC data sets measured in four differentially elongated TAR constructs (Fig. 6b), and a pool of conformation derived from a much longer 8.2 μ s MD trajectory computed on the Anton supercomputer [80] using the updated CHARMM36 force field [75, 81, 82], we constructed an all atom ensemble of TAR.

As shown in Fig. 6c, the agreement between measured and predicted RDCs greatly improved with increasing ensemble size and reached a plateau at $N \sim 20$ where the RMSD of 4 Hz comes very close to the experimental measurement uncertainty. This can be compared to an RMSD of ~ 9 Hz between the measured and predicted RDCs for the entire MD simulation (Fig. 6d). Clearly, this new MD simulation does a better job predicting the RDCs; however it remains unclear whether this is due to the longer timescales and/or use of an updated force field. The ability to incorporate a large number of RDCs made it possible to more rigorously assess the quality of the RDC-derived ensemble using cross-validation, in which parts of the RDCs data are left out from ensemble construction and used to evaluate ensemble quality. The quality of the ensemble was also examined based on its ability to reproduce ^1H experimental chemical shifts measured at the bulge using the program NUCHEMICS [83, 84] to compute ^1H chemical shifts based on structure. The RDC ensemble reproduces the bulge ^1H chemical shifts with RMSD of 0.17 ppm as compared to 0.19 ppm for the entire MD trajectory.

The high-resolution RDC-derived TAR ensemble allowed us develop a comprehensive picture of motions in TAR which includes demarcation of ns- μs motions sensed by RDCs and not spin relaxation. The ensemble revealed a sub-set of bent conformations that yield motional amplitudes consistent with those probed by spin relaxation, and another set of more coaxial conformations. The ensemble revealed that transitions between bent and coaxial conformations are likely to be slow, because they require disruption of stacking and hydrogen binding interactions in and around the bulge. It is possible that these transitions occur at timescales too slow to be accessible by spin relaxation measurements. Importantly, this higher resolution ensemble came even closer to within sampling the TAR conformations observed when bound to various ligands, confirming that intrinsic motions specify the ligand bound states (Fig. 6d).

The above RDC studies as well as other studies comparing MD simulations employing the CHARMM force field [75] with spin relaxation data measured in elongated TAR [85] and a single-stranded RNA [86] show that while MD captures important trends and the distinct modes of motions observed using NMR, the agreement between the measured and predicted NMR data is far from being quantitative. This could be attributed to lack of convergence of the simulations, a mismatch in motional timescales due to inadequate modelling of solvent friction [85, 87], improper modelling of ions [88] which are known to profoundly influence motions in TAR [89], and also to incomplete decoupling of internal and overall motions or inaccuracies in treating any existing correlations.

Transient Hoogsteen base-pairs in canonical duplex DNA

During the course of our TAR studies, we became interested in using relaxation dispersion to assess the timescales of the inter-helical motions observed using RDCs. As it turned out, we did not detect any evidence for μs -ms exchange broadening at the bulge linker and neighboring residues, indicating that the ‘excess’ inter-helical motions sensed by RDCs and not spin relaxation occur at the ns- μs timescale window. RDC studies by Griesinger and colleagues have elegantly established the occurrence of such ‘supra- τ ’ ns- μs motions in proteins [90, 91].

This detour into relaxation dispersion made us interested in more generally exploring chemical exchange in nucleic acid, and to characterize slower motions occurring at μs -ms timescales that are directed towards higher energy conformational sub-states, which often are referred to as ‘excited states’ [92–94]. $R_{1\rho}$ and Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion techniques developed by Kay and Palmer [92–94] have made it possible to characterize excited states in proteins, determine their 3D structure [95, 96], and

thereby establish their importance in catalysis [97], signaling [98], recognition [99], folding [100], and aggregation [101]. However, there were a number of challenges that made it difficult to apply relaxation dispersion techniques to study conformational exchange in nucleic acids, particularly slower millisecond timescale exchange. The shortage of ideal ^{15}N reporter spins makes it necessary to rely on ^{13}C nuclei; however, CPMG experiments are difficult to apply to carbon nuclei in uniformly labeled nucleic acid samples due to extensive ^{13}C – ^{13}C interactions that are difficult to suppress [102, 103]. While appropriate radiofrequency (RF) spin lock fields can be used in the $R_{1\rho}$ experiment to reduce or eliminate unwanted ^{13}C – ^{13}C interactions, the need to suppress C–H scalar coupling evolution and cross-correlated relaxation between C–H dipole-dipole and carbon CSA typically limited the RF spin lock field strength to ~ 1 kHz, setting an upper limit for the exchange timescale that can be characterized to ~ 500 microseconds. This made it difficult to characterize slower millisecond timescale conformational exchange.

To address this problem, we exploited advances in $R_{1\rho}$ experiments introduced by the groups of Palmer and Kay that allow use of much lower spin lock fields [104, 105], on the order of 200 Hz, thus extending sensitivity to exchange timescales on the order of tens of milliseconds [106]. The experiment (Fig. 7a) uses selective Hartmann-Hahn polarization transfers developed by Bodenhausen [107, 108] to excite specific spins of interest and collect data in a 1D manner. This scheme is particularly well-suited for nucleic acids, where exchange is often limited to a small number of residues, making it unnecessary to record full multidimensional experiments. The resultant ~ 100 -fold timesaving makes it possible to comprehensively map out the dependence of $R_{1\rho}$ on spin-lock amplitude (ω_1) and offset (Ω). It also circumvents the need for $^{13}\text{C}/^{15}\text{N}$ isotopic enrichment when working with concentrated nucleic acid samples (>2 mM), allowing studies of chemically modified nucleobases that are difficult to prepare with $^{13}\text{C}/^{15}\text{N}$ isotopic enrichment [106]. Importantly, C–H scalar coupling evolution and cross-correlated relaxation between C–H dipole-dipole and carbon CSA can be efficiently suppressed using a strong ^1H continuous-wave field applied on the resonance of interest (Fig. 7a), allowing use of significantly weaker RF fields in the range of 200–1000 Hz [104, 105].

To our surprise, while we did not observe exchange in the TAR bulge, we did observe clear evidence for chemical exchange at base and sugar carbons in canonical duplex DNA (Fig. 7b). This chemical exchange would have been difficult to detect by conventional methods because the exchange contribution would have already been suppressed by spin lock powers of 1 kHz (Fig. 7b). A two-state analysis of the off-resonance relaxation dispersion data using equations developed by Palmer [109] revealed a single base-pair exchange process that is directed towards minutely populated ($p_B \sim 0.5\%$) excited states that have exceptionally short lifetimes ($\tau_{\text{ex}} = 1/k_{\text{ex}} \sim 1.5$ ms and ~ 0.3 ms for G·C and A·T), and that have ~ 3 ppm downfield-shifted sugar and base carbon chemical shifts. Analysis of the temperature dependence of the relaxation dispersion data yielded forward rate constants that are comparable to those previously measured for base-pair opening, indicating that the transition to the excited state requires the complete disruption of Watson-Crick base-pairs.

Based on a survey of carbon chemical shifts [110] and density functional theory (DFT) calculations [111], the correlated exchange at purine C8 and C1' and large downfield carbon chemical shifts (Fig. 7c) strongly hinted to an *anti-to-syn* transition, where the purine base flips 180° like a pancake. In DNA, such a flip is accomplished by a 180° rotation of the purine base around the glycosidic bond (N9–C1'). Strikingly, the *syn* purine could base-pair with its pyrimidine partner to form a different type of base-pair referred to as a Hoogsteen (HG) base-pair (Fig. 7d) [112, 113] [113]. As in WC base-pairs, the A–T HG base-pair features two hydrogen bonds between thymine and the *syn* adenine base, one of which (thymine O4 and adenine N6) is identical to that seen in WC base-pairs (Fig. 7d).

However, the second hydrogen bond is not between thymine N3 and adenine N1, but rather, between thymine N3 and N7 of the flipped adenine base (Fig. 7d). Likewise, the HG G-C⁺ base-pair preserves hydrogen bonding between cytosine N4 and guanine O6. However, formation of a second hydrogen bond with N7 of the flipped guanine requires protonation of cytosine N3. The latter explained the strong pH-dependence of the dispersion observed in G-C but not A-T base-pairs (Fig. 7d) [114].

While HG base-pairs explained the dispersion data, we need additional evidence to pin down the excited state. Once again, we resorted to sample engineering. Specifically, we introduced chemical modification designed to trap the excited state HG base-pair. Kay and co-workers had previously shown the feasibility of trapping a protein excited state by deleting a segment of the protein [95]. In our study, we methylated N1 positions in guanine and adenine bases, thereby disrupting the ability to form WC while preserving the ability to form HG base-pairs. This not only allowed us to examine whether HG base-pairs could stably form within duplex DNA, it also provided a route for directly measuring the HG base-pair carbon chemical shift signatures, which could in turn be compared with the carbon chemical shifts determined by relaxation dispersion. We were able to confirm formation of G-C⁺ and A-T HG in these N1-methylated samples using NOE measurements. Moreover, the carbon chemical shifts of the trapped HG base-pairs were in excellent agreement with those measured for the excited state using relaxation dispersion (Fig. 7c). In a follow-up study, we used an inverse strategy involving a single atom substitution in the purine base designed to destabilize the HG base-pair, and as expected, this abolished the observed relaxation dispersion [115]. We also extended this methodology to measure of ¹⁵N R_{1ρ} data which was also consistent with a HG excited state [115].

Since their discovery in 1959, Hoogsteen base-pairs continued to surface in DNA duplexes containing alternating AT repeats [116], and in duplex DNA bound to antibiotics [117] and a wide variety of proteins [118, 119]. This stirred a good deal of controversy regarding the true nature of the base-pairing in duplex DNA [113]. More recently, iota polymerases have been shown to employ HG base-pairing as a general mechanism to replicate both damaged and undamaged DNA [120]. Here, HG base pairing makes it possible to replicate nucleotides that are damages in the Watson-Crick face. In many of the DNA-protein complexes, there are specific interactions that stabilize the HG base-pair. However, our results clearly show that the formation of HG base-pairs in naked duplex DNA is already energetically quite favorable; the free energy difference between WC and HG base-pairs is ~ 3 kcal/mol. Rather, the picture that emerges is one, in which every base-pair in DNA exists as rapid superposition of WC and HG base-pairs, with external parameters operating on the DNA resolving one or the other base-pair. This directional flexibility can explain the long and controversial observation of WC versus HG [113] – small changes in conditions can favor one form over the other. It is striking that the difference in the abundance of transient G-C⁺ and A-T base-pairs mirrors the differences in efficiency observed in iota polymerase replication of A/T versus G/C. The HG base-pairs transiently expose the Watson-Crick faces of purines, and may potentially help explain the much greater abundance of damage in the form of N1 methylation in adenine versus guanine.

Finally, we note that our NMR relaxation dispersion studies suggest that ~0.1%–10% of genomes exist as HG base-pairs; for the human genome, which contains ~3 billion base-pairs, this amounts to many millions of HG base-pairs. Because the diameter of the HG helix is compressed relative to that of WC, it may well be that a larger fraction of the supercoiled and compacted eukaryotic genomes exist as HG rather than WC base-pairs. When combined with the current difficulties in resolving WC from HG based on X-ray diffraction data, it may well be the case that there are more HG base-pairs in X-ray structures currently deposited in the PDB that have gone undetected, particularly for A-T base-pairs.

Transient structures of RNA

Unlike DNA, evidence for chemical exchange and transitions towards excited states had long been observed in RNA. An early example are $R_{1\rho}$ studies by Pardi and Hoogstraten which revealed microsecond timescale exchange processes involving an excited state CA^+ base-pair within the active site of the lead-dependent ribozyme that may be important for catalysis [121]. In some cases, the excited state could be trapped by changing pH, as elegantly demonstrated by Butcher and co-workers in their characterization of an excited state structure involving a CA^+ base-pair within a functionally important region of the U6 RNA element [122, 123]. However, in most cases, the structures of RNA excited states were unknown, and difficult to determine, especially given the much larger range of conformations accessible to RNA.

Using the low spin lock field carbon $R_{1\rho}$ relaxation dispersion experiment, we started to observe chemical exchange in literally every RNA studied in the lab [124]. Unlike for DNA, in which the relaxation dispersion was localized at the base-pair level, the relaxation dispersion measured in RNA spanned many residues, typically located in and around non-canonical regions, such as bulges and internal loops. This pointed to a much larger concerted conformational change. The involvement of many residues and the much wider range of conformations that are accessible to RNA made it considerably more difficult to deduce the excited state structures of RNA compared to DNA. We therefore pursued parallel relaxation dispersion studies on three distinct RNA systems hoping that this would help unveil common trends or principles that would make it easier to characterize the structures of the excited states; the apical loop of HIV-1 TAR, the HIV-1 stem-loop 1 element, and the ribosomal A-site [124]. A consistent picture started to emerge from qualitative assessment of the ground and excited state carbon chemical shifts in these three RNA systems; flipped out residues in the ground state tended to become more stacked in the excited state while some of the residues that were base-paired in the ground state became less stacked in the excited state. This hinted to conformational exchange involving localized changes in secondary structure.

Fortunately, methods exist that can be used to accurately predict RNA secondary structure based on sequence. Using a secondary structure prediction program [125], we computed secondary structures for each RNA targets. Strikingly, in each case, the ground state was predicted to be the most favorable secondary structure, while the secondary structure that is predicted to be the second most energetically favorable featured rearrangements localized precisely at those residues that exhibit relaxation dispersion, and what's more, the rearrangements could qualitatively explain the differences in chemical shifts between the ground and excited state. Thus, we used a 'mutate and finger print' strategy to test whether these second most favorable secondary structures represent the RNA excited state. In this approach, we trapped a target RNA secondary structure by introducing single mutations [124]. We then used NOEs to confirm that the mutations trap the targeted RNA secondary structure, and then compared the chemical shifts of the trapped RNA with those determined for the excited state using relaxation dispersion. Using this approach, we visualized the secondary structures of the excited states for all three RNA systems, and in all cases, they featured local reshuffling of base-pairing partners in and around non-canonical residues (Fig. 8) [124].

What is the biological significance of the RNA excited states and how might this be established? One of the benefits of using mutations to trap excited states is that one can then preform follow up functional studies and ask how mutations that stabilize excited states affect function as compared to mutations that are predicted to stabilize the ground state or have little effect. Fortunately, much of this mutagenesis work had already been done on our

three RNA systems, not because of any interest in excited states, but as part of general biochemical/functional studies [124]. Strikingly, we found that mutations that are predicted to stabilize the TAR ES over the GS inhibit Tat/Cyclin T1 binding and transcriptional activation whereas mutations that stabilize the A-site ES over the GS increase the rates of ribosomal stop-codon read-through and frame-shifting or inhibit binding of Initiation Factor. These observations could be explained based on the ES structures of TAR and A-site. The TAR ES sequesters four residues (U31, G34, C30, A35) (Fig. 8a) into base-pairs that are otherwise exposed and available for Tat/Cyclin T1 recognition whereas the A-site ES sequesters A92 and A93 into base-pairs making them unavailable to decode incoming tRNAs (Fig. 8b). This points to an expose/sequester type on/off switch analogous to that used by other regulatory RNA switches but that operates on a much smaller localized scale.

To date, we have not worked with an RNA molecule that does not exhibit signs of chemical exchange in non-canonical regions, and it is very likely that many of these transitions will also involve local changes in secondary structure, though we cannot rule out other types of excited state conformations. It is remarkable that the ground and excited state corresponded to the lowest and second lowest energy secondary structures computed using secondary structure prediction programs. Thus, we can expect that secondary structure prediction programs will provide a valuable tool for whole genome analysis of RNA excited states, including assessments of their evolutionary conservation.

Conclusion

Our NMR studies have shown that nucleic acid structures specify directional flexibility that makes it more likely to sample certain conformations as opposed to others. Thus, RNA helices sample a small specific sub-set of inter-helical orientations, and exclude the vast majority of orientations. Likewise, RNA structures can sample a few distinct secondary structures from a large number of possibilities and DNA can switch between Watson-Crick and Hoogsteen base-pairs. Cellular cues in turn take advantage of this pre-existing flexibility to robustly effect specific conformational changes and biological outcomes. Future studies should explore precisely how cellular cues interact with nucleic acids and specifically alter their free energy landscapes as well as examine in greater depth how barriers between conformations might help temporally tune biologically important conformational changes.

Recently, at the Vilcek Prize ceremony recognizing work from American scientists who were born outside the US, American cellist Yo-Yo Ma emphasized the power of bringing together people from diverse cultures by referring to what ecologists call 'edge effects', which describe the advantageous changes in community structures that occur at the boundary of two habitats due to increased biodiversity. In our work, we have found it particularly beneficial to work at the edges between experimental and theoretical aspects of NMR, sample engineering, and computation. Theoretical treatment of motional couplings guided our sample engineering efforts, and the domain-elongation strategy, which turned it possible to combine NMR data with MD simulations to construct atomic resolution ensembles; low field $R_{1\rho}$ spin lock experiments made it possible to probe transient states in uniformly labeled nucleic acids, and in combination with chemical modifications and mutations, allow their structure to be deduced and their biological significance to be established. We can anticipate many further developments along these edges in the future from which a more rigorous quantitative understanding of nucleic acid dynamic behavior will undoubtedly emerge.

Acknowledgments

I would like to dedicate this article to my father, Dr. Murtadha Al-Hashimi, who has profoundly shaped my personal and professional life. I am indebted to many outstanding students and postdoctoral fellows who carried out the work described in this review. I would also like to thank Dr. Alex Kurochkin and Dr. Vivekanandan Subramanian for maintaining NMR instruments at the University of Michigan. Finally, I would like to acknowledge collaborations involving nucleic acids with the groups of Ioan Andricioaei at the University of California- Irvine and Mark Banaszak Holl, Patrick O'Brien, Charles Brooks, Carol Fierke, David Markovitz, and Nils Walter at the University of Michigan. Supported by the US National Institutes of Health (R01AI066975, R01GM089846, R21AI096985, and R21GM096156) and an Agilent award.

References

1. Pauling L. Molecular basis of biological specificity. *Nature*. 1974; 248:769–771. [PubMed: 4599082]
2. Frauenfelder H, Sligar SG, Wolynes PG. The energy landscapes and motions of proteins. *Science*. 1991; 254:1598–1603. [PubMed: 1749933]
3. Cruz JA, Westhof E. The dynamic landscapes of RNA architecture. *Cell*. 2009; 136:604–609. [PubMed: 19239882]
4. Al-Hashimi HM. Dynamics-Based Amplification of RNA Function and Its Characterization by Using NMR Spectroscopy. *ChemBioChem*. 2005; 6:1506–1519. [PubMed: 16138302]
5. Bothe JR, Nikolova EN, Eichhorn CD, Chugh J, Hansen AL, Al-Hashimi HM. Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat Methods*. 2011; 8:919–931. [PubMed: 22036746]
6. 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. [PubMed: 21894962]
7. Bardaro MF Jr, Varani G. Examining the relationship between RNA function and motion using nuclear magnetic resonance. *Wiley Interdiscip Rev: RNA*. 2012; 3:122–132. [PubMed: 22180312]
8. Clery A, Schubert M, Allain FH. NMR spectroscopy: an excellent tool to understand RNA and carbohydrate recognition by proteins. *Chimia (Aarau)*. 2012; 66:741–746. [PubMed: 23146258]
9. Kay LE, Torchia DA, Bax A. Backbone Dynamics of Proteins As Studied by ¹⁵N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry*. 1989; 28:8972–8979. [PubMed: 2690953]
10. Lipari G, Szabo A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J Am Chem Soc*. 1982; 104:4546–4559.
11. Tjandra N, Bax A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*. 1997; 278:1111–1114. [PubMed: 9353189]
12. Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH. NMR evidence for slow collective motions in cyanometmyoglobin. *Nat Struct Biol*. 1997; 4:292–297. [PubMed: 9095197]
13. Zhang Q, Sun X, Watt ED, Al-Hashimi HM. Resolving the motional modes that code for RNA adaptation. *Science*. 2006; 311:653–656. [PubMed: 16456078]
14. Zhang Q, Throolin R, Pitt SW, Serganov A, Al-Hashimi HM. 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. [PubMed: 12940730]
15. Showalter SA, Hall KB. Isotropic reorientational eigenmode dynamics complements NMR relaxation measurements for RNA. *Methods Enzymol*. 2005; 394:465–480. [PubMed: 15808233]
16. Showalter SA, Baker NA, Tang CG, Hall K. Iron responsive element RNA flexibility described by NMR and isotropic reorientational eigenmode dynamics. *J Biomol NMR*. 2005; 32:179–193. [PubMed: 16132819]
17. Musselman C, Al-Hashimi HM, Andricioaei I. iRED analysis of TAR RNA reveals motional coupling, long-range correlations, and a dynamical hinge. *Biophys J*. 2007; 93:411–422. [PubMed: 17449677]

18. Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH. Nuclear Magnetic Dipole Interactions in Field-Oriented Proteins – Information For Structure Determination in Solution. *Proc Natl Acad Sci USA*. 1995; 92:9279–9283. [PubMed: 7568117]
19. Tolman JR, Ruan K. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem Rev*. 2006; 106:1720–1736. [PubMed: 16683751]
20. Peti W, Meiler J, Bruschweiler R, Griesinger C. Model-free analysis of protein backbone motion from residual dipolar couplings. *J Am Chem Soc*. 2002; 124:5822–5833. [PubMed: 12010057]
21. Getz M, Sun X, Casiano-Negroni A, Zhang Q, Al-Hashimi HM. NMR studies of RNA dynamics and structural plasticity using NMR residual dipolar couplings. *Biopolymers*. 2007; 86:384–402. [PubMed: 17594140]
22. Bax A, Grishaev A. Weak alignment NMR: a hawk-eyed view of biomolecular structure. *Curr Opin Struct Biol*. 2005; 15:563–570. [PubMed: 16140525]
23. Ramirez BE, Bax A. Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium. *J Am Chem Soc*. 1998; 120:9106–9107.
24. Al-Hashimi HM, Valafar H, Terrell M, Zartler ER, Eidsness MK, Prestegard JH. Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J Magn Reson*. 2000; 143:402–406. [PubMed: 10729267]
25. Yao L, Bax A. Modulating protein alignment in a liquid-crystalline medium through conservative mutagenesis. *J Am Chem Soc*. 2007; 129:11326–11327. [PubMed: 17718572]
26. Bondensgaard K, Mollova ET, Pardi A. The global conformation of the hammerhead ribozyme determined using residual dipolar couplings. *Biochemistry*. 2002; 41:11532–11542. [PubMed: 12269797]
27. Latham MP, Hanson P, Brown DJ, 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. [PubMed: 18026844]
28. Bothner-By, AA. Magnetic field induced alignment of molecules. In: Grant, DM.; Harris, RK., editors. *Encyclopedia of Nuclear Magnetic Resonance*. Wiley: Chichester; 1995. p. 2932-2938.
29. Al-Hashimi HM, Majumdar A, Gorin A, Kettani A, Skripkin E, Patel DJ. Field- and phage-induced dipolar couplings in a homodimeric DNA quadruplex, relative orientation of G center dot(C-A) triad and G-tetrad motifs and direct determination of C2 symmetry axis orientation. *J Am Chem Soc*. 2001; 123:633–640. [PubMed: 11456575]
30. Dethoff EA, Hansen AL, Zhang Q, Al-Hashimi HM. Variable helix elongation as a tool to modulate RNA alignment and motional couplings. *J Magn Reson*. 2010; 202:117–121. [PubMed: 19854083]
31. Getz MM, Andrews AJ, Fierke CA, Al-Hashimi HM. Structural plasticity and Mg²⁺ binding properties of RNase P P4 from combined analysis of NMR residual dipolar couplings and motionally decoupled spin relaxation. *RNA*. 2007; 13:251–256. [PubMed: 17194721]
32. Sun X, Zhang Q, Al-Hashimi HM. Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg²⁺ binding: role in the kissing-duplex structural transition. *Nucleic Acids Res*. 2007; 35:1698–1713. [PubMed: 17311812]
33. Shajani Z, Drobny G, Varani G. Binding of U1A Protein Changes RNA Dynamics As Observed by (13)C NMR Relaxation Studies. *Biochemistry*. 2007; 46:5875–5883. [PubMed: 17469848]
34. Bardaro MF Jr, Shajani Z, Patora-Komisarska K, Robinson JA, Varani G. How binding of small molecule and peptide ligands to HIV-1 TAR alters the RNA motional landscape. *Nucleic Acids Res*. 2009; 37:1529–1540. [PubMed: 19139066]
35. Bardaro MF Jr, Varani G. Independent alignment of RNA for dynamic studies using residual dipolar couplings. *J Biomol NMR*. 2012; 54:69–80. [PubMed: 22806132]
36. Muesing MA, Smith DH, Capon DJ. Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. *Cell*. 1987; 48:691–701. [PubMed: 3643816]
37. Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*. 1992; 257:76–80. [PubMed: 1621097]
38. Aboul-ela F, Karn J, Varani G. The Structure of the Human-Immunodeficiency-Virus Type-1 Tar RNA Reveals Principles of RNA Recognition By Tat Protein. *J Mol Biol*. 1995; 253:313–332. [PubMed: 7563092]

39. Aboul-ela F, Karn J, Varani G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res.* 1996; 24:3974–3981. [PubMed: 8918800]
40. Ippolito JA, Steitz TA. A 1.3-angstrom resolution crystal structure of the HIV-1 transactivation response region RNA stem reveals a metal ion- dependent bulge conformation. *Proc Nat Acad Sci USA.* 1998; 95:9819–9824. [PubMed: 9707559]
41. Faber C, Sticht H, Schweimer K, Rosch P. Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. *J Biol Chem.* 2000; 275:20660–20666. [PubMed: 10747964]
42. Du Z, Lind KE, James TL. Structure of TAR RNA complexed with a Tat-TAR interaction nanomolar inhibitor that was identified by computational screening. *Chem Biol.* 2002; 9:707–712. [PubMed: 12079782]
43. Davis B, Afshar M, Varani G, Murchie AI, Karn J, Lentzen G, Drysdale M, Bower J, Potter AJ, Starkey ID, Swarbrick T, Aboul-ela F. Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic “hot spots”. *J Mol Biol.* 2004; 336:343–356. [PubMed: 14757049]
44. Murchie AI, Davis B, Isel C, Afshar M, Drysdale MJ, Bower J, Potter AJ, Starkey ID, Swarbrick TM, Mirza S, Prescott CD, Vaglio P, Aboul-ela F, Karn J. Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. *J Mol Biol.* 2004; 336:625–638. [PubMed: 15095977]
45. Hermann T, Patel DJ. Adaptive recognition by nucleic acid aptamers. *Science.* 2000; 287:820–825. [PubMed: 10657289]
46. Williamson JR. Induced fit in RNA-protein recognition. *Nat Struct Biol.* 2000; 7:834–837. [PubMed: 11017187]
47. Leulliot N, Varani G. Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture. *Biochemistry.* 2001; 40:7947–7956. [PubMed: 11434763]
48. Hammes GG, Chang YC, Oas TG. Conformational selection or induced fit: a flux description of reaction mechanism. *Proc Natl Acad Sci USA.* 2009; 106:13737–13741. [PubMed: 19666553]
49. Al-Hashimi HM, Gosser Y, Gorin A, Hu W, Majumdar A, Patel DJ. Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. *J Mol Biol.* 2002; 315:95–102. [PubMed: 11779230]
50. Palmer AG 3rd. NMR characterization of the dynamics of biomacromolecules. *Chem Rev.* 2004; 104:3623–3640. [PubMed: 15303831]
51. Pitt SW, Majumdar A, Serganov A, Patel DJ, Al-Hashimi HM. Argininamide binding arrests global motions in HIV-1 TAR RNA: comparison with Mg²⁺-induced conformational stabilization. *J Mol Biol.* 2004; 338:7–16. [PubMed: 15050819]
52. Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J Am Chem Soc.* 1990; 112:4989–4991.
53. Hansen AL, Al-Hashimi HM. Dynamics of large elongated RNA by NMR carbon relaxation. *J Am Chem Soc.* 2007; 129:16072–16082. [PubMed: 18047338]
54. Nikolova EN, Al-Hashimi HM. Preparation, resonance assignment, and preliminary dynamics characterization of residue specific ¹³C/¹⁵N-labeled elongated DNA for the study of sequence-directed dynamics by NMR. *J Biomol NMR.* 2009; 45:9–16. [PubMed: 19636798]
55. Meiler J, Prompers JJ, Peti W, Griesinger C, Bruschweiler R. Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J Am Chem Soc.* 2001; 123:6098–6107. [PubMed: 11414844]
56. Al-Hashimi HM, Gorin A, Majumdar A, Patel DJ. Alignment of the HTLV-I rex peptide bound to its target RNA aptamer from magnetic field-induced residual dipolar couplings and intermolecular hydrogen bonds. *J Am Chem Soc.* 2001; 123:3179–3180. [PubMed: 11457047]
57. Al-Hashimi HM, Tolman JR, Majumdar A, Gorin A, Patel DJ. Determining stoichiometry in homomultimeric nucleic acid complexes using magnetic field induced residual dipolar couplings. *J Am Chem Soc.* 2001; 123:5806–5807. [PubMed: 11403617]

58. Tjandra N, Garrett DS, Gronenborn AM, Bax A, Clore GM. 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. [PubMed: 9187651]
59. Zhang Q, Al-Hashimi HM. Extending the NMR spatial resolution limit for RNA by motional couplings. *Nat Methods.* 2008; 5:243–245. [PubMed: 18246076]
60. Fisher CK, Al-Hashimi HM. Approximate reconstruction of continuous spatially complex domain motions by multialignment NMR residual dipolar couplings. *J Phys Chem B.* 2009; 113:6173–6176. [PubMed: 19358547]
61. Fisher CK, Zhang Q, Stelzer A, Al-Hashimi HM. Ultrahigh resolution characterization of domain motions and correlations by multialignment and multireference residual dipolar coupling NMR. *J Phys Chem B.* 2008; 112:16815–16822. [PubMed: 19367865]
62. Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM. Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature.* 2007; 450:1263–1267. [PubMed: 18097416]
63. Saupe A. Recent results in the field of liquid crystals. *Angew. Chem Int Ed Engl.* 1968; 7:97–112.
64. Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, Zoroddu MA. Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc Natl Acad Sci.* 2004; 101:6841–6846. [PubMed: 15100408]
65. Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L, Yuan J. Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *J Am Chem Soc.* 2007; 129:12786–12794. [PubMed: 17910448]
66. Hansen MR, Mueller L, Pardi A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat Struct Biol.* 1998; 5:1065–1074. [PubMed: 9846877]
67. Clore GM, Starich MR, Gronenborn AM. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J Am Chem Soc.* 1998; 120:10571–10572.
68. Clore GM, Schwieters CD. Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry.* 2004; 43:10678–10691. [PubMed: 15311929]
69. Schwieters CD, Clore GM. A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. *Biochemistry.* 2007; 46:1152–1166. [PubMed: 17260945]
70. Frank AT, Stelzer AC, Al-Hashimi HM, 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. [PubMed: 19369218]
71. Bailor MH, Sun X, Al-Hashimi HM. Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science.* 2010; 327:202–206. [PubMed: 20056889]
72. Bailor MH, Mustoe AM, Brooks CL 3rd, Al-Hashimi HM. Topological constraints: using RNA secondary structure to model 3D conformation, folding pathways, and dynamic adaptation. *Curr Opin Struct Biol.* 2011; 21:296–305. [PubMed: 21497083]
73. Mustoe AM, Bailor MH, Teixeira RM, Brooks CL 3rd, Al-Hashimi HM. New insights into the fundamental role of topological constraints as a determinant of two-way junction conformation. *Nucleic Acids Res.* 2012; 40:892–904. [PubMed: 21937512]
74. Chu VB, Lipfert J, Bai Y, Pande VS, Doniach S, Herschlag D. Do conformational biases of simple helical junctions influence RNA folding stability and specificity? *RNA.* 2009; 15:2195–2205. [PubMed: 19850914]
75. MacKerell AD, Banavali N, Foloppe N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers.* 2000; 56:257–265. [PubMed: 11754339]
76. Chen Y, Campbell SL, Dokholyan NV. Deciphering protein dynamics from NMR data using explicit structure sampling and selection. *Biophys J.* 2007; 93:2300–2306. [PubMed: 17557784]
77. Stelzer AC, Frank AT, Bailor MH, Andricioaei I, Al-Hashimi HM. Constructing atomic-resolution RNA structural ensembles using MD and motionally decoupled NMR RDCs. *Methods.* 2009; 49:167–173. [PubMed: 19699798]

78. Salmon L, Bascom G, Andricioaei I, Al-Hashimi HM. 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. [PubMed: 23473378]
79. 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.
80. Shaw DE, Deneroff MM, Dror RO, Kuskin JS, Larson RH, Salmon JK, Young C, Batson B, Bowers KJ, Chao JC. Anton, a special-purpose machine for molecular dynamics simulation. *Commun ACM.* 2008; 51:91–97.
81. Denning EJ, Priyakumar UD, Nilsson L, Mackerell AD Jr. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: Update of the CHARMM all-atom additive force field for RNA. *J Comput Chem.* 2011; 32:1929–1943. [PubMed: 21469161]
82. Foloppe N, MacKerell AD Jr. All – atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem.* 2000; 21:86–104.
83. Wijmenga SS, Kruithof M, Hilbers CW. Analysis of (1)H chemical shifts in DNA: Assessment of the reliability of (1)H chemical shift calculations for use in structure refinement. *J Biomol NMR.* 1997; 10:337–350. [PubMed: 20859781]
84. Frank AT, Horowitz S, Andricioaei I, Al-Hashimi HM. Utility of (1)H NMR Chemical Shifts in Determining RNA Structure and Dynamics. *J Phys Chem B.* 2013
85. Musselman C, Zhang Q, Al-Hashimi HM, Andricioaei I. Referencing strategy for the direct comparison of nuclear magnetic resonance and molecular dynamics motional parameters in RNA. *J Phys Chem B.* 2010; 114:929–939. [PubMed: 20039757]
86. Eichhorn CD, Feng J, Suddala KC, Walter NG, Brooks CL 3rd, Al-Hashimi HM. Unraveling the structural complexity in a single-stranded RNA tail: implications for efficient ligand binding in the prequeuosine riboswitch. *Nucleic Acids Res.* 2012; 40:1345–1355. [PubMed: 22009676]
87. Wong V, Case DA. Evaluating rotational diffusion from protein MD simulations. *J Phys Chem B.* 2008; 112:6013–6024. [PubMed: 18052365]
88. Chen AA, Pappu RV. Parameters of monovalent ions in the AMBER-99 forcefield: assessment of inaccuracies and proposed improvements. *J Phys Chem B.* 2007; 111:11884–11887. [PubMed: 17887792]
89. Casiano-Negroni A, Sun X, Al-Hashimi HM. Probing Na(+)-induced changes in the HIV-1 TAR conformational dynamics using NMR residual dipolar couplings: new insights into the role of counterions and electrostatic interactions in adaptive recognition. *Biochemistry.* 2007; 46:6525–6535. [PubMed: 17488097]
90. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science.* 2008; 320:1471–1475. [PubMed: 18556554]
91. Ban D, Funk M, Gulich R, Egger D, Sabo TM, Walter KF, Fenwick RB, Giller K, Pichierri F, de Groot BL, Lange OF, Grubmuller H, Salvatella X, Wolf M, Loidl A, Kree R, Becker S, Lakomek NA, Lee D, Lunkenheimer P, Griesinger C. Kinetics of conformational sampling in ubiquitin. *Angew Chem Int Ed Engl.* 2011; 50:11437–11440. [PubMed: 22113802]
92. Palmer AG 3rd, Kroenke CD, Loria JP. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* 2001; 339:204–238. [PubMed: 11462813]
93. Korzhnev DM, Kay LE. Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Acc Chem Res.* 2008; 41:442–451. [PubMed: 18275162]
94. Sekhar A, Kay LE. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc Natl Acad Sci USA.* 2013
95. Korzhnev DM, Religa TL, Banachewicz W, Fersht AR, Kay LE. A transient and low-populated protein-folding intermediate at atomic resolution. *Science.* 2010; 329:1312–1316. [PubMed: 20829478]

96. Neudecker P, Robustelli P, Cavalli A, Walsh P, Lundstrom P, Zarrine-Afsar A, Sharpe S, Vendruscolo M, Kay LE. Structure of an intermediate state in protein folding and aggregation. *Science*. 2012; 336:362–366. [PubMed: 22517863]
97. Eisenmesser EZ, Bosco DA, Akke M, Kern D. Enzyme dynamics during catalysis. *Science*. 2002; 295:1520–1523. [PubMed: 11859194]
98. Li P, Martins IR, Amarasinghe GK, Rosen MK. Internal dynamics control activation and activity of the autoinhibited Vav DH domain. *Nat Struct Mol Biol*. 2008; 15:613–618. [PubMed: 18488041]
99. Boehr DD, McElheny D, Dyson HJ, Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*. 2006; 313:1638–1642. [PubMed: 16973882]
100. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, Kay LE. Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature*. 2004; 430:586–590. [PubMed: 15282609]
101. Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE. Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science*. 2013; 339:1080–1083. [PubMed: 23393091]
102. Yamazaki T, Muhandiram R, Kay LE. NMR Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly ^{13}C , ^{15}N -Labeled Proteins: Application to $^{13}\text{C}\alpha$ Carbons. *J Am Chem Soc*. 1994; 116:8266–8278.
103. Johnson JE Jr, Hoogstraten CG. Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using ^{13}C NMR spin relaxation and specific isotope labeling. *J Am Chem Soc*. 2008; 130:16757–16769. [PubMed: 19049467]
104. Massi F, Johnson E, Wang CY, Rance M, Palmer AG. NMR R-1 rho rotating-frame relaxation with weak radio frequency fields. *J Am Chem Soc*. 2004; 126:2247–2256. [PubMed: 14971961]
105. Korzhnev DM, Orekhov VY, Kay LE. Off-resonance R(1rho) NMR studies of exchange dynamics in proteins with low spin-lock fields: an application to a Fyn SH3 domain. *J Am Chem Soc*. 2005; 127:713–721. [PubMed: 15643897]
106. Hansen AL, Nikolova EN, Casiano-Negroni A, Al-Hashimi HM. Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R(1rho) NMR spectroscopy. *J Am Chem Soc*. 2009; 131:3818–3819. [PubMed: 19243182]
107. Pelupessy P, Chiarparin E, Bodenhausen G. Excitation of selected proton signals in NMR of isotopically labeled macromolecules. *J Magn Reson*. 1999; 138:178–181. [PubMed: 10329243]
108. Ferrage F, Eykyn TR, Bodenhausen G. Frequency-switched single-transition cross-polarization: a tool for selective experiments in biomolecular NMR. *ChemPhysChem*. 2004; 5:76–84. [PubMed: 14999846]
109. Miloushev VZ, Palmer AG 3rd. R(1rho) relaxation for two-site chemical exchange: general approximations and some exact solutions. *J Magn Reson*. 2005; 177:221–227. [PubMed: 16143548]
110. Greene KL, Wang Y, Live D. Influence of the glycosidic torsion angle on ^{13}C and ^{15}N shifts in guanosine nucleotides: investigations of G-tetrad models with alternating syn and anti bases. *J Biomol NMR*. 1995; 5:333–338. [PubMed: 7647551]
111. Xu XP, Au-Yeung SCF. Investigation of chemical shift and structure relationships in nucleic acids using NMR and density functional theory methods. *J Phys Chem B*. 2000; 104:5641–5650.
112. Hoogsteen K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta crystallographica Section A, Foundations of crystallography*. 1963; 16:907.
113. Hoogsteen K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta crystallographica*. 1959; 12:822–823.
114. Nikolova EN, Goh GB, Brooks CL 3rd, Al-Hashimi HM. Characterizing the Protonation State of Cytosine in Transient G.C Hoogsteen Base Pairs in Duplex DNA. *J Am Chem Soc*. 2013
115. Nikolova EN, Gottardo FL, Al-Hashimi HM. Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J Am Chem Soc*. 2012; 134:3667–3670. [PubMed: 22309937]

116. Abrescia NG, Thompson A, Huynh-Dinh T, Subirana JA. Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc Natl Acad Sci USA*. 2002; 99:2806–2811. [PubMed: 11880632]
117. Ughetto G, Wang AH, Quigley GJ, van der Marel GA, van Boom JH, Rich A. A comparison of the structure of echinomycin and triostin A complexed to a DNA fragment. *Nucleic Acids Res*. 1985; 13:2305–2323. [PubMed: 4000957]
118. Rice PA, Yang S, Mizuuchi K, Nash HA. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*. 1996; 87:1295–1306. [PubMed: 8980235]
119. Nikolova EN, Zhou H, Gottardo FL, Alvey HS, Kimsey IJ, Al-Hashimi HM. A historical account of hoogsteen base-pairs in duplex DNA. *Biopolymers*. 2013
120. Nair DT, Johnson RE, Prakash L, Prakash S, Aggarwal AK. Human DNA polymerase ϵ incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure*. 2005; 13:1569–1577. [PubMed: 16216587]
121. Hoogstraten CG, Wank JR, Pardi A. Active site dynamics in the lead-dependent ribozyme. *Biochemistry*. 2000; 39:9951–9958. [PubMed: 10933815]
122. Venditti V, Clos L 2nd, Niccolai N, Butcher SE. Minimum-energy path for a u6 RNA conformational change involving protonation, base-pair rearrangement and base flipping. *J Mol Biol*. 2009; 391:894–905. [PubMed: 19591840]
123. Reiter NJ, Blad H, Abildgaard F, Butcher SE. Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry*. 2004; 43:13739–13747. [PubMed: 15504036]
124. Dethoff EA, Petzold K, Chugh J, Casiano-Negroni A, Al-Hashimi HM. Visualizing transient low-populated structures of RNA. *Nature*. 2012; 491:724–728. [PubMed: 23041928]
125. Parisien M, Major F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*. 2008; 452:51–55. [PubMed: 18322526]
126. Cornilescu G, Marquardt JL, Ottiger M, Bax A. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J Am Chem Soc*. 1998; 120:6836–6837.

Highlights

- Domain-elongation, RDCs, and MD simulations yield ensembles of RNA.
- Relaxation dispersion and mutagenesis yield structures of DNA and RNA excited states
- Directional flexibility allows nucleic acids to change conformation specifically.

Fig. 1. Domain-elongation strategy for decoupling internal and overall motions in RNA. (a) Collective motions of helices can lead to coupled changes in the overall rotational diffusion (D_{zz}) and/or partial alignment (S_{zz}) of the RNA, which can be decoupled by (b) domain elongation. (c) Isotopic labeling strategy for invisible elongation of RNA. $^{13}\text{C}/^{15}\text{N}$ labeled and unlabeled residues are shown in color and grey respectively. (d) Comparison of 2D ^1H - ^{13}C HSQC spectra of elongated (in red) and non-elongated TAR (in black) in the free state and when bound to the ligand argininamide.

Fig. 2. Picosecond-to-nanosecond motions in HIV-1 TAR by domain-elongation and spin relaxation measurements. (a) HIV-1 TAR conformational adaptation upon binding to different ligands and small molecules. (b) Resolving local and collective motions in TAR. Shown are normalized sugar and base C-H resonance intensities (peak heights) measured from 2D ^1H - ^{13}C HSQC spectra E-TAR and TAR. (c) The ratio (R_2/R_1) of imino ^{15}N transverse (R_2) to longitudinal (R_1) relaxation rates measured for guanine (filled circles) and uridine (filled diamonds) residues in elongated (E-TAR) and non-elongated TAR (TAR) in free and argininamide (ARG) bound states. (d) Directional flexibility specify ligand-bound TAR conformations. Correlation plot between the mean-angular-difference ($\langle \Delta\theta \rangle$) in the orientation of sugar and base C-H bond vectors across eight different HIV-1 TAR structures and the corresponding free E-TAR normalized resonance intensities.

Fig. 3.

Simultaneous anchoring of NMR frames onto individual helices in a two-helix RNA using motional couplings. RDCs measured in a given helix (indicated with a star) are modulated by collective motions, which lead to coupled changes in the total χ -tensor governing overall alignment of the RNA relative to the magnetic field. This provides a basis for measuring helix motions relative to frames of reference that are anchored along the χ -tensor of each helix. The angle β is the bend angle between the two helices, and α and γ specify rotations around the helix axes of HI and HII respectively.

Fig. 4.

Measurement of RNA helix motions using domain elongation and RDCs. (a) NMR reference frames are anchored onto individual helices using domain-elongation (indicated with arrow). (b) Quality factor (Q) [126] for HI (in black) and HII (in orange) as a function of the inter-helical ensemble size N . Also shown is comparison of measured RDCs and values back-predicted using the best-fit three-state ensemble. (c) The three-state TAR ensembles (in green) in which each conformer is specified by three Euler angles and the TAR conformations (in gray) bound to peptide derivatives of Tat (1ARJ) [37, 38], divalent ions (397D) [40], and five different small molecules (1QD3, 1LVJ, 1UUD, 1UUI, 1UTS) [41–44]. Shown on each 2D plane is the correlation coefficient (R) between angles for the ligand bound conformations. (d) The three-state TAR ensemble implies a spatially structured motional trajectory in which the two helices bend and twist in a correlated manner.

Fig. 5. Topological confinement of A-form helices in RNA. (a) 3D inter-helical orientation maps showing for bulges that are four (S_4S_0), three (S_3S_0 , HIV-1 TAR), two (S_2S_0 , HIV-2 TAR), and one (S_1S_0) nucleotide long. The NMR-MD, PDB-derived, and topologically computed inter-helical distributions are shown in blue, red, and grey, respectively. (b) The topologically allowed inter-helical orientations computed for two different junction topologies showing correlations between the twist angles α and γ for the RNA two-way junctions but not for two centrally connected cylinders. (c) Topological constraints result in correlations between twisting and bending motions in RNA two-way junctions.

Fig. 6. Constructing HIV-1 TAR dynamic ensembles using variable elongation RDCs, MD simulations, and structure based prediction of alignment. (a) Flow chart describing the 'sample and select' scheme. (b) Modulation of TAR alignment by partial helix elongation. Shown are secondary structures for four differentially elongated TAR samples used to measure multiple sets of RDCs. (c) Comparison of measured and predicted RDCs in terms of reduced χ^2 as a function of the size of the ensemble (N). (d) Comparison of RDCs measured in the four TAR constructs (color coded according to Fig. 6b) and those predicted for the MD trajectory and an ensemble of 20 conformers selected using all of the RDCs. (e) Comparison of conformers from the 20-state TAR ensemble (in orange) and ligand bound TAR conformations (in grey).

Fig. 7. Probing exchange between Watson-Crick and Hoogsteen base-pairs in canonical duplex DNA using $R_{1\rho}$ relaxation dispersion. (a) Selective ^{13}C $R_{1\rho}$ pulse sequence for quantifying microsecond- to-millisecond exchange in nucleic acids. (b) DNA duplex highlighting A-T and G-C base-pairs with significant carbon chemical exchange. Shown are on-resonance and off-resonance ^{13}C $R_{1\rho}$ relaxation dispersion profiles as a function of spinlock power ($\omega_{\text{eff}}/2\pi$) and both spin lock power ($\omega_{\text{eff}}/2\pi$) and offset ($\Omega_{\text{eff}}/2\pi$) measured at 14.1 T and pH 6.8. (c) Comparison of excited state carbon chemical shifts measured by relaxation dispersion and chemical shifts measured for trapped Hoogsteen base-pairs by N1-methylating guanine and computed for Hoogsteen base-pairs using DFT, all referenced relative to the Watson-Crick ground state chemical shift (arbitrarily set to 0 ppm). (d) Equilibrium between WC and HG A-T and G-C base-pairs showing the relative populations and exchange rate constants at pH 6.8 and the estimated pKa for a HG G-C base-pair [114].

Fig. 8. Probing chemical exchange in RNA involving local changes in secondary structures using $R_{1\rho}$ relaxation dispersion. Shown are two examples involving (a) the apical loop in HIV-1 TAR and (b) the ribosomal A-site. The conformational transitions serve to sequester or expose functionally important residues that are involved in binding proteins or other RNAs.