Journal of Biomolecular NMR (2005) 32: 179–193
DOI 10.1007/s10858-005-7948-2

Article

Iron responsive element RNA flexibility described by NMR and isotropic reorientational eigenmode dynamics

Scott A. Showalter[a], Nathan A. Baker[a,b], Changguo Tang[a] & Kathleen B. Hall[a,*]
[a]*Department of Biochemistry & Molecular Biophysics;* [b]*Center for Computational Biology, Washington University School of Medicine, St Louis, MO, 63110, USA*

Received 3 December 2004; Accepted 1 May 2005

Key words: IRE RNA, molecular dynamics, NMR structure and dynamics, RED analysis

Abstract

The first example of the application of reorientational eigenmode dynamics (RED) to RNA is shown here for the small and floppy Iron Responsive Element (IRE) RNA hairpin. Order parameters calculated for bases and riboses from a 12 ns molecular dynamics trajectory are compared to experimentally determined order parameters from ^{13}C-^{1}H NMR relaxation experiments, and shown to be in qualitative agreement. Given the small size of the IRE hairpin and its very flexible loop, isotropic RED (iRED) was also used to analyze the trajectory in order to describe its dynamic motions. iRED analysis shows that the global and internal dynamics of the IRE are not rigorously separable, which will result in inaccurate experimental order parameters. In addition, the iRED analysis described the many correlated motions that comprise the dynamics of the IRE RNA. The combined use of NMR relaxation, RED, and iRED provide a uniquely detailed description of IRE RNA dynamics.

Introduction

Short RNA sequences that participate in intermolecular interactions are often displayed as single stranded structures in order to offer the maximum number of physical contacts with the biological partner. Structural characterization of small RNAs is best done using solution NMR (e.g. TAR (Puglisi et al., 1992), the GNRA tetraloop (Heus et al., 1991), and the tetraloop receptor (Butcher et al., 1997)), but some of these RNAs appear to be quite flexible, such as the U6 intramolecular stem-loop (Reiter et al., 2004); the IRE loop (Laing and Hall, 1996); and the active site of ribozymes (Hoogstraten et al., 2000). Thus, equally important is an appreciation of RNA dynamic motion, which can be addressed in part by NMR relaxation methods (King et al., 1995; Hall and Tang, 1998; Dayie et al., 2002). To obtain a physical description of the rapid (ps-ns) motions present in an RNA molecule, the most comprehensive description is potentially available from molecular dynamics simulations (Miller and Kollman, 1997; Cheng et al., 2004; Hall and Williams, 2004), with the caveats that motions are sufficiently sampled and accurately portrayed.

The Iron Responsive Element (IRE) RNA hairpin is examined here by a combination of NMR relaxation experiments and molecular dynamics simulations. It is an example of a small and flexible RNA hairpin loop that is the binding site for a protein. One copy of the IRE is present in the 5′-untranslated region of the ferritin mRNA, and five copies are found in the 3′-untranslated region of the transferrin mRNA; in both sites, it is

*To whom correspondence should be addressed. E-mail: hall@bionmr3.wustl.edu

bound by the Iron Responsive Element-Binding Protein (IRE-BP). The IRE loop sequence C6A7G8U9G10C11, sits on an AU loop closing base pair, which is conserved in the ferritin mRNA (Casey et al., 1988) (Structure shown in Figure 1; numbering refers to the hairpin construct used in these experiments). NMR structures of the IRE hairpin loop showed a base pair formed between Cytidine 6 and Guanosine 10, which accounts for the higher than predicted melting temperature of the IRE hairpin (Laing and Hall, 1996; Addess et al., 1997). As a consequence of this base pair interaction, Cytidine 11 is extruded from the loop. Similarly, the Uridine 9 base was not located within the loop structure, and structures generated from the NMR data placed it in solution at the turn of the loop. On the 5′ side of the loop, Adenosine 7 was stacked over Cytosine 6, but the Guanosine 8 base did not appear to be proximal to any element of the loop. The ribose of A7 is best described by an alternating *C2′-endo*/*C3′-endo* pucker (although the IRE structure of Addess et al. (1997) shows it to be *C3′-endo*). Addess et al. (1997) measured the $^{3}J_{H1',H2'}$ of the loop riboses, and concluded that G8 and U9 were *C2′-endo*, while G10 and C11 riboses interconvert between the two puckers. The repuckering of the sugars and the absence of physical restraints for four of six bases in the loop has led to the description of this structure as "floppy."

The IRE-BP recognizes the short IRE RNA loop sequence, 5′CAGUGN, where N is C, U, or A, in addition to a conserved bulge in the IRE hairpin stem (Barton et al., 1990; Leibold et al., 1990; Bettany et al., 1992; Jaffrey et al., 1993; Kikinis et al., 1995). The mechanism of RNA:protein interaction is not known; the 90 kDa phylogenetically conserved metazoan IRE-BP is presumed to be structurally homologous to the mitochondrial aconitase (Robbins et al., 1989), and the binding site for RNA was modeled in the cleft of the two domains (Basilion et al., 1994; Hirling et al., 1994; Kaldy et al., 1999). There are no biophysical data describing this RNA:protein interaction, but based on the floppy loop structure of the IRE in solution, we postulate that loop flexibility is an important feature of sequence-specific recognition.

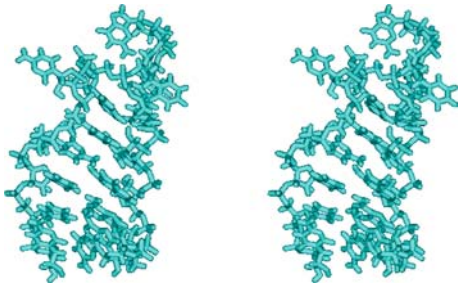

Figure 1. A structure of the IRE RNA consistent with NMR structural data, in which the base of Guanosine 8 is unstacked, and those of Uridine 9 and Cytosine 11 are solvent-exposed. Stereo figure is generated by MOLMOL (Kraulis, 1991).

The lack of NMR constraints to describe the loop structure suggested that several bases in the loop were mobile. To measure the dynamic motions of the bases, a ^{13}C NMR relaxation study was carried out on the IRE hairpin (Hall and Tang, 1998), observing only the purine bases. Results showed that at 20 °C, where the NMR structure was solved, all purines in the loop appeared to experience the same amplitude of dynamic motion. The implication is that although the G8 base could not be located within a structured loop, it did not appear to be more flexible than the G10 base in its C6:G10 base pair. However, at 37 °C, approximately 20° below the melting temperature in this buffer (30 mM NaCl, 10 mM sodium phosphate, pH 6), ^{13}C NMR data showed a significant increase in the motions experienced by all three purine bases in the loop, indicating that the entire loop had become more mobile on the ps-ns timescale (Hall and Tang, 1998).

Subsequent fluorescence studies of IRE loops used 2-aminopurine in place of either A7 or G8 (Hall and Williams, 2004). Estimates of the timescale of motions of these two bases were made on the basis of time-resolved fluorescence. 2AP at the A7 position appeared to be mostly stacked on its C6:G10 neighbor, with only occasional excursions away (on the ns timescale); 2AP at position G8 was best described as undergoing a significant proportion of its dynamic motions on the ns timescale. A physical description of the motions in the IRE was provided by several stochastic dynamics simulations (Hall and Williams, 2004), which showed that G8 made large amplitude excursions away from the other nucleotides in the loop, while its neighboring A7 was far more restricted.

The dynamic properties of short RNA sequences are a necessary component of their bio-

logical function, since nucleotides must be accessible to an incoming protein in order to provide sequence-specific recognition and high affinity binding. While RNA structure can be solved using NMR, descriptions of the RNA dynamics using NMR are usually incomplete, given the span of timescales present, the number of applicable NMR experiments, and the paucity of isolated NMR relaxation-active reporters. Analysis of NMR relaxation data typically relies on the Lipari–Szabo model-free formalism (Lipari and Szabo, 1982a, b), which in its basic form produces a correlation time for global tumbling (τ_c) as well as an order parameter (S^2) representing the amplitude of motion and a local correlation time (τ_e) associated with that motion for each spin pair (typically $^{15}N—^{1}H$ or $^{13}C—^{1}H$).

One necessary condition for application of Lipari–Szabo formalism is that the global reorientation of the system be separable from the internal motion. To a good approximation, this condition suggests that the internal motion should occur on a timescale at least an order of magnitude faster than global tumbling if the Lipari–Szabo model is to be applied with confidence. The IRE RNA is small enough that it tumbles with a rotational correlation time of 3.6 ns at 20 °C, and 2.4 ns at 37 °C (Hall and Tang, 1998). Unfortunately, both NMR relaxation data and independently measured fluorescence data indicate that several bases which protrude from the core structure are moving on timescales similar to the overall tumbling time of the IRE. These conditions could indicate correlation between internal and global motion and thus invalidation of the separability assumption, rendering calculated Lipari–Szabo order parameters inaccurate. The most likely outcome in this case would be an overestimation of the order parameter for most sites (Vugmeyster et al., 2003).

NMR relaxation measurements are not the only experiments capable of generating order parameters and other dynamic information. The recent development of the Reorientational Eigenmode Dynamics (RED) formalism allows for facile calculation of order parameters and correlation functions for internal motions from molecular dynamics trajectories, and allows for their comparison with results obtained from NMR experiments (Prompers and Bruschweiler, 2001). In addition, RED calculations extend the description of NMR relaxation active local motions by unmasking correlations between them. Assuming that the molecular dynamics trajectory is of sufficient length to sample motions on timescales comparable to those accessed by standard NMR dynamics (ps-ns), then RED analysis can provide a detailed description of dynamic motion in a molecule.

Analogous to Cartesian principle component analysis (Kitao et al., 1991) or essential dynamics (Amadei et al., 1993), the RED formalism is applied by constructing a covariance matrix from an MD trajectory with diagonal elements describing positional fluctuations at a subset of sites interesting to the investigator, and off-diagonals describing the covariance between those sites. Unlike the other commonly applied covariance methods, which tend to focus on the Cartesian position of sites within the molecule, RED analysis tracks the (NMR relaxation active) reorientational fluctuations of internuclear bond vectors, typically corresponding to those which the investigator can study experimentally by NMR methods. The diagonal elements of the RED covariance matrix are directly proportional to the Lipari–Szabo order parameter for each site and thus the link between MD and NMR experiments is established.

The information encoded in the RED covariance matrix is much richer than the order parameters, however, in that it also describes the potential correlations between the reorientations of each site. These correlations are most readily visualized through calculation of the eigenvectors of the matrix. These eigenvectors can be represented such that the fractional contribution to the order parameter (δS^2) of the motion described by the vector is displayed for each spin pair. For example, a single site in a molecule (corresponding to a ^{13}C-^{1}H NMR vector) may have motions that are uncorrelated to any other site; such isolated motion would be identified by an eigenvector that contained a contribution from this site and no other, neither would this site contribute any amplitude to other eigenvectors. In contrast, it is probable that there are sites that experience motion predominantly or exclusively in concert with other sites (correlated motion). In that case, the likely outcome is that several eigenvectors will contain simultaneous contributions from this site and others, although the amplitude of the contribution will vary. Presumably, small floppy RNA

molecules will exhibit both uncorrelated and correlated motions.

The RED formalism is best applied to molecular systems in which the correlation function describing global motion is separable from that of local motion, analogous to the Lipari–Szabo formalism for analysis of NMR relaxation data. Recently, RED was extended such that this separation need not be assumed *a priori*, leading to the formalism termed isotropic RED, or iRED (Prompers and Brüschweiler, 2002). In addition to eliminating the need to assume separability, iRED offers the potential to study the nature of a molecule's global reorientation by MD, a property not accessible in the original formulation of RED, while still providing a description of NMR spin relaxation active motions. Small floppy RNA molecules, specifically the IRE RNA hairpin, are excellent candidates for iRED analysis, since MD trajectories can be run for tens of nanoseconds, providing information on motions on the ps to low ns timescale, and because it is likely that global and internal motion are not rigorously separable for this class of molecule.

To disentangle the dynamics of the small IRE molecule, we have used RED and iRED to analyze a 12 ns MD trajectory of the IRE, and then compared the RED order parameters predicted from the trajectory to those obtained from NMR. In addition, the correlated motions of bases and riboses are clearly revealed by the iRED protocol. This powerful formalism gives a unique and comprehensive description of the dynamic motion of this small floppy RNA.

Materials and methods

RNA samples

IRE molecules were synthesized in vitro using SP6 RNA polymerase, incorporating [^{13}C]-ATP, [^{13}C]-GTP, [$^{13}C_{1'}$]-ATP, [$^{13}C_{1'}$]-GTP or [$^{13}C_{1'}$]-CTP (Laing and Hall, 1996). The labeled nucleotides were prepared as described (Nikonowicz et al., 1992). Unfortunately, our yield of [$^{13}C_{1'}$]-UTP was insufficient for synthesis. RNAs were purified by HPLC on a Dionex column and desalted. RNA samples were prepared in 30 mM NaCl, 10 mM sodium phosphate, pH 6, at RNA concentrations from 0.8 to 1.3 mM, and lyophilized three times for exchange into D_2O. NMR experiments used a Nalorac microprobe on the 500 and 600 MHz spectrometers; sample volumes were typically 160 μl.

NMR relaxation

Data were collected at 20 °C, where the structure was solved, and at 37 °C, the physiological temperature that is approximately 20° below the melting temperature in this buffer. ^{13}C NMR relaxation measurements of the IRE RNA required five samples, each one containing a single selectively labeled RNA: $^{13}C_{1'}$-Adenosine, $^{13}C_{1'}$-Guanosine, ^{13}C-1′–Cytosine, u-^{13}C-adenosine and u-^{13}C-Guanosine RNAs. The chemical shift overlap of the $^{13}C_{1'}$ ribose carbons and protons and the purine C_8 carbons precluded the use of a single sample. Those samples where the purine nucleotides were uniformly ^{13}C-labeled were used only for base relaxation experiments, since their $^1H_8$-$^{13}C_8$ and $^1H_2$-$^{13}C_2$ systems are isolated spin pairs, but their ribose carbons are not.

NMR T_1, $T_{1\rho}$ and NOE experiments were adapted from those of Yamazaki et al. (1994) for use with RNA (Hall and Tang, 1998). T_1 delays were 5, 35, 85, 125, 165, 205, 245, 285, 325, 365, 405, 445, 485, 525, 565, and 605 ms; $T_{1\rho}$ delays were 4, 8, 12, 16...to 72 ms. Relaxation dispersion experiments used fields from 1 to 6 kHz. One result of the use of separate samples for the $T_{1\rho}$ experiments was that the carrier could be placed effectively on-resonance to the ^{13}C frequencies. $T_{1\rho}$ spin-lock fields centered on C_2 (~155 ppm) carbons might also contain aromatic C_4 (~148 ppm) or C_6 carbons (~158 ppm), while fields centered on the C_8 (~140 ppm) carbons could contain the C_4 carbons. Although these aromatic carbons are passively coupled to the carbons of interest, they could constitute a relaxation pathway, and so introduce an error into the determination of T_2. However, this dipole:dipole relaxation will be a fraction of the dominant dipole:dipole relaxation through the attached protons of C_2 and C_8 (e.g. the distance between C_8 and its attached proton is 1.1 Å, while the distance between C_8 and its nearest carbon (C_4) is 2.19 Å; given the $1/r^6$ dependence of dipole:dipole interactions, there is a 62-fold difference in efficiency between the two pathways, far below our experimental uncertainty). Thus we consider the $^{13}C_2$-^{1}H and $^{13}C_8$-^{1}H spin

systems as effectively isolated for these measurements.

Relaxation data were acquired at 500 MHz, and fit using the Lipari–Szabo formalism (Hall and Tang, 1998). In our previous paper (Hall and Tang, 1998), a chemical shift anisotropy (CSA) of $\Delta\sigma = 189$ ppm was used for all purine carbons; now, based on density functional calculations, $\Delta\sigma = 120$ ppm for C_8 of adenosine and guanosine (Fiala et al., 2000) and for the adenosine C_2, $\Delta\sigma = 143$ ppm (slightly lower than the value of 152 ppm reported in Fiala et al., 2004). The CSA value for ribose $C_{1'}$ nuclei was 25 ppm (lower than the value of $\Delta\sigma \approx 40$ ppm (Fiala et al., 2000)).

To select the order parameter for each site, the data were fit with the selected τ_c and examined with respect to the residuals for each motional model. Errors in the calculated order parameters from experimental data come from at least two assumptions: we have not corrected for the angle between the C–H bond vector and σ_{33} of the CSA tensor (Fiala et al., 2000); and, we have assumed a constant value of CSA for all purine carbons. An estimate of the error from the first approximation can be made for the C_8 carbons (where the effects are likely to be the greatest), where here CSA = 120 ppm, $r_{CH} = 1.09$ Å, and the angle between the C–H bond and σ_{33} is 20 degrees (Sitkoff and Case, 1998; Fiala et al., 2000). We calculate that assuming an angle of zero degrees introduces an error of $< 1\%$ in R_1 and R_2, and an error in S^2 of 0.01; these errors are less than those from the experiment, and so we consider them to be negligible. By comparison, changes as small as 0.02 Å in the value of the effective bond length typically cause about a 10% change in R_1 and R_2.

Use of a constant value of the CSA for the purine bases ignores any site-specific variation that arises through local geometry and motions, particularly those that contribute to $J(0)$. In the absence of a rigorous experimental investigation of base CSA tensor variation as a function of position within RNA, the error introduced by neglect of site-specific variation cannot be estimated with accuracy or precision. To our knowledge, investigation into the variability of these tensors by solid state or solution NMR methods has not been reported (see Stueber and Grant (2002) for comments on the sensitivity of the tensor to electrostatic crystal potentials). Although it is difficult to extrapolate from ^{15}N to ^{13}C CSA tensor properties with certainty, investigation of site variation in amide ^{15}N CSA tensors for the protein RNase H led to the conclusion that as relaxation experiments are now measured, the precision of the data does not warrant correction of CSA for variation at 11.7 T (Kroenke et al., 1999). Given that CSA is only a minor contributor to the overall ^{13}C spin relaxation rate at 11.7 T for the nuclei investigated in this study, even large effects due to variation are likely to produce negligibly minor effects in the experimentally determined relaxation rates and derived order parameters for the IRE as well.

To calculate the overall correlation time of the RNAs, the new relaxation data for the riboses were included in global calculations of τ_c. As previously, an array of τ_c values was used as input for a minimization in parameter space using the Lipari–Szabo formalism. For each value of τ_c, T_1, $T_{1\rho}$, and NOE data from each residue were fit to each model of motion (S^2 only; S^2 and τ_e; S^2 and R_{ex}; S^2, τ_e, and R_{ex}; or S_f^2, S_s^2, and τ_f for each site (Mandel et al. 1995). The residuals of the target function χ^2 were calculated for each residue using each model,

$$\chi^2 = (T_1^{exp} - T_1^c)^2/\sigma_1^2 + (T_2^{exp} - T_2^c)^2/\sigma_2^2 + (NOE^{exp} - NOE^c)^2/\sigma_3^2$$

where *exp* is experimental value, *c* is calculated, and the standard deviations σ_i come from experimental data. The best correlation time for the molecule was that value that produced the smallest sum of all residuals.To select the order parameter for each site, the data were fit with the selected τ_c and examined with respect to the residual, and the values calculated for τ_e, R_{ex}, S_f^2, S_s^2, and τ_f. For ribose $^{13}C_{1'}$ sites, the simplest model was always the best; for the bases, other models were more appropriate (Hall and Tang, 1998). However, because application of the Lipari–Szabo formalism for this RNA molecule was shown to be only qualitatively correct, only the experimental order parameters are reported here.

Molecular dynamics simulations

Explicit solvent molecular dynamics (MD) simulations were performed using the AMBER7 software (Pearlman et al., 1995; Case et al., 2002) with

the AMBER94 force field (Cornell et al., 1995; Ponder and Case, 2003). The IRE starting conformation was taken from the solved NMR structures (Laing and Hall, 1996) in a manner consistent with earlier implicit solvent simulations by Hall and Williams (2004).

The IRE was initially solvated with 9050 TIP3P water molecules (Jorgensen et al., 1983) in a cubic box with sides of 67 Å length. This large simulation box served two purposes: to minimize interactions between periodic images of the RNA and to allow for reasonable numbers of monovalent and divalent counterions. Na^+, Cl^-, and Mg^{++} counterions were added to the simulation with the ISIM grand canonical Monte Carlo (GCMC) software (Vitalis et al., 2004) using excess chemical potentials calibrated to give bulk (i.e., in the absence of RNA) concentrations of 50 mM NaCl and 1 mM $MgCl_2$. The resulting GCMC simulations placed 45 Na^+, 32 Cl^-, and 1 Mg^{++} ions in the (67 Å) (Butcher et al., 1997) domain to give an electroneutral system with effective local concentrations of 250 mM Na^+, 180 mM Cl^-, and 5 mM Mg^{++}.

The initial RNA structure was minimized via steepest descent minimization to relax the initial structure. The water was then heated to 50 K through 10 ps of canonical (NVT) MD with 2 fs time steps and SHAKE constraints (Ryckaert et al., 1977) on hydrogen-heavy atom bonds. Next, the entire system was slowly heated through a series of isothermal-isobaric (NPT) simulations at 50, 100, 150, 200, 250, and 300 K. Each simulation was performed for 10 ps with 2 fs time steps, SHAKE constraints on hydrogen-heavy atom bonds, and particle mesh Ewald (PME) electrostatics (Darden, 2001) with a 10 Å direct space cutoff. Upon completion of this heating process, an additional 11.93 ns of NPT MD (Berendsen et al. 1984) was performed to generate a total trajectory of 12 ns duration.

Order parameters from RED analysis

The reorientational eigenmode dynamics formalism and its isotropic extension have been described in detail elsewhere (Prompers and Brüschweiler, 2001, 2002). All calculations were performed using in house programs written in either perl or C; our protocol for implementation of RED has been published elsewhere and required only superficial modification to adapt to RNA (Showalter and Hall, 2002). The polar coordinates representing 36 spin interactions were calculated every 2 ps along the final 10 ns of the 12 ns trajectory. These sites were the 16 backbone C1′–H1′, the purine C8-H8, Adenine C2-H2, and the pyrimidine C5-H5. The saved coordinates were used to generate the 36×36 reorientational covariance matrix M, which has elements (Prompers and Brüschweiler, 2001):

$$M_{ij} = \sum_{M=-2}^{2} \overline{|\Delta Y_{2M}(\Omega_i) >< \Delta Y_{2M}(\Omega_j)|} \qquad (1)$$

where $\Delta Y_{2M} = Y_{2M}(\Omega) - \overline{Y_{2M}(\Omega)}$, Y_{2M} are the normalized spherical harmonics of rank 2, and the bar indicates ensemble averaging over the coordinates from the simulation. One important property of the matrix M in this representation is that its diagonal elements are proportional to the generalized order parameter S^2 of Lipari and Szabo (Prompers and Brüschweiler, 2001; Brüschweiler and Wright, 1994):

$$1 - S_i^2 = \frac{4\pi}{5} \sum_{M=-2}^{2} \sigma^2_{Y_{2M,i}} = \frac{4\pi}{5} M_{ii} \qquad (2)$$

where σ^2 is the variance of spherical harmonic Y_{2M}.

iRED analysis

All calculations were performed using in house programs written in perl or C and the details of our protocol for applying iRED to RNA trajectories have been published (Showalter and Hall, 2005). Briefly, over the course of an MD trajectory the principle axis directions $\Omega_j(t) = (\theta_j(t),\phi_j(t), j = 1, \ldots, n)$ of the spin interaction tensors for n sites are stored and used to compute a covariance matrix of the $Y_{2M}(\theta,\phi)$. After isotropic averaging, an $n \times n$ real symmetric matrix is produced with elements of the form (Prompers and Brüschweiler, 2002):

$$M_{ij} = \overline{P_2(\cos(\Omega_i - \Omega_j))} \qquad (3)$$

where $P_2(x) = (3x^2 - 1)/2$ is the second Legendre Polynomial, $\Omega_i - \Omega_j$ is the angle between spin interactions i and j in a given snapshot, and the bar indicates averaging over all snapshots of the trajectory. For this work, the calculated covariance matrix was 36×36 and included the same

sites as described for RED analysis. The covariance matrix can be diagonalized by solving the eigenvalue problem $M|\mathrm{m}>=\lambda_{\mathrm{m}}|\mathrm{m}>(\mathrm{m}=1, ..., n$; $n = 36$ for the IRE system). The resulting normalized reorientational eigenvectors $|\mathrm{m}>$ contain information about which spin interactions reorient in concert under the influence of each motional mode, and the eigenvalues λ_{m} represent the amplitude of the observed motion.While the eigenvectors of M contain detailed information about the dynamic correlations between sites, they contain no information about the timescale on which the observed motion occurs. This information can be extracted by constructing correlation functions describing the decay of motion along each eigenvector over the course of the trajectory:

$$C_{\mathrm{m}}(t) = \sum_{l=-2}^{2} < a_{\mathrm{m,l}}{}^{*}(\tau + t)a_{\mathrm{m,l}}(\tau) >_{\tau} \qquad (4)$$

where averaging is done over the snapshots of the simulation and the $a_{\mathrm{m,l}}(t)$ are constructed from the instantaneous projection of the eigenvectors onto the snapshots (Prompers and Brüschweiler, 2002). Assuming that these correlation functions decay mono-exponentially, a lifetime τ_{m} associated each motional mode can be established (Lipari and Szabo, 1982a,b).

Eigenvector collectivity

The collectivity of an eigenvector is defined by the parameter κ, which is roughly proportional to the percent of sites significantly reoriented by the motion represented by that vector (Brüschweiler, 1995):

$$\kappa_m = \frac{1}{N}\exp\left\{-\sum_{n=1}^{N} \left|\left|\mathrm{m}>_n\right|^2 \log\left|\left|\mathrm{m}>_n\right|^2\right.\right\} \qquad (5)$$

where $|\mathrm{m}>_n$ is the nth component of eigenvector $|\mathrm{m}>$ and N is the number of spin interactions. κ_{m} ranges from $1/N$ to 1.

Results and discussion

IRE dynamics from NMR relaxation

^{13}C NMR relaxation experiments to investigate IRE base dynamics were previously described (Hall and Tang, 1998). In those experiments, relaxation parameters were obtained from data collected at 20 °C, where the structure was solved, and at 37 °C (where the structure could not be solved, due to weak or nonexistent NOEs). Analysis was complicated by having few NMR relaxation-active probes, since only the ^{13}C-purine bases were used in those experiments.

Subsequent experiments to probe ribose dynamics have now been done, using cytidine, adenine and guanine specifically labeled at $^{13}C_{1'}$. Because spectral overlap in both the carbon and proton dimensions was severe, separate RNA samples had to be synthesized that contained only one labeled nucleotide species. The $^{1}H_{1'}$ resonances were assigned from our previous NMR structure determination (Laing and Hall, 1996). Each sample was used in standard T_1, $T_{1\rho}$, and heteronuclear NOE relaxation experiments, as described (Hall and Tang, 1998). These data were then combined with the previous IRE relaxation data for purine bases, and all data were processed together to select an overall correlation time (τ_{c}); the recalculated values are $\tau_{\mathrm{c}} = 3.6$ ns at 20 °C, and $\tau_{\mathrm{c}} = 2.8$ ns at 37 °C (the previous study produced $\tau_{\mathrm{c(20\ C)}} = 3.65$ ns and $\tau_{\mathrm{c(37\ C)}} = 2.45$ ns). Providing more input data from the riboses (presumably less flexible than the bases) for the calculation of the global tumbling time prior to selection of the appropriate model of motion for each site should improve the accuracy of the results.

Estimating the τ_{c} of the IRE using data from the purine bases alone was likely to be biased by any unusual motions of those bases in the loop. By using the riboses and bases in the stems for fitting of the overall correlation time, the input data should be more reflective of a stable core structure, analogous to the backbone amides of a protein core. In fact, the effect was to more precisely define τ_{c} from fitting with the $^{13}C_{1'}$ ribose sites from the stem. Unfortunately, inclusion of the ribose data resulted in an increase in the selected order parameters of all bases compared to our previous analysis (Hall and Tang, 1998), suggesting that although τ_{c} may be better determined, the dynamics of many of the purine bases are not well described by the formalism. The order parameters selected for each ^{13}C labeled position are shown in Figure 2b for the bases and Figure 2a for the riboses, at 20 and 37 °C. In general, the results show: (1) lower order parameters for the riboses than for their bases; (2) uniform (high) order

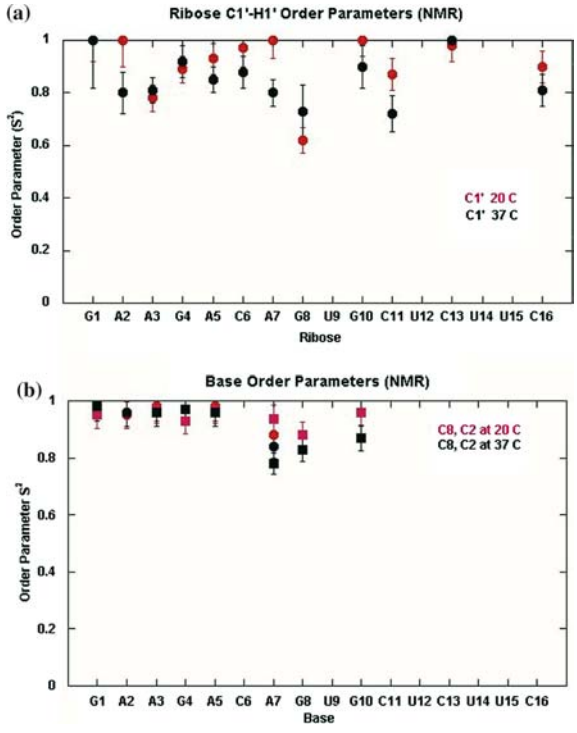

Figure 2. ^{13}C NMR order parameters for the IRE at 20 and 37 °C. (a) S^2 for guanosine, adenosine, and cytosine ribose $C_{1'}$. (b) S^2 for purine bases, guanine C_8 and adenine C_2, C_8. Errors were calculated from Monte Carlo simulation.

parameters for the bases in the stem; (3) lower order parameters for the bases A7, G8, and G10 in the loop, a trend that is more pronounced at high temperature; and (4) ribose order parameters appear independent of those of their base, the most striking example of which is the G8 nucleotide, where at 20 °C, the G8 $^{13}C_8$ has an order parameter of 0.85, but its $^{13}C_{1'}$ has $S^2 = 0.6$.

The values of the order parameters for each spin pair should reflect the amplitude of motion. Completely unrestricted motion of the internuclear vector would have $S^2 = 0.0$, while a vector held rigid in the molecular frame would have $S^2 = 1.0$. Since the base ^{13}C–^{1}H spin pairs used in the NMR experiments are fixed by sp^2 geometry in the plane of the aromatic purine base, the motion reported is that of the base itself. Assuming that the A-form RNA duplex structure does not tolerate sugar repuckering and that the bases are tightly constrained, both riboses and bases from the stem should have high order parameters. Since the loop is floppy, and several of its sugars undergo repuckering, the order parameters of its bases and riboses should be measurably lower than those of the stem. The recalculated order parameters (S^2) for purines are often unreasonably high, but there is clearly a difference between the bases in the stem and those in the loop (noting that G10 has higher order parameters at 20 °C, consistent with its base pairing to C6). Although several ribose $^{13}C_{1'}$ positions are not adequately fit by any model, as indicated by the value of $S^2 = 1.0$ (an allowed, but physically unrealistic value), there does appear to be a recognizable difference in the order parameters for those riboses in the loop and stem.

In addition to motions of the nucleotides on the ps-ns timescale reported by the Lipari–Szabo order parameters, some nucleotides also undergo motions on longer timescales. As noted previously (Hall and Tang, 1998), the three terminal base pairs of the stem fray, leading to the observation of two and sometimes three resonances for a single nucleus (slow exchange regime). The upper limit of the exchange, which results from local duplex melting, can be simply estimated from the chemical shift difference ($\Delta\delta$) between the resonances. The two $^{13}C_2$ and $^{13}C_{1'}$ resonances from A2 and the two $^{13}C_{1'}$ resonances from G1 have an upper limit of exchange of 24 ± 1 ms. Adenosine 3 also shows two resonances of its $^{13}C_2$ peak, but that exchange is slower (72 ms), probably because melting of the duplex at this position occurs more seldom. This slow conformational exchange will not be reported by the order parameters derived from NMR relaxation experiments, since the latter are dominated by much faster motions occurring on the ps-ns timescale. Only single resonances were observed from bases and riboses in the loop. The one exception is the ribose of G8, which has two $^{13}C_{1'}$ resonances at 37 °C; based on the chemical shift difference between the resonances, the conformational exchange time is 23 ms. To further explore the timescales of motion, relaxation dispersion experiments (on-resonance $R_{1\rho}$ experiments, where the spin lock power was varied) were done at 20, 30, and 37 °C at 500 and 600 MHz for the purine bases and the $^{13}C_{1'}$-labeled riboses (data not shown). No resonance showed a significant power dependence of $R_{1\rho}$, indicating no exchange on the ms timescale.

While these trends observed for dynamics of nucleotides in the IRE are plausible, there is a potential problem with the extraction of order parameters from these data. The Lipari–Szabo

formalism requires that the internal motion of the ^{13}C-^{1}H vector be separable from overall tumbling (Lipari and Szabo, 1982a, b). Both of the bases in the IRE loop which have been studied by fluorescence spectroscopy have demonstrated motions on the ns timescale, comparable to the global tumbling time of the RNA (Hall and Williams, 2004). This suggests a high probability of a breakdown in the separability of global and internal motion, which would render the NMR derived order parameters inaccurate.

RED analysis – order parameters

The experimentally determined order parameters provide minimal data about some regions of the IRE, specifically the 3′ stem, owing to the lack of a $^{13}C_{1'}$ Uridine labeled sample, and a complete lack of isolated spin pairs in the pyrimidine bases. To compensate for the incomplete description of IRE dynamics offered by the NMR experiments, a 12 ns molecular dynamics simulation of the molecule was run and analyzed using the RED formalism. Snapshots collected every 2 ps over the final 10 ns of the trajectory were superposed and used to construct the RED covariance matrix (methods, Equation 1), for a total of 36 spin pairs: the 16 backbone $C_{1'}$–$H_{1'}$, 8 purine C_8–H_8, 4 Adenine C_2–H_2, and 8 pyrimidine C_5–H_5. Order parameters for each of these sites were computed using Equation 2 and are shown in Figure 3.

Figure 3a shows the computed ribose $C_{1'}$ order parameters. The trends suggest that the general features of the IRE have been preserved by the simulation, notably slight end fraying (more dominant at the 3′ end), lower order parameters in the loop, and a stabilization of C6 and G10 by their base pairing interaction. The average order parameter of 0.86 ± 0.06 for the ten stem riboses, inclusive of the fraying terminal base pair, is consistent with a relatively rigid A-form structure.

Figure 3b shows the computed order parameters for each of the bases. Notably, the order parameters of 0.86 for C6 and 0.89 for G10 fall within the expected range for a stable, well ordered, base pairing interaction (compare to $S^2 = 0.88 \pm 0.02$ for the bases in the stem). Although the order parameters for A7 are lower than those of bases in base pairing interactions, they are much higher than those for G8, U9, and C11, consistent with preservation of the stacking interaction between C6 and A7. The notably lower order parameters for G8, U9, and C11, along with their reduced ribose order parameters, demonstrate that they truly are more mobile than the rest of the molecule and highlight the flexible nature of the loop structure.

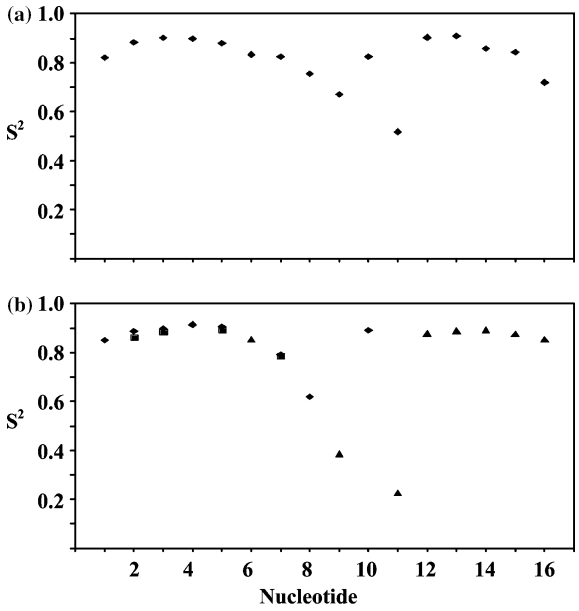

Figure 3. The square of the generalized order parameter, S^2, calculated from the diagonal elements of the non-diagonalized RED matrix M (methods, Equation 2). (a) Ribose $C_{1'}$ (b) ◆ Purine C_8, ■ adenine C_2, ▲ pyrimidine C_5.

iRED analysis

Analysis of experimental relaxation data predicts a global tumbling time (τ_c) on the order of 2.8–3.7 ns depending on temperature. Fluorescence anisotropy measurements of 2-aminopurine substituted IRE suggest internal (local) dynamic motions on the nanosecond timescale as well (Hall and Williams, 2004). When the timescale of local motions matches that of global tumbling, the Lipari–Szabo formalism is only conditionally valid and interpretation of order parameters may lead to significant errors if the separability assumption cannot be independently justified (Lipari and Szabo, 1982a,b). The isotropic Reorientational Eigenmode Dynamics (iRED) formalism (Prompers and Brüschweiler, 2002) was applied to the last 10 ns of the IRE trajectory in order to test the validity of the separability assumption. Even

though the decay times of the five largest amplitude modes are expected to be shorter than the experimental global tumbling time, the eigenvalues of these global tumbling modes should be distinct from those of internal modes if separability is satisfied. In addition, this formalism provides detailed physical information about the motions experienced by each RNA site, and demonstrates the extent of any dynamic correlation between these sites.

The real symmetric iRED covariance matrix M (methods, Equation 3) was constructed for the 36 spin interactions of interest using snapshots collected every 2 ps over the final 10 ns of the trajectory. The matrix was diagonalized, yielding 36 normalized reorientational eigenvectors $|m>$ (referred to here as e1, e2, etc. for simplicity), containing information about which spin interactions reorient in concert under the influence of each motional mode, and the eigenvalues λ_m, which represent the amplitude of the observed motion (see Table 1 for a key relating position in the eigenvectors to spin interactions in the RNA).

Testing the separability assumption

The eigenvalue (λ_m) associated with each eigenvector $|m>$ of the iRED covariance matrix M is the variance of the amplitude of motion along the reorientational mode (Prompers and Brüschweiler, 2002). The collectivity (κ_m, Equation 5) of a given eigenvector reflects the percent of spin interactions significantly reorienting under the influence of the mode represented by the eigenvector (Brüschweiler, 1995), and so indicates the degree of correlation between all sites in the molecule imposed by the eigenmode. For an internally rigid molecule, the matrix M will have at most five non-zero eigenvalues (when the rank 2 spherical harmonics are used to construct M) and the collectivities of their associated eigenvectors should be very high ($\kappa_m > 0.50$). Additional non-zero eigenvalues reflect internal motion of the system; separability is characterized by a large gap in both eigenvalue and collectivity between the five largest amplitude (global) modes and the $n - 5$ internal modes.

A plot of κ_m vs. λ_m (Figure 4) calculated from the IRE trajectory shows a continuous distribution with no apparent gap separating internal and global modes in either the collectivity or amplitude dimension. Given that the trajectory was calculated to four times the experimental global tumbling time, this suggests the separability assumption is invalid for the IRE construct and brings the absolute value of the experimentally determined order parameters into question. Order parameters could be calculated from iRED, but when separability is shown to be unsatisfied, it is

Table 1. iRED eigenvector spin interaction key

Spin interaction	Nucleotide	Heavy atom	Spin interaction	Nucleotide	Heavy atom
1	G1	C1′	19	G8	C1′
2	G1	C8	20	G8	C8
3	A2	C1′	21	U9	C1′
4	A2	C8	22	U9	C5
5	A2	C2	23	G10	C1′
6	A3	C1′	24	G10	C8
7	A3	C8	25	C11	C1′
8	A3	C2	26	C11	C5
9	G4	C1′	27	U12	C1′
10	G4	C8	28	U12	C5
11	A5	C1′	29	C13	C1′
12	A5	C8	30	C13	C5
13	A5	C2	31	U14	C1′
14	C6	C1′	32	U14	C5
15	C6	C5	33	U15	C1′
16	A7	C1′	34	U15	C5
17	A7	C8	35	C16	C1′
18	A7	C2	36	C16	C5

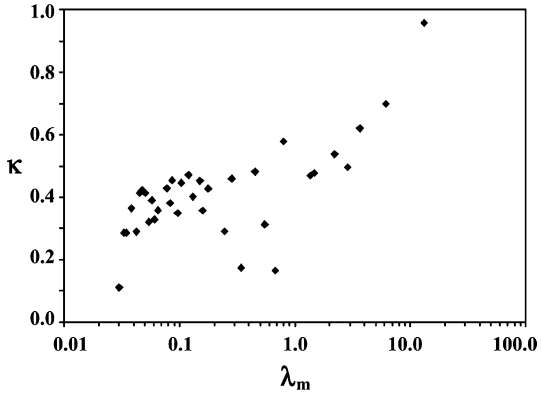

Figure 4. The distribution of the collectivity parameter κ vs. eigenvalue λ_m from iRED analysis. The absence of a gap separating the five points with largest λ_m from the rest of the distribution suggests that global and internal motion are not separable for the IRE.

unclear which eigenvectors to include in the calculation (Prompers and Brüschweiler, 2002). Although separability is also strictly necessary for RED calculations, those calculations more closely resemble the experimentally derived values in their assumptions, thus they were used for comparison.

Correlation times of the eigenmodes

While the eigenmodes of the iRED covariance matrix provide information on the amplitude of motions experienced by each investigated site in the IRE, all time information is lost during the construction of the matrix, leaving the timescale of the motions sampled undetermined. Correlation functions describing the loss of memory along each eigenmode were calculated using equation 4 in order to reconstruct the timescale information contained within the trajectory. Assuming that these correlation functions decay mono-exponentially, correlation times can be extracted that represent the timescale of motion along each mode. Of the 36 correlation functions calculated, 32 decayed to zero mono-exponentially; e1, e4, e5 and e6 have correlation functions that do not reach a stable plateau or that have a plateau at a value well above zero. These four correlation functions can be found in the supplementary material.

In theory, the five eigenvectors with the largest eigenvalues should have τ_m on the order of 2–3 ns, approximately equal to the global tumbling time from Lipari–Szabo analysis. In practice, the longest lifetimes sampled in the MD trajectory fell well short of the experimental tumbling time (Figure 5), with the largest calculated $\tau_m = 790$ ps. Interestingly, 13 eigenmodes, 36% of the total, have lifetimes within an order of magnitude of the experimentally determined global tumbling time. Although the longest tumbling times must be underestimated by the simulation, as the five longest should correspond to the experimental value, this result is consistent with the time resolved fluorescence data which demonstrates that a significant portion of the internal motion occurs on the same timescale as global tumbling (Hall and Williams, 2004).

Correlated motion

The eigenvectors of the iRED matrix *M* predict the pattern and extent of the correlation networks that couple the various sites in the IRE. Of course, eigenvectors could also be calculated from RED analysis, but iRED is preferred here because, given the intermingling between global and internal dynamics in the IRE, the superposition procedure required for RED would remove some potentially important information. Among the thirty six eigenvectors, two (e1, e2) contain large amplitude contributions from all sites in the RNA; e3 and e4 contain large amplitude contributions from the loop plus the 5′ or 3′ stem, respectively. Smaller amplitude contributions come from all sites in e5 as well, consistent with the expectation that the five eigenvectors with the largest eigenvalues will represent global motion. Of the remaining eigenvectors, many are no less "global" in their distribution

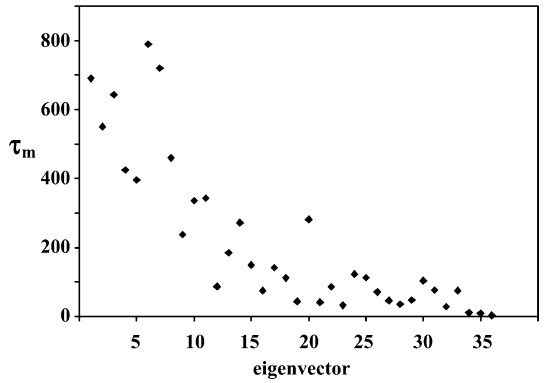

Figure 5. Characteristic lifetime in picoseconds (τ_m) associated with each eigenvector $|m>$ of the matrix *M* generated from iRED analysis of the IRE simulation. τ_m is the lifetime associated with the reorientational mode represented by $|m>$ and therefore is not a direct estimate of τ_e from Lipari–Szabo analysis.

of affected sites: e7, e19 and e25 contain contributions sampling all regions of the RNA; e8 contains contributions from most ribose sites; and a total of nineteen eigenvectors contain contributions from most or all of the sites in the loop, with some correlation to one or both stems in seven of these nineteen. The remaining eigenvectors contain contributions from the 5′ stem, 3′ stem, or both and are generally amongst the lowest in eigenvalue. Several eigenvectors are especially noteworthy for what they reveal about the motions in the IRE.

The bases G8, U9, and C11 exhibited low order parameters in RED analysis of the trajectory; G8 was shown by NMR and fluorescence experiments to have considerable motion. These three bases might be expected to undergo local motion, which would result in a set of three iRED eigenvectors, each dominated by amplitude from one of these bases individually. Indeed, e9, e10, and e12 correspond to such eigenvectors; there is little significant correlation to other sites in the molecule (Figure 6a,b,c). However, their motions are not exclusively local: e6 shows strong correlation between these bases and other loop sites, as well as the loop-closing base pair (Figure 6d). Eigenvector 6 also has the distinction of having the longest correlation time ($\tau_m = 790$ ps) calculated from its correlation function. More generally, the extensive motions of G8, U9 and C11 are reflected in the number of eigenvectors of the iRED matrix that contain projections onto these sites.

Guanosine 8

An example of the utility of iRED is seen in the description of Guanosine 8, one of the mobile nucleotides in the IRE loop. Experimentally determined NMR order parameters for this base and its ribose are significantly lower than the flanking A7 nucleotide. The NMR relaxation data are in agreement with time-resolved fluorescence data that show that this base (2-aminopurine substituted for G8) experiences dynamic motion on a ns timescale (Hall and Williams, 2004). What the iRED analysis shows is that the motions that contribute to the experimentally measured dynamics come from several highly correlated modes, not simply from independent motion of the base and/or ribose. In fact, amplitude from the G8 base is found not only in e10, where it is uniquely present, but also in e1 and e4 (global eigenvectors), e6, e11, e13, and e14. Amplitude from the G8 ribose appears in e18, which it shares only with U9 $C_{1'}$, and also in e1 and e2 (global) and e7, e11, e15, and e20. Curiously, it does not appear in e8, which has contributions from most of the riboses in the IRE molecule. The correlated motions revealed by the iRED analysis will be inextricably part of the order parameter calculated from NMR relaxation data, rendering that value a complicated function of global, local, and correlated motions.

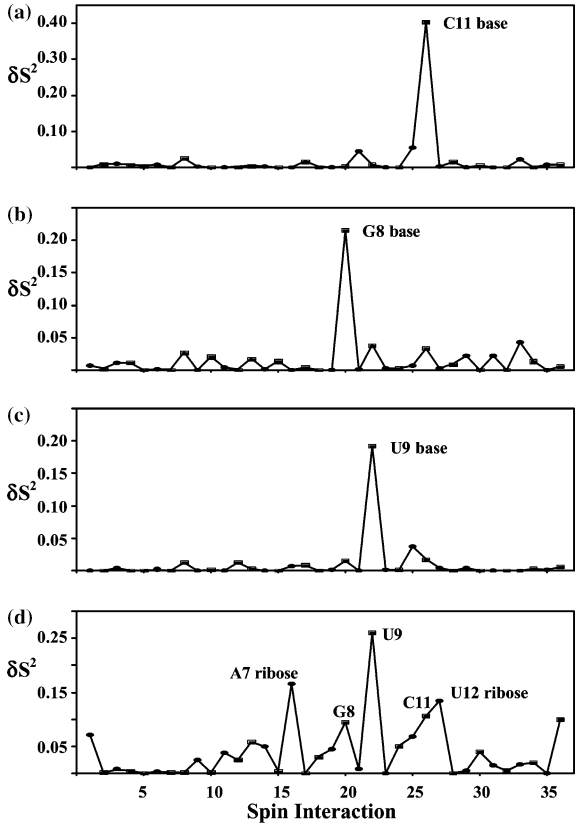

Figure 6. Reorientational eigenvectors from iRED analysis of the IRE trajectory shown as the principal order parameter components $\delta S_j^2 = \lambda_m |m>_j|^2$ for each spin pair. (a) e9 displays predominantly independent C11 base motion, (b) e10 displays predominantly independent G8 base motion, and (c) e12 displays predominantly independent U9 base motion. (d) e6 shows correlations spanning the loop and including the bases of G8, U9, and C11, even though these bases are unrestrained in the NMR structure. (●) Ribose positions; (■) Base positions; the full key relating interaction number to site in the IRE can be found in Table 1.

Uridine 9 and Cytosine 11

There are no NMR relaxation data for the pyrimidines in the IRE, but the simulations describe their motions. Given the good agreement between

in vitro and *in silico* descriptions of the purine nucleotides, the RED and iRED data should provide a reasonable picture of the properties of the pyrimidines. The two pyrimidine nucleotides in the loop, U9 and C11, have low order parameters for both base (0.4 and 0.22, respectively) and ribose (0.65 and 0.5, respectively) from RED analysis. NMR relaxation data also produce a low order parameter of C11 $C_{1'}$, especially at 37 °C. NMR structural data were unable to provide constraints for these base positions in the loop. The iRED calculation shows the extent to which motions of U9 and C11 nucleotides are correlated to other positions in the IRE. iRED data show that the U9 base appears in e1 (global), predominates in e6, and is alone in e12, making it relatively isolated. In contrast, its ribose has amplitude in e1, e2, e4, and e5 (all global motions), as well as e8, e11 (predominant contributions from A5 and C11 bases, followed by the amplitude of the U9 $C_{1'}$, with smaller amplitude from A7 and G8 bases, and G8 ribose), e14 (where U9 $C_{1'}$ is correlated only with C11 $C_{1'}$), e15 (riboses of A7, G8, and U9, and the A7 base), and e16 (riboses of G8 and U9). This picture of the ribose shows its strong connection to other sugars in the loop, with some base motions also contributing.

The C11 base, extruded from the loop, has amplitude in e1 (global), e6, is alone in e9, and has the major contribution to e11. Its ribose appears in e1, e2, and e4 (global eigenvectors), as well as in e6, e7, e8, e13, e25 (the dominant amplitude). In general, these data show that loop nucleotides appear to be moving in concert, in addition to their idiosyncratic dynamics. There is a strong connection among the riboses throughout the loop, independent of their base motions. The consequence of these correlations is that any single site order parameter determined by NMR experiments (and RED analysis) will be influenced by dynamics of many other sites in the RNA.

The C6:G10 pair

Two bases in the loop, C6 and G10, are shown to be hydrogen bonded in the NMR structure. If this pairing is stable on the timescale of the simulation, then little motion should be observed for these two bases. Indeed, this is indicated by iRED analysis, where e23 and e28 (Figure 7a,b) are the eigenvectors with the largest eigenvalues displaying significant motion from these two bases (excepting the truly global modes, which show correlation from all sites). These eigenvectors represent specific correlated motion of these two bases and A7, which stacks with C6. Again, however, their motion is not isolated from other sites in the IRE.

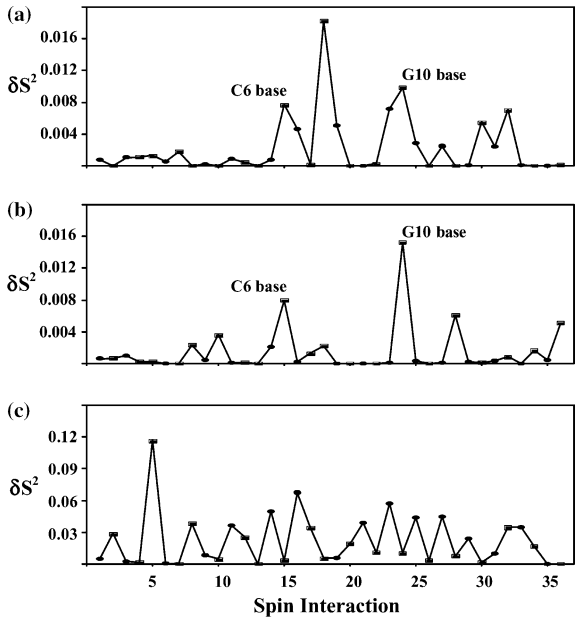

Figure 7. Reorientational eigenvectors from iRED analysis of the IRE trajectory shown as the principal order parameter components $\delta S_j^2 = \lambda_m |\langle m|_j|^2$ for each spin pair. (a) e23 and (b) e28 display correlated motion spanning the C6:G10 basepair. (c) e8 is dominated by ribose correlations along the majority of the sugar-phosphate backbone. (●) Ribose positions; (■) Base positions; the full key relating interaction number to site in the IRE can be found in Table 1.

Noteworthy eigenvectors with correlated sites

Eigenvector 8 is unique in that its contributions come predominantly from the ribose vectors (Figure 7c). While riboses in a mostly rigid stem might be expected to have some correlated motion, the extension of these correlations through the flexible loop is unexpected. The observation here is that the correlation starts with three bases at the 5′ end of the stem, then extends from the A5 ribose through C6 (stacked on A5) and A7 (some stacking on C6); breaks at G8, then continues with the riboses of U9, C11, U12, C13, and U15.

The eight eigenvectors with the lowest eigenvalues contain amplitude from the stem regions of the IRE, either the duplex or one strand. Motions of the stem should indeed occur seldom on the timescale of the simulation, for the A-form struc-

ture is both rigid and stable. Motion of the 3′ terminal ribose is indicated here by its slightly lower order parameter, but fraying of the duplex terminus is slow (ms timescale), and so that motion will not be sampled by the simulation.

Conclusions

One implication of the extensive correlated motions we observe in the IRE is the degree to which base motion can affect an RNA structure. The large scale motions of the bases in the IRE loop, as they appear during the course of the simulation, illustrate how the structure of the entire RNA is altered in concert (see Figure 8). The large amplitude of the modes dominated by the loop bases suggest that the motion they represent may be deforming the overall structure, a situation that would couple internal motion and the global tumbling of the IRE.

Standard practice for experimental analysis of RNA secondary and tertiary structure is to mutate a given site, then probe the RNA for structural changes. If sites are coupled through their dynamics, then a base mutation could propagate not only a structural change throughout the molecule, but could alter its dynamic motions as well. If dynamic motion is necessary to prepare a site for interaction, then a change in the extent of flexibility could alter the available contact surface and so affect binding.

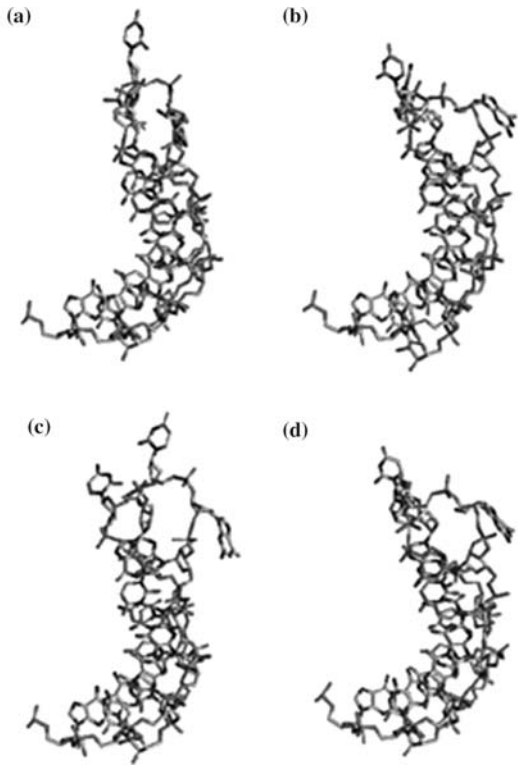

Figure 8. Snapshots from the trajectory showing the deformations of the IRE loop (a) 2 ns; (b) 5 ns; (c) 7 ns; (d) 9 ns.

The iRED analysis has provided a graphic confirmation of what our intuition predicted: that there is no clean separation between global tumbling of the IRE RNA molecule and the local motions of its nucleotides. In addition, iRED is able to provide a detailed description of which motions are correlated, and the extent to which they contribute to the overall system description (e.g. small eigenvalues, small contribution). These features of the IRE are likely to be shared by other RNA structures, such as the large hairpin loops in snRNAs, small hairpin loops like the TAR loop, GNRA loops, and tRNA anticodon loops.

This study is the first application of iRED formalism to the description of motions in a small and floppy RNA structure. It shows that global and local motions cannot be rigorously separable; thus order parameters calculated from NMR relaxation experiments will be unreliable. While RED analysis of the MD trajectory qualitatively reproduces the experimental NMR order parameters, iRED analysis shows that the differences may be due to inseparability of global and internal motion. Both RED and iRED results show that several modes contribute to the motion at each site. We propose that iRED analysis of MD trajectories provides an indispensable complement to experimental descriptions of dynamic motions in small floppy RNA molecules.

Supplementary material available

Plots of correlation functions for the eigenvectors illustrated in Figures 6 and 7, as well as the four that do not converge are available as electronic supplementary material at http://dx.doi.org/10.1007/s10858-005-7948-2.

Acknowledgements

SAS is the recipient of an NSF pre-doctoral fellowship, and funding from the Molecular

Biophysics Training Grant. This work is supported in part by the NIH (to KBH and NAB) and the Alfred P. Sloan Foundation (to NAB).

References

Addess, K.J., Basilion, J.P., Klausner, R.D., Rouault, T.A. and Pardi, A. (1997) *J. Mol. Biol.* **274**, 72–83.

Amadei, A., Linssen, A.B.M. and Berendsen, H.J.C. (1993) *Proteins* **17**, 412–425.

Barton, H.A., Eisenstein, R.S., Bomford, A. and Munro, H.N. (1990) *J. Biol. Chem.* **265**, 7000–7008.

Basilion, J.P., Rouault, T.A., Massinople, C.M., Klausner, R.D. and Burgess, W.H. (1994) *Proc. Natl. Acad. Sci. USA.* **91**, 574–578.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A. and Haak, J.R. (1984) *J. Phys. Chem.* **81**(8), 3684–3690.

Bettany, A.J., Eisenstein, R.S. and Munro, H.N. (1992) *J. Biol. Chem.* **267**, 16531–16537.

Brüschweiler, R. (1995) *J. Chem. Phys.* **102**, 3396–3403.

Brüschweiler, R. and Wright, P.E. (1994) *J. Am. Chem. Soc.* **116**, 8426–8427.

Butcher, S.E., Dieckmann, T. and Feigon, J. (1997) *EMBO J.* **16**, 7490–7499.

Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham III, T.E., Wang, J., Ross, W.S., Simmerling, C., Darden, T., Merz, K.M., Stanton, R.V., Cheng, A., Vincent, J.J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R., Duan, Y., Pitera, J., Massova, I., Siebel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A. (2002) *AMBER 7 User's Manual*. University of California.

Casey, J.L., Henze, M.W., Koeller, D.M., Caughman, S.W., Rouault, T.A., Klausner, R.D. and Harford, J.B. (1988) *Science* **240**, 924–928.

Cheng, X., Hornak, V. and Simmerling, C. (2004) *J. Phys. Chem. B* **108**, 426–437.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1995) *J. Am. Chem. Soc.* **117**, 5179–5197.

Darden, T.A. (2001) In *Computational Biochemistry and Biophysics Treatment of long-range forces and potential*, Becker, O.M., Mackerell, A.D., Roux, B. and Watanabe, M. (Eds.), Marcel Dekker, Inc., New York, pp. 91.

Dayie, K.T., Brodsky, A.S. and Williamson, J.R. (2002) *J. Mol. Biol.* **317**, 263–278.

Fiala, R., Czernek, J. and Sklenář, V. (2000) *J. Biomolec. NMR* **16**, 291–302.

Fiala, R., Munzarová, M.L. and Sklenář, V. (2004) *J. Biomol. NMR* **29**, 477–490.

Hall, K.B. and Tang, C. (1998) *Biochemistry* **37**, 9323–9332.

Hall, K.B. and Williams, D.J. (2004) *RNA* **10**, 34–47.

Heus, H.A. and Pardi, A. (1991) *Science* **253**, 191–194.

Hirling, H., Henderson, B.R. and Kuhn, L.C. (1994) *EMBO J.* **13**, 453–461.

Hoogstraten, C., Wank, J.R. and Pardi, A. (2000) *Biochemistry* **39**, 9951–9958.

Jaffrey, S.R., Haile, D.J., Klausner, R.D. and Harford, J.B. (1993) *Nucleic Acids Res.* **21**, 4627–4631.

Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D. (1983) *J. Chem. Phys.* **79**, 926–935.

Kaldy, P., Menotti, E., Moret, R. and Kuhn, L.C. (1999) *EMBO J.* **18**, 6073–6083.

Kikinis, Z., Eisenstein, R.S., Bettany, A.J.E. and Munro, H.N. (1995) *Nucleic Acids Res.* **23**, 4190–4195.

King, G.C., Harper, J.W. and Xi, Z. (1995) *Methods Enzymol.* **261**, 436–450.

Kitao, A., Hirata, F. and Go, N. (1991) *Chem. Phys.* **158**, 447–472.

Kraulis, P.J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.

Kroenke, C.D., Rance, M. and Palmer, III, A.G. (1999) *J. Am. Chem. Soc.* **121**, 10119–10125.

Laing, L.G. and Hall, K.B. (1996) *Biochemistry* **35**, 13586–13594.

Leibold, E.A., Laudano, A. and Yu, Y. (1990) *Nucleic Acids Res.* **18**, 1819–1824.

Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.* **104**, 4546–4559.

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.* **104**, 4559–4570.

Mandel, A.M., Akke, M. and Palmer, A.G. (1995) *J. Mol. Biol* **245**, 144–163.

Miller, J.L. and Kollman, P.A. (1997) *J. Mol. Biol.* **270**, 436–450.

Nikonowicz, E.P., Sirr, A., Legault, F.M., Jucker, P., Baer, L.M. and Pardi, A. (1992) *Nucleic Acids Res.* **20**, 4507–4526.

Pearlmann, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, III, T.E., DeBolt, S., Ferguson, D., Seibel, G. and Kollman, P. (1995) *Comput. Phys. Commun.* **91**, 1–41.

Ponder, J.W. and Case, D.A. (2003) *Adv. Protein Chem.* **66**, 27–85.

Prompers, J.J. and Brüschweiler, R. (2001) *J. Am. Chem. Soc.* **123**, 7305–7313.

Prompers, J.J. and Brüschweiler, R. (2002) *J. Am. Chem. Soc.* **124**, 4522–4534.

Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) *Science* **257**, 76–80.

Reiter, N.J., Blad, H., Abildgaard, F. and Butcher, S.E. (2004) *Biochemistry* **43**, 13739–13747.

Robbins, A.H. and Stout, C.D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3639–3643.

Ryckaert, J.-P., Ciccotti, G. and van Gunsteren, W.F. (1977) *J. Comput. Phys.* **23**, 327–341.

Showalter, S.A. and Hall, K.B. (2002) *J. Mol. Biol.* **322**, 533–542.

Showalter S.A. and Hall, K.B. (2005) *Methods Enzymol.* **394**, in press.

Sitkoff, D. and Case, D.A. (1998) *Prog. NMR Spectrosc.* **32**, 165–229.

Stueber, D. and Grant, D.M. (2002) *J. Am. Chem. Soc.* **124**, 10539–10551.

Vitalis, A., Baker, N.A. and McCammon, J.A. (2004) *Mol. Simulation* **30**, 45–61.

Vugmeyster, L., Raleigh, D.P., Palmer, III A.G. and Vugmeister, B.E. (2003) *J. Am. Chem. Soc.* **125**, 8400–8404.

Yamazaki, T., Muhandiram, R. and Kay, L.E. (1994) *J. Am. Chem. Soc.* **116**, 8266–8278.