J Biomol NMR (2012) 54:69–80
DOI 10.1007/s10858-012-9655-0

ARTICLE

Independent alignment of RNA for dynamic studies using residual dipolar couplings

Michael F. Bardaro Jr. · Gabriele Varani

Received: 7 December 2011 / Accepted: 26 June 2012 / Published online: 18 July 2012

Abstract Molecular motion and dynamics play an essential role in the biological function of many RNAs. An important source of information on biomolecular motion can be found in residual dipolar couplings which contain dynamics information over the entire ms-ps timescale. However, these methods are not fully applicable to RNA because nucleic acid molecules tend to align in a highly collinear manner in different alignment media. As a consequence, information on dynamics that can be obtained with this method is limited. In order to overcome this limitation, we have generated a chimeric RNA containing both the wild type TAR RNA, the target of our investigation of dynamics, as well as the binding site for U1A protein. When U1A protein was bound to the portion of the chimeric RNA containing its binding site, we obtained independent alignment of TAR by exploiting the physical chemical characteristics of this protein. This technique can allow the extraction of new information on RNA dynamics, which is particularly important for time scales not covered by relaxation methods where important RNA motions occur.

Keywords RNA dynamics · NMR · Residual dipolar couplings · Independent alignment

M. F. Bardaro Jr. · G. Varani (✉)
Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195, USA
e-mail: varani@chem.washington.edu

G. Varani
Department of Biochemistry, University of Washington, Box 357350, Seattle, WA 98195, USA

Introduction

The HIV-1 TAR RNA is a paradigmatic example of biological function requiring a conformational change in RNA (Al-Hashimi 2005; Zhang et al. 2006; Leulliot and Varani 2001; Bardaro et al. 2011). Specifically, binding of HIV-1 Tat protein and the subsequent recruitment of human cyclin T1 for transcriptional activation of the HIV promoter depend on the ability of TAR to alter its structure in response to ligand binding (Karn 1999). TAR binds to a variety of ligands resulting in a large ensemble of known structures (Bardaro et al. 2009; Zhang et al. 2006). Although ligand binding induces a considerable conformational change in TAR, traditional NMR relaxation and $T_{1\rho}$ power dependence experiments (Hansen and Al-Hashimi 2007; Dayie et al. 2002; Bardaro et al. 2009) have been unable to sample the timescales where relevant dynamics occur (Olsen et al. 2010; Zhang et al. 2007). Residual dipolar couplings, however, indicate there are large scale global motions occurring between the two helices with the bulge region acting as a hinge, which occur at a rate too slow to be observed by relaxation methods (Al-Hashimi et al. 2002; Zhang et al. 2006, 2007) yet too fast for relaxation dispersion studies.

Consistent with those RDC studies and in contrast to solution state relaxation experiments, solid state NMR (ssNMR) studies of TAR dynamics discovered a plethora of motions occurring in the μs-ns timescale and described their characteristics at both the domain and residue level (Olsen et al. 2008; Olsen et al. 2010; Olsen et al. 2009). The ssNMR deuterium techniques generate robust results after motional modeling of the data, but they require considerable acquisition time in addition to cumbersome deuterium labeling; progress is slow because single site labeling is required at the present time to extract detailed

motional information. Furthermore, an independent method of validation of the solid state data under solution conditions would be of great value. For proteins, it has been possible to attain accurate as well as detailed information regarding amplitudes of motion (although not rates) in the time scale where solution relaxation methods are essentially blind by using residual dipolar couplings (Meiler et al. 2001; Blackledge 2005; Lange et al. 2008). Thus, performing a motional analysis using RDCs may address the fundamental question of how TAR RNA moves as it performs its function by allowing for the observation of local (atomic site) dynamics over the entire ps-ms timescale, a much wider timescale than is probed by relaxation experiments. Although exact rates of motion would remain unknown, this would undoubtedly be a considerable advancement over current knowledge.

In order to account for the orientational dependence of RDCs, data analysis requires a high quality structure (establishing the bond vector orientation relative to a local axis of motion) and the measurement of RDCs in multiple independent media which induce the molecule to assume a particular orientation determined by its interaction with the alignment media. The use of multiple independent media effectively allows for the separation of the orientational effects on the residual dipolar couplings. Because this technique relies on the exploitation of orientational averaging properties of RDCs to characterize the dynamic modes of RNA bases and riboses, the accuracy of this analysis depends on each medium inducing a sufficiently different orientation for TAR. Therefore, the success of the technique relies on the different dependence of RDCs on internal anisotropic motion in the presence of differently oriented tensors.

Thus far, with one exception, it has not been possible to attain multiple data sets of aligned RDCs for RNA with sufficient orthogonality (Latham et al. 2008; Higman et al. 2011). In the only exception, independent alignment was achieved between two data sets by elongating the lower stem of TAR until it became the dominant factor in aligning the RNA (Dethoff et al. 2010). Since nucleic acids have a fairly uniform negative electrostatic distribution on a mesoscopic scale, they interact with each alignment media in a similar fashion. Indeed when samples of TAR were subjected by us to a number of different alignment media, three of these media (Pf1, Peg/Hexanol and Glucopone/Hexanol) produced RDC data sets that were highly correlated with each other, indicating TAR assumes highly similar orientations in each of these alignment media.

In order to overcome RNA's inherent reluctance to align with any sort of independence, we reasoned that perhaps binding the RNA to a protein would generate a molecular complex with sufficient electrostatic anisotropy to obtain alignments with varying levels of orthogonality. For this approach to work, however, the protein must bind tightly so as to form a complex that does not readily dissociate or, more importantly, lead to resonance broadening, and the protein itself cannot perturb the RNA being investigated. Thus, we created a chimeric RNA containing both TAR and the binding site for human U1A protein in a chimeric dumbbell construct. The protein binds to its cognate site which is separated from TAR, and forms a sub-nM complex with extremely slow off rates. The RDC data set collected on this sample is minimally correlated with the same data set collected on the original TAR RNA in the same media, thereby demonstrating independent alignment.

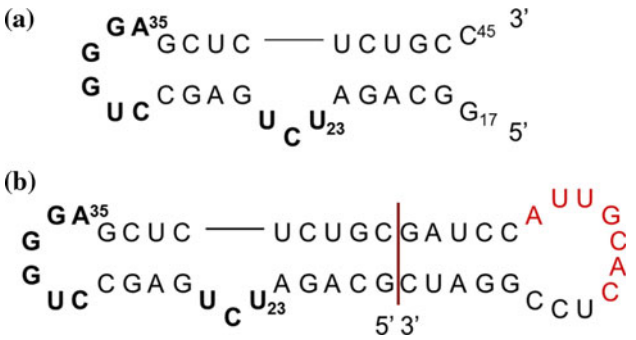

Fig. 1 **a** Sequence and secondary structure of the TAR construct studied here; **b** chimeric TAR1A sequence with the U1A binding site shown in red. The *maroon line* shows the connection point between the two RNAs

Materials and methods

RNA sample preparation

All RNAs used in this study were prepared enzymatically via the phage T7 RNA polymerase in vitro transcription method (Milligan et al. 1987; Milligan and Uhlenbeck 1989) using ^{13}C–^{15}N-labeled nucleotides purchased from Isotec and *in house* purified polymerase. DNA templates were purchased from IDT and included 2′-O-methyl groups in the last two residues at the 5′ end of the template in order to reduce further addition of nucleotides past the template end (Varani et al. 1996). The RNAs were purified by denaturing gel electrophoresis followed by electro-elution and ethanol precipitation. Finally, microdialysis was performed in a 10 mM potassium phosphate buffer (pH 6.6) with 0.01 mM EDTA.

In addition to preparing TAR fully labeled samples, the chimeric TAR-U1A construct (referred to as TAR1A) samples were also made with a partial labeling scheme containing either GU or AC-labeled nucleotides to reduce spectral overlap. All samples were dialyzed into 10 mM potassium phosphate and 10 mM KCl buffer (pH 6.2) with 0.01 mM EDTA. The TAR1A sequence, Fig. 1, combines

the wt-TAR sequence (without G17 and C45) with the U1 snRNA hairpin II sequence (A62-U79) in a dumbbell construct.

U1A preparation

The wt-U1A 1:98 double mutant (Tyr31-His and Gln36-Arg) plasmid (Allain et al. 1996) was cloned into the Pet28 kanamycin expression vector and transformed into Rosetta cells with a 45 s heat shock. The batch of ^{15}N labeled U1A (prepared to observe proper binding to the RNA) was made and purified the same way, except that minimal M9 medium (containing ^{15}N labeled NH_4Cl) was used instead of 2XYT broth.

Complex formation

The complex of unlabeled TAR1A with ^{15}N labeled U1A was prepared by slowly titrating the desired amount of protein (110 % the molar amount of RNA) into a TAR1A sample and running a ^{15}N-HSQC after each addition to observe the progression in complex formation. The final U1A HSQC in the titration series was compared to spectra of U1A bound to the 3′-UTR of its mRNA which contains the same seven nucleotide that U1A protein recognizes (Gubser and Varani 1996; Allain et al. 1996) and found to be nearly identical. The initial spectrum of the unbound U1A protein was also found to be identical with spectra from previous studies (Avis et al. 1996; Howe et al. 1994).

The complex of ^{13}C/^{15}N labeled TAR1A with unlabeled U1A was prepared by slowly titrating the desired amount of protein (110 % the molar amount of RNA) into a TAR1A sample and running a ^{13}C-HSQC after each addition to observe the progression of complex formation. A final molar ratio of 1.1:1 of U1A to ^{13}C/^{15}N TAR1A was used for all RDC experiments on TAR1A.

Sample alignment

Four different media were used to align TAR, and one of the media was also used to align the TAR1A-U1A complex. Alignment with filamentous Pf1 bacteriophage was accomplished by adding phage to a final concentration of ~17 mg/mL (Getz et al. 2007; Bardaro et al. 2011; Latham et al. 2008). Pf1 was the only medium to be used on both wt-TAR and the chimeric variant, all other media were applied thus far only to TAR. Alignment was also achieved using a n-alkyl-poly(ethylene glycol)/n-hexanol mixture as well as a glucopone/n-hexanol mixture (Ruckert and Otting 2000). Specifically, $C_{12}E_6$ (where 12 is the number of carbons in the n-alkyl group and 6 is the number of glycol units in poly(ethylene glycol)) was added to the TAR sample to a final concentration of 5 %. Hexanol was then added until a final molar ratio of $C_{12}E_6$:hexanol of 0.64 was reached. The process for glucopone was similar. First, glucopone was added to the RNA sample until a final concentration of 4 % was achieved. Then hexanol was added to a final concentration of 0.57 %. In the case of both alcohol mixtures, samples were biphasic at low hexanol concentrations. However, once the desired concentration was reached, the samples became monophasic as well as transparent.

The final media used for the induction of RNA alignment was a negatively charged 6.0 % acrylamide gel (Chou et al. 2001; Ishii et al. 2001). In order to introduce the negative charge, a 5 % molar amount of acrylamide was replaced by an equimolar amount of 2-acryloylamino-2-methylpropane-1-sulfonic acid (AMPS). The gel was then added to a special NMR tube (with a typical 5 mm diameter) by previously published means (Chou et al. 2001).

The alignment of the TAR samples was verified by measuring deuterium splitting (Ottiger and Bax 1998) on a different 499 MHz Avance instrument at 25 °C. The deuterium splitting of TAR1A bound to U1A protein was measured on the same instrument at both 25 and 32 °C (Ottiger and Bax 1998).

NMR experiments

All TAR data collection was performed on a Bruker Avance-500 instrument in the aforementioned buffers with 99.9 % D_2O and at 25 °C. For TAR1A samples bound to the U1A protein, a temperature of 32 °C was used due to the enhancement in sensitivity and the potential for the use of additional alignment media. The data for all samples was collected using a cryoprobe at a sample concentration of ~0.8 mM.

The In-Phase Anti-Phase (IPAP) CT-HSQC pulse sequence (Ottiger et al. 1998), HaCaCb (Fiala and Sklenar 2007), and 3D MQ-HCN-QJ (or its TROSY-HCN-QJ version) experiments (Jaroniec et al. 2005) were used to measure RDCs on the uniformly ^{13}C/^{15}N labeled HIV-1 TAR samples. Only the IPAP HSQC experiment was used on the ^{13}C/^{15}N TAR1A-U1A complex as the other experiments were too insensitive to obtain usable spectra. Only single bond couplings were used for the purpose of this study.

The IPAP data for the bases ($^1D_{C6H6}$, $^1D_{C8H8}$) were collected with the ^{13}C frequency set at 144.5 ppm using 172 points, while the data for the sugars ($^1D_{C1'H1'}$, $^1D_{C2'H2'}$, $^1D_{C3'H3'}$, $^1D_{C4'H4'}$, $^1D_{C5'H5'}$, and $^1D_{C5''H5''}$) were collected with the ^{13}C frequency set at 73.5 ppm using 300 points; the spectral width was set at 24 ppm for both datasets. The carrier in the ^{1}H dimension was set to the resonance

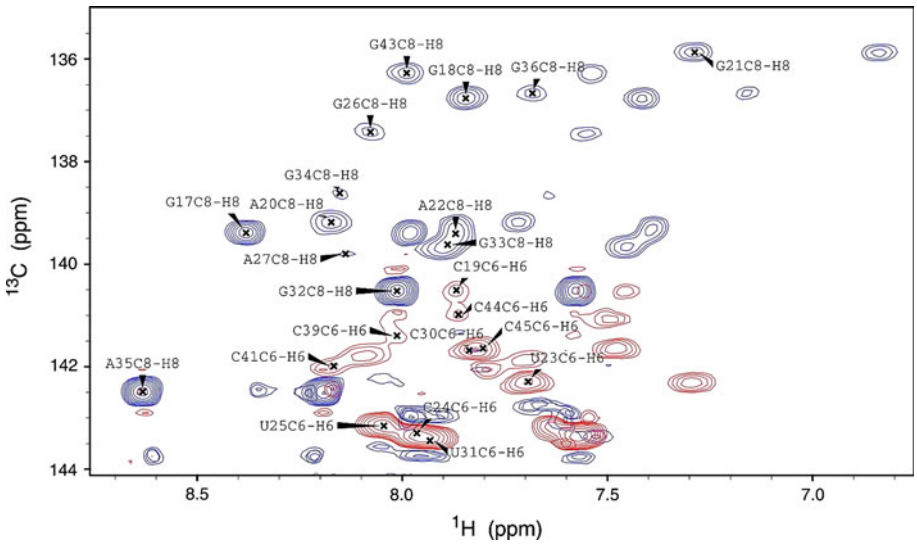

Fig. 2 [^{1}H,^{13}C] HSQC IPAP spectrum collected on wt-TAR in a negatively charged acrylamide gel showing the data used to extract the $^{1}D_{C6H6}$ and $^{1}D_{C8H8}$ RNA base couplings. The singlets of both the "In-Phase" and "Anti-Phase" experiments have been overlapped to show the doublets resulting from both J-coupling and dipolar coupling

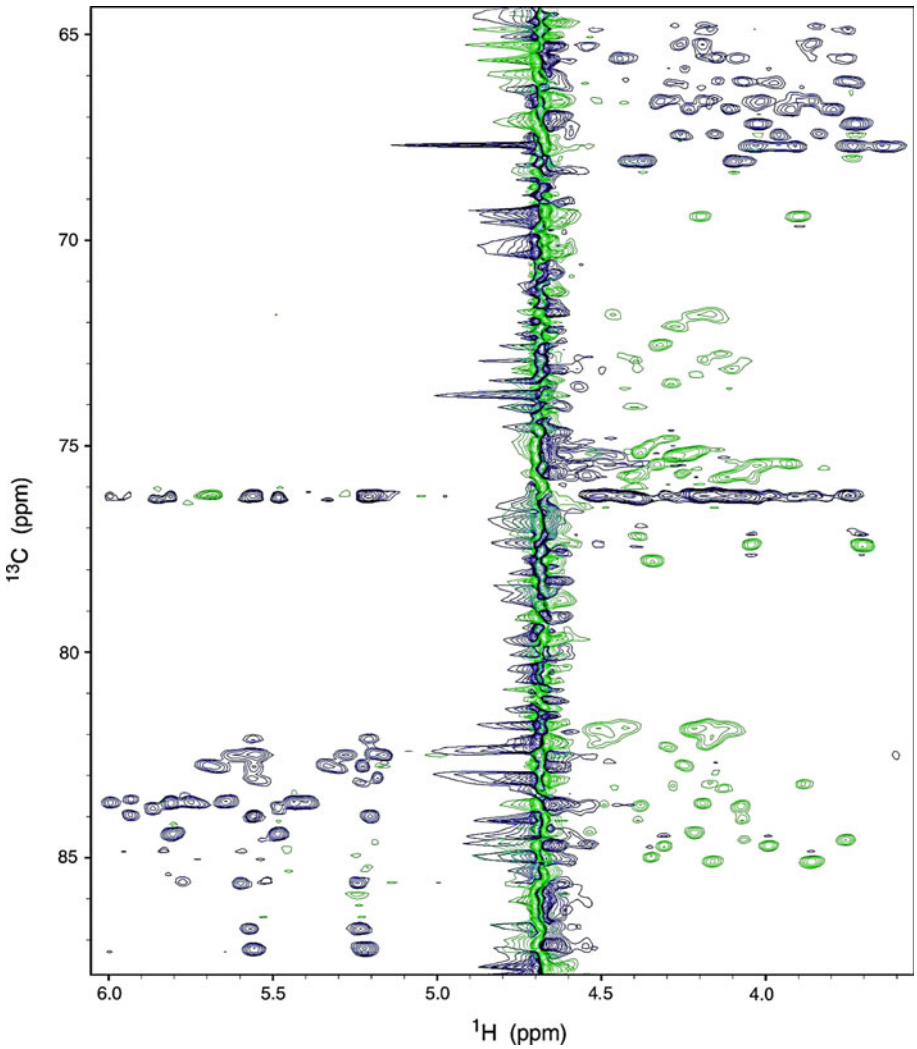

Fig. 3 [^{1}H,^{13}C] IPAP spectrum collected on wt-TAR in the Pf1 medium showing data used to extract $^{1}D_{C1'H1'}$, $^{1}D_{C2'H2'}$, $^{1}D_{C3'H3'}$, $^{1}D_{C4'H4'}$, $^{1}D_{C5'H5'}$, and $^{1}D_{C5''H5''}$ ribose couplings. The singlets of both the "In-Phase" and "Anti-Phase" experiments have been overlapped to show the doublets resulting from both J-coupling and dipolar coupling

frequency of H_2O ($\sim$ 4.67 ppm) with a spectral width of 12 ppm. Figures 2 and 3 show IPAP data of TAR collected in two different alignment media for the base and sugar regions, respectively. The acquisition times for the IPAP experiments on the bases were 113 ms (^{1}H) and 28 ms (^{13}C) and 205 ms (^{1}H) and 49 ms (^{13}C) for the sugars.

The data for the HaCaCb (pyrimidine $^{1}D_{C5C6}$) were collected with the ^{13}C frequency set at 142 ppm. The spectral width in this dimension was 30 ppm. The carrier in the ^{1}H dimension was set to a frequency of ~7.80 ppm with a spectral width of 5 ppm. 412 points were collected in this dimension. The resulting acquisition times were 272 ms (^{13}C) × 82 ms (^{1}H). Figure 4 shows a representative HaCaCb spectrum of unaligned TAR.

The data for the 3D MQ-HCN-QJ (from which purine $^{1}D_{N9C1'}$ and pyrimidine $^{1}D_{N1C1'}$ couplings were extracted) were collected with the ^{13}C frequency set at 90 ppm, 72 points and a spectral width of 28 ppm. The carrier in the ^{1}H dimension was set to a frequency of ~4.70 ppm with a spectral width of 6.0 ppm and 648 points were collected. Lastly, the ^{15}N dimension had a carrier set to 156 ppm, 200 points and a sweep width of 37 ppm. These parameters lead to acquisition times of 107 ms (^{1}H), 53 ms (^{15}N) and 10 ms (^{13}C).

The analysis of these NMR experiments, resulting in the reported RDC values, was performed as stated in previous work (Bardaro et al. 2011; Jaroniec et al. 2005; Fiala and Sklenar 2007).

Results

TAR is not independently aligned in four different media

In order to analyze intermediate time scale motion in TAR, we aligned this RNA in four media and collected four extensive sets of RDC's comprising $^{1}D_{C6H6}$, $^{1}D_{C8H8}$, $^{1}D_{C1'H1'}$, $^{1}D_{C4'H4'}$, $^{1}D_{C5'H5'}$, and $^{1}D_{C5''H5''}$, pyrimidine $^{1}D_{C5C6}$, purine $^{1}D_{N9C1'}$ and pyrimidine $^{1}D_{N1C1'}$ couplings. However, an accurate motional analysis requires the collection of multiple set of RDC's from sufficiently independently aligned media (Bouvignies et al. 2005). Despite the collection of such a large amount of data, Fig. 5 (illustrating the $^{1}D_{C6H6}$, $^{1}D_{C8H8}$ and $^{1}D_{C1'H1'}$couplings), intuitively indicates the data sets are highly correlated with each other. The noticeable correlation of these data sets suggests that TAR align in a collinear manner in all media evaluated, and that all alignment tensors are highly similar to the alignment tensor for Pf1 phage medium.

To confirm this intuitive conclusion, a Pearson's correlation analysis (Eq. 1) was performed (Latham et al. 2008) in order to quantitate the level of correlation between the RDC data sets collected in different alignment media. The results of these analyses are shown in Table 1. If all possible sites in TAR have corresponding RDC value measurements, there are 29 degrees of freedom (for example in the bases, there are 29 nucleotides for which there are measured RDCs in Pf1 and 26 in the PEG/Hexanol mixture, meaning there are only 26 degrees of freedom in these media). In this case, any R value greater than 0.46 would indicate a very high probability (at a 99 % confidence level) of correlation between the nominally independent data sets. The lowest correlation observed between data sets on TAR is between Pf1 and the negatively charged acrylamide gel, showing a correlation of 0.61. Thus, all RDC datasets (except perhaps the RDCs observed when wt-TAR is aligned in the acrylamide gel and in Pf1 phages) appear too highly correlated with each other to allow for the determination of atomic level motions.

$$R = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (1)$$

Other studies have reported similar conclusions when attempting to obtain independent alignments of RNA (Latham et al. 2008; Higman et al. 2011). Clearly, a different approach is needed if RNA is to be aligned independently, leading to sufficiently different alignment tensors to perform a complete analysis of dynamic amplitudes.

Formation of the TAR1A-U1A complex

We reasoned that, since many proteins align independently in different media (Bax 2003), RNA may align differently if it is bound to a protein. In order to bind TAR to a protein in such a way that RNA secondary structure and dynamics are not perturbed, we generated the TAR1A construct (Fig. 1). This chimeric RNA forms a dumbbell structure which contains the wt-TAR sequence (without G17 and C45) spliced to the U1 snRNA hairpin II (A62-U79) (Oubridge et al. 1994; Howe et al. 1994). The addition of the U1 snRNA hairpin II sequence allows for binding of the U1A protein to this RNA. We used the 1:98 double mutant (Tyr31-His and Gln36-Arg) version of the protein with full binding activity and better NMR behavior (Oubridge et al. 1994; Avis et al. 1996; Allain et al. 1996, 1997). By creating a new RNA binding site for the protein away from the TAR structure, we assumed structural and motional properties for this RNA would be preserved.

In order to verify formation of the desired secondary structure, several HSQCs experiments were recorded. First [^{1}H, ^{15}N]-HSQCs of unbound and TAR1A-bound ^{15}N labeled protein were collected, as shown in Fig. 6. The ^{15}N HSQCs on the free protein showed the protein structure to be exactly the same as in previous studies (Avis et al. 1996). Furthermore, HSQCs on the complex indicated that the U1A protein bound to the RNA in a highly similar fashion to what has been reported before (Howe et al. 1994).

We then recorded [^{1}H, ^{13}C]-HSQC experiments on the unbound and U1A-bound $^{13}C/^{15}N$ labeled TAR1A, as

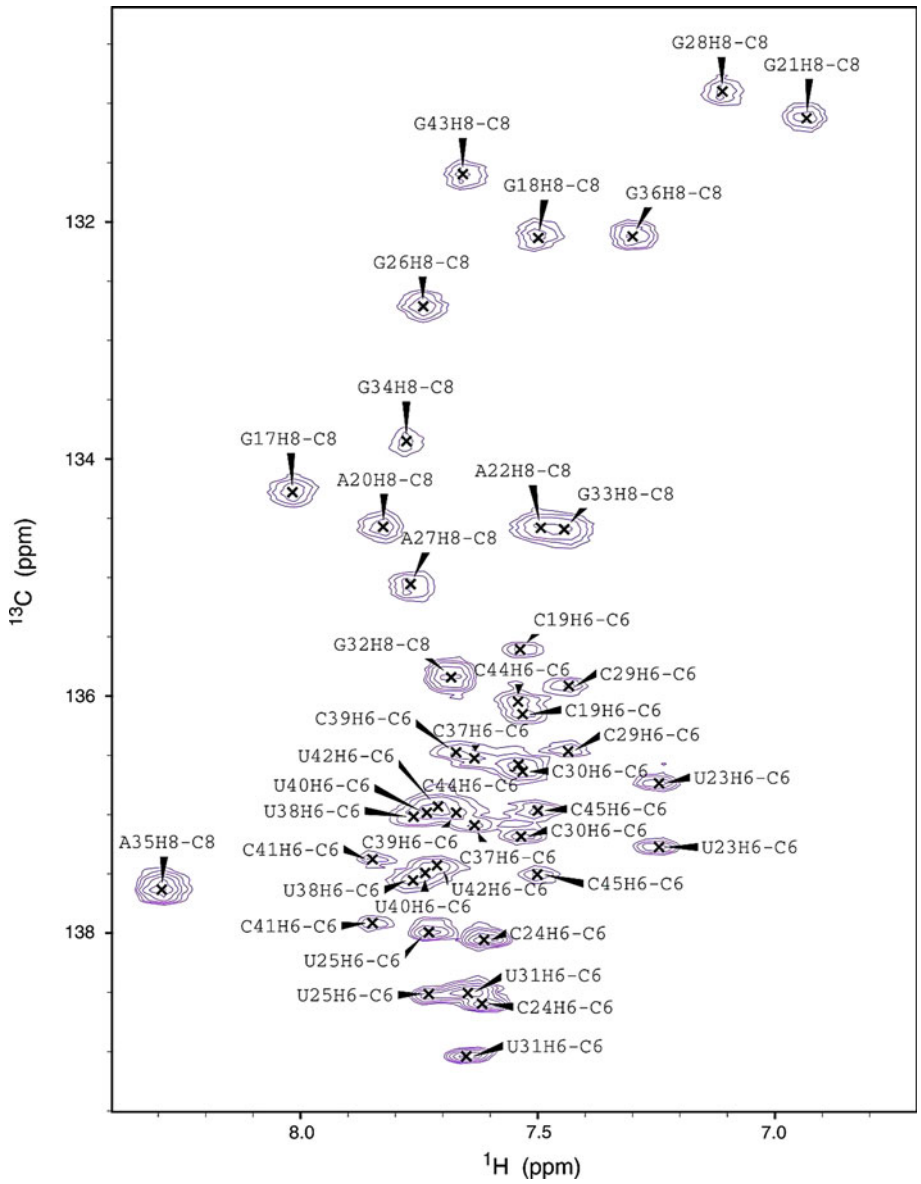

Fig. 4 A HaCaCb spectrum on wt-TAR in the absence of alignment media showing data used to extract the pyrimidine $^{1}D_{C5C6}$ couplings. Peak assignments are also indicated on the figure

shown in Fig. 7a. The ^{13}C HSQCs show that the secondary structure of the RNA remains the same in the presence of the protein, with only a few chemical shift perturbations to resonances belonging to the TAR part of the variant sequence, aside from terminal residues. Figure 7a shows that the only wt-TAR residue significantly perturbed in the presence of U1A is the extra-helical A35 H2 resonance, with smaller effects for the H8 of the same residue. We believe that these changes are most likely to arise from changes in solvent conditions upon addition of U1A, because the changes are small and limited to this flexible and solvent exposed residue.

Another set of HSQC experiments were performed on the bound TAR1A at 32 °C and were found to have considerably increased sensitivity (due to the decreased solvent viscosity) without affecting chemical shifts in any significant way. As a result of this increased sensitivity, all RDC experiments on TAR1A were performed at 32 °C. A fortunate side effect of the change in temperature is that more alignment media become available at this higher temperature.

RDCs for the TAR1A-U1A complex

A full set of RDCs (as collected for the TAR construct) was extremely difficult to collect on the ^{13}C/^{15}N TAR1A-U1A for a multitude of reasons. First, spectral overlap became a significant problem as most RDC experiments require the measurement of doublets rather than single peaks. To reduce these problems, two samples were prepared. The first consisted of TAR1A having ^{13}C/^{15}N uniform labels on A and C, while the second consisted of uniformly labeled G

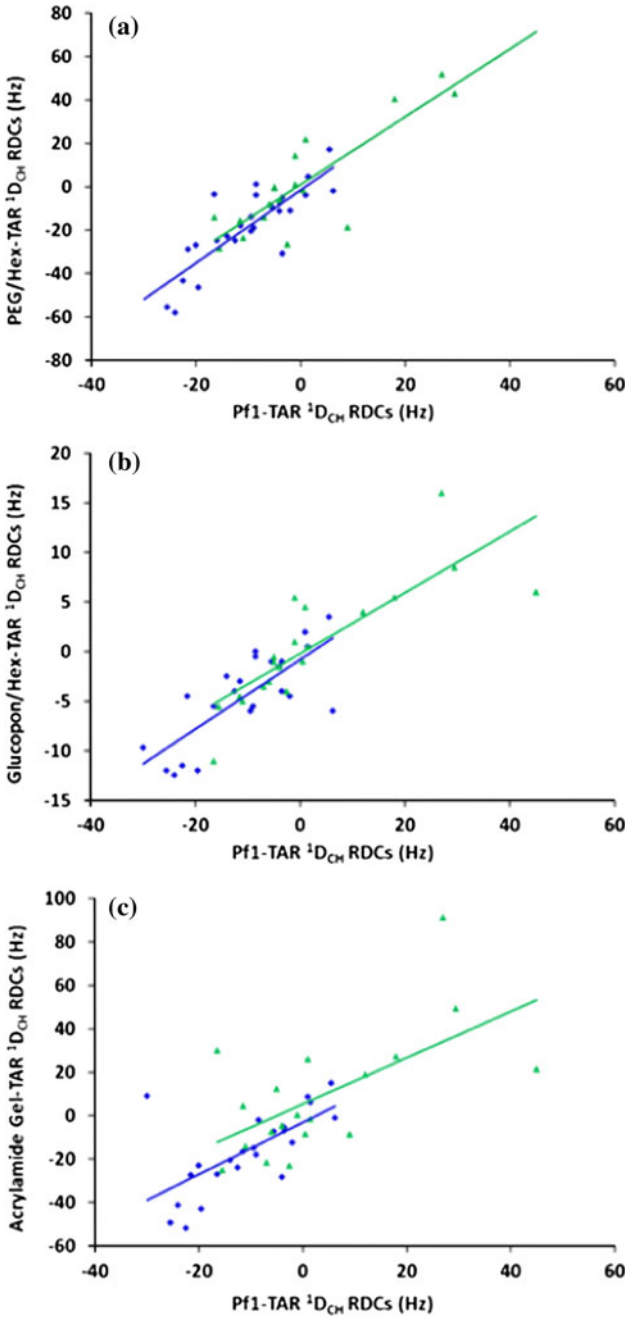

Fig. 5 Scatter plots of C6-H6, C8-H8 (*blue diamonds*) and C1′-H1′ (*green triangles*) RDCs (Hz) collected for wt-TAR aligned in Pf1 phage plotted versus the same couplings collected in **a** Peg/Hexanol, **b** Glucopone/Hexanol **c** stretched negatively charged acrylamide gel. The *lines* are the best linear fit for each plot and are colored to match their corresponding data set

and U sites. Figure 7b shows an unaligned IPAP spectrum for the base region taken on the variant TAR RNA bound to the U1A protein with only A's and C's being uniformly ^{13}C labeled ('AC' labeled sample). Since the RDCs were collected for two separate RNA–protein complexes ('AC' and 'GU' $^{13}C/^{15}N$ labeled samples), the degree of

Table 1 Pearson's correlation coefficients calculated for each data set when compared to the same set of RDC's when collected for wt-TAR in the Pf1 phage alignment medium

Alignment media	Correlation value (C6/C8)	Correlation value (C1′)
Peg:Hexanol	0.82	0.85
Glucopone:Hexanol	0.77	0.82
Neg. acrylamide gel	0.65	0.61
Pf1 (TAR1A)	0.50	NA

alignment was slightly different, because the volume of Pf1 in each sample cannot be perfectly reproduced. In order to overcome this problem, the ^2H splittings were measured for both samples and used to normalize the data sets.

In addition to the spectral overlap, the sensitivity loss due to the considerable increase in complex size (from a 9 kDa RNA to a 28 k Da complex) had a detrimental effect on the quality of the data. The data recorded for the TAR1A-U1A complex in Pf1 phages had peak widths that were on average $\sim$4.0 Hz larger in the proton dimension compared to the wt-TAR RNA aligned in the same medium. In order to partially offset the effects of increased line width on sensitivity, the temperature was increased to 32 °C. The effects of temperature on RDC values were established to be negligible by using 1D-^2H splitting experiments on aligned samples at 25 and 32 °C.

Due to the reduced sensitivity, only IPAP experiments could be recorded with sufficient signal/noise to measure RDC's accurately. Using this approach, $^1D_{C6H6}$, $^1D_{C8H8}$, $^1D_{C1'H1'}$, $^1D_{C4'H4'}$, $^1D_{C5'H5'}$, and $^1D_{C5''H5''}$ couplings were collected, but only $^1D_{C6H6}$, $^1D_{C8H8}$, and $^1D_{C1'H1'}$ were analyzed because the other measurements were too overlapped to provide accurate RDC values. Figure 7c shows an unaligned spectrum of TAR1A bound to U1A and a Pf1-aligned spectrum of TAR1A bound to U1A, overlapped onto each another, highlighting the C8–H8 resonances in the 'GU'-labeled sample.

The RDCs collected for the TAR region of the $^{13}C/^{15}N$ TAR1A-U1A complex were then compared to data obtained in the previous alignment on TAR in the same media (Pf1). The scatter plot shown in Fig. 8 indicates that these data sets are highly uncorrelated with each other, suggesting that the U1A protein induces substantially independent alignment of RNA in the Pf1 medium.

A Pearson's correlation analysis was performed in order to quantitate the level of any possible correlation. The determined Pearson's correlation coefficient in Table 1 for these two data sets is 0.50. In this case there are only 18 data points (nucleotides with an accurate RDC measurement), meaning that an R-value of 0.56 is necessary to state that there is a correlation at the 99 % confidence level. Clearly, 0.50 falls below this value, strongly suggesting

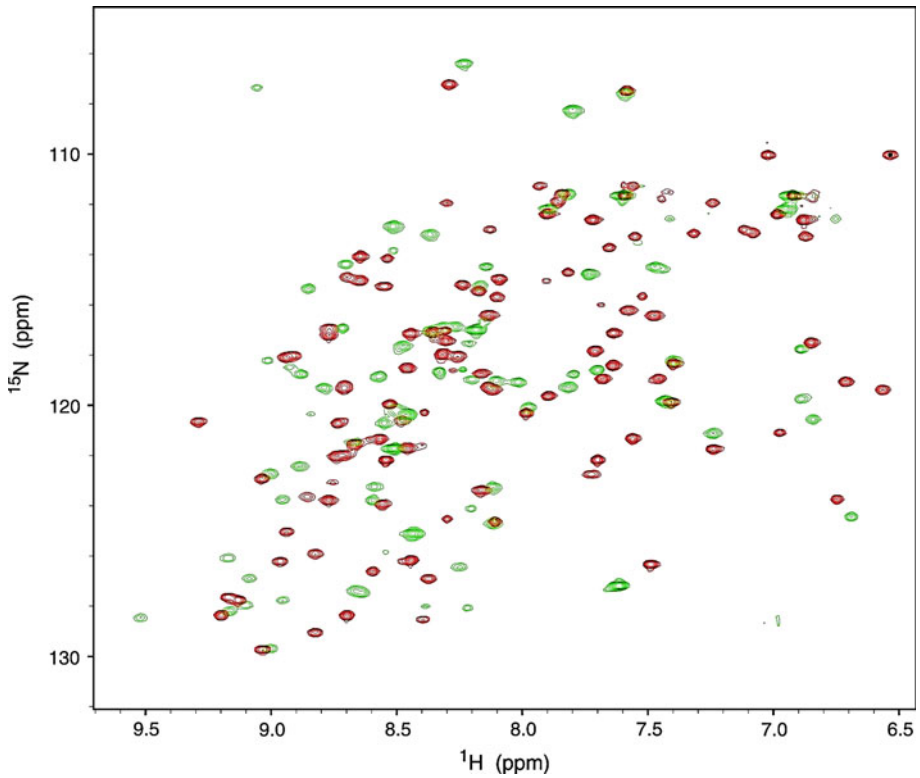

Fig. 6 [^{1}H, ^{15}N]-HSQC spectra of U1A free in solution (*red*) and bound to TAR1A (1:1) (*green*). Spectral changes observed upon RNA addition demonstrate that the protein binds to the U1A portion of the construct as previously reported

that these two data sets are likely to be independent of each other, suggesting that at least partially independent alignment was achieved.

In order to validate the results of the Pearson's analysis of the raw RDC data, alignment tensor parameters were determined for each of TAR's two helical stems in all four different alignment media using the deposited TAR structure (PDB 1ANR model 1) to perform a singular value decomposition analysis through the PALES program (Zweckstetter 2008; Zweckstetter and Bax 2000; Zweckstetter et al. 2004). The tensors for the two helical stems of TAR were calculated separately, as done in the past (Al-Hashimi et al. 2002), because a single alignment tensor cannot accurately reproduce experimental RDCs due to the inter-helical motions. The accuracy of the calculated tensors was then tested by using them to back-calculate the RDCs, as predicted from the TAR structure. The results of this analysis are shown in Table 2, where it is immediately possible to observe the high degree of correlation between the experimental RDCs and those back calculated using the alignment tensors calculated with PALES.

Thus, the analysis of the rhombicity values leads to the same conclusion derived from the Pearson's correlation analysis of the experimental data. Only in one case do the rhombicity values for wt-TAR show a significant deviation from what is calculated for the Pf1 dataset, namely this occurs for the analysis of RDC's for the lower stem of TAR in the negatively charged acrylamide gel. This result is consistent with the Pearson analysis of the raw data reported earlier. All other alignment media lead to very similar alignment tensors. In contrast, the TAR1A tensor parameters are dissimilar from the tensors determined for wt-TAR in the same Pf1 alignment medium. Namely, the alignment of TAR1A in Pf1 leads to a significantly different rhombicity, over three times greater, than determined for wt-TAR in the same media. These differences indicate that TAR1A adopts a different preferential orientation, thus achieving at least partial orthogonality. This desired feature was not observed for any pairs of wt-TAR data sets aligned in different media.

In order to quantitate the differences in tensor orientation further, normalized scalar products between the fitted alignment tensors for wt-TAR and the TAR1A in the presence of each alignment media were determined and are shown in Table 3 (Sass et al. 1999). Values close to 1 indicate a high degree of collinearity between tensors, while values close to 0 indicate orthogonal alignment tensors. The scalar products can be used to obtain an "angle" between different tensors by taking the inverse cosine of the normalized scalar product value. The scalar products in Table 3 indicate that the negatively charged acrylamide gel used on wt-TAR and the TAR1A-U1A complex in the presence of Pf1 yield orientation tensors that are largely orthogonal with respect to each other and to all other media. The analysis of the tensors extracted for the remaining three media used on wt-TAR indicate a largely collinear alignment, supporting the more qualitative analysis performed using Pearson's coefficients.

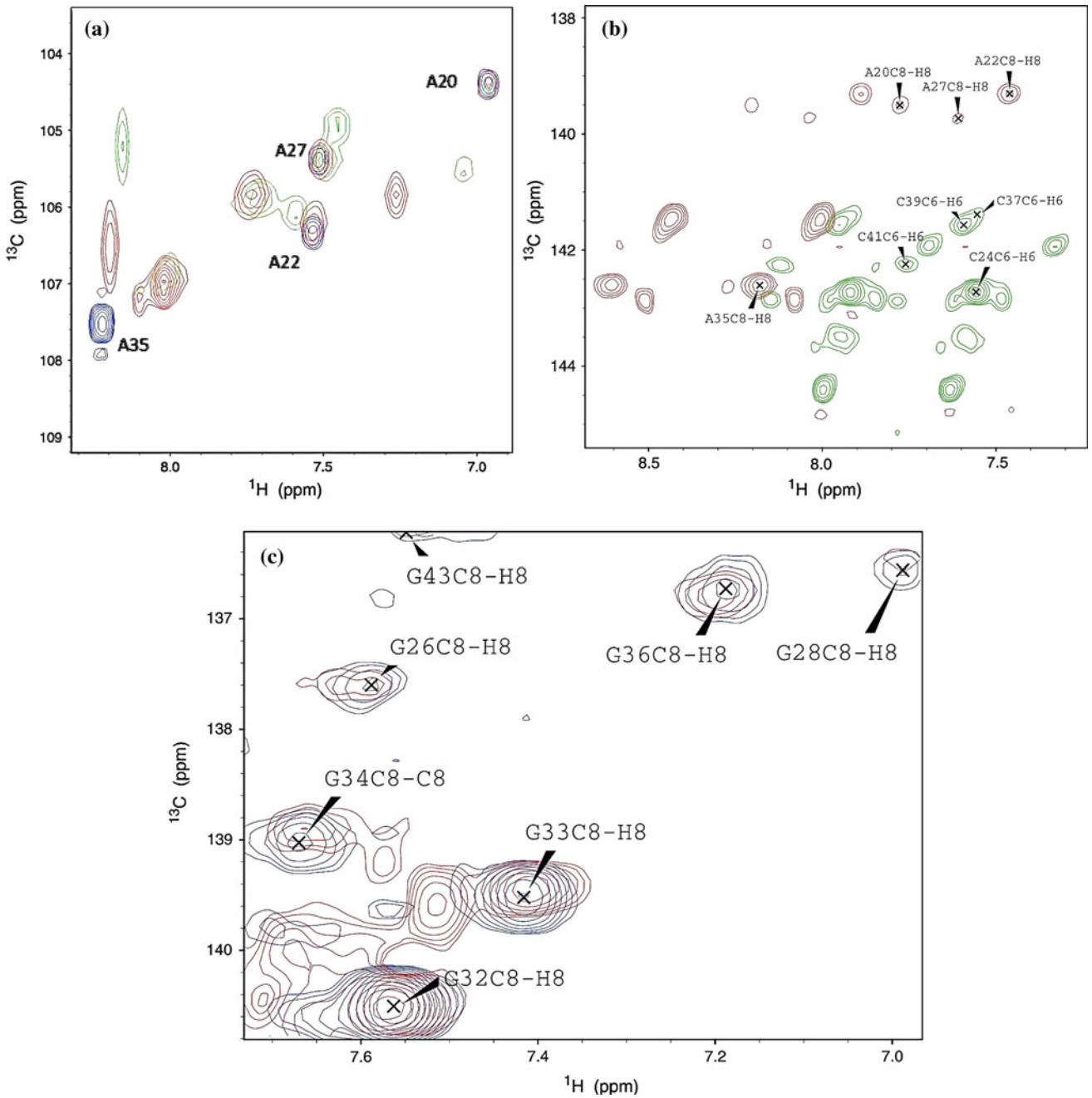

Fig. 7 **a** Overlapped [^{1}H,^{13}C]-HSQC spectra of the C2-H2 region of wt-TAR (*blue*), TAR1A (*red*) and TAR1A bound to U1A protein (*green*). **b** Unaligned [^{1}H,^{13}C]-IPAP spectra (both components were overlapped to show the splittings) of the base region of TAR1A bound to U1A. **c** [^{1}H,^{13}C]-IPAP spectra of the base region of an unaligned GC labeled TAR1A sample bound to U1A protein with a Pf1-aligned and GC-labeled TAR1A also bound to U1A protein. Only the in-phase peaks of each spectrum are shown in order to reduce overlap

Discussion

Independent alignment of RNA

Despite the physical–chemical variety of the media used to align RNA, all RDC data sets collected by us were determined to be highly correlated with each other using a Pearson's analysis or when explicitly determining alignment tensors using Pales. Unfortunately, this phenomenon is common to other RNAs as well, indicating that this is probably a universal concern when aligning RNA in different media (Latham et al. 2008; Higman et al. 2011).

The inability to align RNA with sufficient independence multiple times using different external media can be explained by its physical properties. RNA (and DNA as well) has a highly uniform and negatively charged

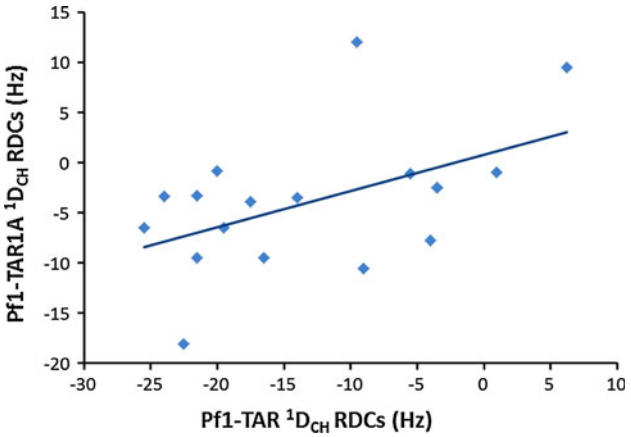

Fig. 8 Scatter plots of C6–H6 and C8–H8 RDCs (Hz) collected on wt-TAR aligned in Pf1 medium versus the same coupling collected for TAR1A bound to U1A protein and aligned in the same Pf1 medium. The line is the best linear fit for the plotted data

electrostatic distribution (Higman et al. 2011; Latham et al. 2008). Furthermore, many of the RNAs currently subjected to multiple alignment media have elongated rod-like structures. Since molecular alignment tensors in liquid crystal media can be accurately predicted through the use of either a steric obstruction model (Zweckstetter and Bax 2000) or a mixed steric obstruction and electrostatic repulsion model (Zweckstetter et al. 2004), it is unlikely that a molecule with the properties mentioned above can align differently when exposed to negatively charged or neutral media. In principle, alignment mechanisms that do not rely solely on steric or electrostatic repulsion could generate independent alignment. However, even magnetic alignment of RNA failed to generate a sufficiently independent alignment tensor (Latham et al. 2008). Moreover, positively charged media have a high probability of interacting with the negatively charged RNA, leading to aggregation.

We reasoned that the alignment properties of proteins could be used to independently align RNA molecules. In order to test this hypothesis, a chimeric RNA was designed containing a binding site for the U1A protein spliced onto the TAR element. The construct was used to test whether the RNA would take on the alignment properties of the bound protein and align independently in different external media.

The results indicate that the U1A protein has a direct effect on the physical alignment of the RNA. U1A is similar to other proteins, such as GB3, GB1 and ubiquitin which have all been independently aligned using different media, due to their small size and similar steric and electrostatic properties. Proteins resistant to independent alignment tend to be larger protein complexes with more uniform electrostatics and elongated shapes (Higman et al. 2011).

Our strategy was therefore successful, but several words of caution are warranted. Using the method presented here, we can successfully measure several RDCs per nucleotide by using IPAP experiments. However, the sensitivity of the HaCaCb experiment is too low to measure the pyrimidine $^1D_{C5C6}$ accurately enough, unless partial deuteration is used to reduce the line widths, yet these couplings provide

Table 2 Calculated alignment tensor parameters, D_a and R, determined from analysis of the experimental RDCs are shown alongside linear correlation coefficient, Rcorr, rmsd and Q, between observed RDCs and those predicted from the TAR three dimensional structure using the *Best-Fit* module in the program PALES

RNA and stem	Media	Da (Hz)[a]	R[b]	RMSD[c]	Rcorr[d]	Q[e]	$^1D^f_{CH}$
wt-TAR lower stem	Pf1	10.7	0.08	0.65	0.972	0.163	10
wt-TAR upper stem	Pf1	30.0	0.13	0.58	0.996	0.060	10
wt-TAR lower stem	PegHex	24.4	0.06	0.86	0.984	0.114	7
wt-TAR upper stem	PegHex	51.0	0.15	1.53	0.994	0.097	8
wt-TAR lower stem	GluHex	5.8	0.21	0.33	0.981	0.149	9
wt-TAR upper stem	GluHex	15.5	0.20	0.22	0.998	0.050	7
wt-TAR lower stem	NegGel	24.4	0.31	1.30	0.987	0.159	8
wt-TAR upper stem	NegGel	50.1	0.15	1.82	0.995	0.098	8
TAR1A upper stem	Pf1	15.1	0.38	0.11	0.999	0.026	6

[a] The axial component of the alignment tensor normalized to the dipolar interaction constant of the one-bond CH internuclear vector ($^1D^{CH}_{max}$ * Da)

[b] The rhombicity of the alignment tensor (D_r/D_a)

[c] The root-mean-square deviation between predicted and experimental RDCs

[d] Pearson's linear correlation coefficient between predicted and experimental RDCs

[e] The average rmsd normalized by the average magnitude of experimental RDCs

[f] The number of CH couplings available for use in tensor determination

Table 3 Normalized scalar products between fitted alignment tensors for samples of TAR in the presence of different alignment media

	Peg-hexanol	Glucopon-hexanol	Acrylamide gel	Pf1 (*mutant TAR*)
Pf1 (*wt-TAR*)	0.974	0.994	0.400	−0.322
Peg-hexanol		0.953	0.560	−0.134
Glu-hexanol			0.367	−0.358
Acrylamide gel				0.732

The final column shows all four media in the presence of wt-TAR compared to the mutant construct bound to U1A in the presence of Pf1

the "motionless" or "rigid" couplings used to calculate the initial alignment tensor. The minimization of dynamics is important when calculating initial alignment tensors, since motion is obviously a source of systematic error when determining an experimental alignment tensor using a steric model (Bernado and Blackledge 2004). The $^{1}D_{C4'H4'}$ couplings from the IPAP experiment are also lost due to poor sensitivity, and experiments used to measure the valuable purine $^{1}D_{N9C1'}$ and pyrimidine $^{1}D_{N1C1'}$ couplings are also too insensitive. However, it is possible to increase the sensitivity of NMR experiments on RNA considerably by using perdeuterated NTPs. Based on our experience in other projects, we are confident that the use of partially deuterated and $^{13}C/^{15}N$ NTPs will result in the enhanced sensitivity needed to measure additional RDCs.

Conclusions

Like other RNAs, TAR does not align with a sufficient level of independence in multiple alignment media used to generate partial alignment, limiting the dynamic information that can be extracted from RDC studies. However, by constructing a chimeric molecule containing the U1 snRNA hairpin II sequence (A62-U79) spliced to TAR and by binding U1A protein to this construct, we generated a sample that is independently aligned when compared to the TAR sequence placed in the same alignment medium. Despite drawbacks due to the increased size of the RNA and the resulting decreased sensitivity of NMR experiments, this design has the potential to solve a major obstacle in using RDCs for interpreting RNA dynamics.

Acknowledgments This work was supported by grant #0642253 to GV from the National Science Foundation.

References

Al-Hashimi HM (2005) Dynamics-based amplification of RNA function and its characterization by using NMR spectroscopy. ChemBioChem 6:1506–1519

Al-Hashimi HM, Gosser Y, Gorin A, Hu W, Majumdar A, Patel DJ (2002) Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. J Mol Biol 315:95–102

Allain FH, Gubser CC, Howe PW, Nagai K, Neuhaus D, Varani G (1996) Specificity of ribonucleoprotein interaction determined by RNA folding during complex formulation. Nature 380(6575): 646–650

Allain FH, Howe PW, Neuhaus D, Varani G (1997) Structural basis of the RNA-binding specificity of human U1A protein. EMBO J 16(18):5764–5772

Avis JM, Allain FH, Howe PW, Varani G, Nagai K, Neuhaus D (1996) Solution structure of the N-terminal RNP domain of U1A protein: the role of C-terminal residues in structure stability and RNA binding. J Mol Biol 257(2):398–411

Bardaro MF Jr, Shajani Z, Patora-Komisarska K, Robinson JA, Varani G (2009) How binding of small molecule and peptide ligands to HIV-1 TAR alters the RNA motional landscape. Nucleic Acids Res 37:1529–1540

Bardaro MF Jr, Pederson K, Drobny G, Varani G (2011) Biomolecular NMR spectroscopy. In: Andrew J, Dingley SMP (eds) Advances in biomedical spectroscopy, vol 3. IOS Press, The Netherlands, pp 279–301

Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 12(1):1–16

Bernado P, Blackledge M (2004) Anisotropic small amplitude peptide plane dynamics in proteins from residual dipolar couplings. J Am Chem Soc 126:4907–4920

Blackledge M (2005) Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. Prog Nucl Mag Res Spectrosc 46:23–61

Bouvignies G, Bernadó P, Blackledge M (2005) Protein backbone dynamcs from N–H dipolar couplings in partially aligned systems: a comparison of motional models in the presences of structural noise. J Magn Reson 173:328–338

Chou JJ, Gaemers S, Howder B, Louis JM, Bax A (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. J Biomol NMR 21:377–382

Dayie KT, Brodsky AS, Williamson JR (2002) Base flexibility in HIV-2 TAR RNA mapped by solution ^{15}N, ^{13}C NMR relaxation. J Mol Biol 317:263–278

Dethoff EA, Hansen AL, Zhang Q, Al-Hashimi HM (2010) Variable helix elongation as a tool to modulate RNA alignment and motional couplings. J Magn Reson 202(1):117–121. doi:10.1016/j.jmr.2009.09.022

Fiala R, Sklenar V (2007) ^{13}C-detected NMR experiments for measuring chemical shifts and coupling constants in nucleic acid bases. J Biomol NMR 39:153–163

Getz M, Sun X, Casiano-Negroni A, Zhang Q, Al-Hashimi HM (2007) NMR studies of RNA dynamics and structual plasticity using NMR residual dipolar couplings. Biopolymers 86:384–402

Gubser CC, Varani G (1996) Structure of the polyadenylation regulatory element of the human U1A pre-mRNA 3′-untranslated region and interaction with the U1A protein. Biochemistry 35(7):2253–2267

Hansen AL, Al-Hashimi HM (2007) RNA dynamics by carbon relaxation and domain elongation. J Am Chem Soc 129:16072–16082

Higman VA, Boyd J, Smith LJ, Redfield C (2011) Residual dipolar couplings: are multiple independent alignments always possible? J Biomol NMR 49:53–60

Howe PW, Nagai K, Neuhaus D, Varani G (1994) NMR studies of U1 snRNA recognition by the N-terminal RNP domain of the human U1A protein. EMBO J 13(16):3873–3881

Ishii Y, Markus MA, Tycko R (2001) Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. J Biomol NMR 21:141–155

Jaroniec CP, Boisbouvier J, Tworowska I, Nikonowicz EP, Bax A (2005) Accurate measurement of 15N–13C residual dipolar couplings in nucleic acids. J Biomol NMR 31:231–241

Karn J (1999) Tackling Tat. J Mol Biol 293:235–254

Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. Science 320:1471–1475

Latham MP, Hanson P, Brown DJ, Pardi A (2008) Comparison of alignment tensors generated for native tRNAVal using magnetic fields and liquid crystalline media. J Biomol NMR 40:83–94

Leulliot N, Varani G (2001) Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. Biochemistry 40(27):7947–7956

Meiler J, Prompers JJ, Peti W, Griesinger C, Bruschweiler R (2001) Model-free approach to the dynamics interpretation of residual dipolar couplings in globular proteins. J Am Chem Soc 123:6098–6107

Milligan JF, Uhlenbeck OC (1989) Synthesis of small RNAs using T7 RNA polymerase and synthetic DNA templates. Methods Enzymol 180:51–62

Milligan JF, Groebe DR, Witherell WG, Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res 15:8783–8798

Olsen GL, Echodu CD, Shajani Z, Bardaro MF Jr, Varani G, Drobny GP (2008) Solid-state deuterium NMR studies reveal micro-ns motions in the HIV-1 transactivation response RNA recognition site. J Am Chem Soc 10:2896–2897

Olsen GL, Bardaro MF Jr, Echodu CD, Drobny GP, Varani G (2009) Hydration dependent dynamics in RNA. J Biomol NMR 45(1–2):133–142

Olsen GL, Bardaro MF Jr, Echodu CD, Drobny GP, Varani G (2010) Intermediate rate atomic trajectories of RNA by solid-state NMR spectroscopy. J Am Chem Soc 132(1):303–308

Ottiger M, Bax A (1998) Characterization of magnetically oriented phospholipids micelles for measurement of dipolar couplings in macromolecules. J Bio Mol NMR 12:361–372

Ottiger M, Delaglio F, Bax A (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 131:373–378

Oubridge C, Ito N, Evans PR, Teo CH, Nagai K (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372(6505):432–438

Ruckert M, Otting G (2000) Alignment of biological macromolecules in novel nonionic liquid crystal media for NMR experiments. J Am Chem Soc 122:7793–7797

Sass J, Cordier F, Hoffmann A, Rogowski M, Cousin A, Omichinski JG, Lowen H, Grzesiek S (1999) Purple membrane induced alignment of biological macromolecules in the magnetic field. J Am Chem Soc 121:2047–2055

Varani G, Abou-ela F, Allain FHT (1996) NMR investigation of RNA structure. Progr Nuclear Magn Spectrosc 29:51–127

Zhang Q, Sun X, Watt ED, Al-Hashimi HM (2006) Resolving the motional modes that code for RNA adaption. Science 311:653–656

Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. Nature 450:1263–1268

Zweckstetter M (2008) NMR: prediction of molecular alignment from structure using the PALES software. Nat Protoc 3(4):679–690

Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. J Am Chem Soc 122(15):3791–3792. doi:10.1021/ja0000908

Zweckstetter M, Hummer G, Bax A (2004) Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. Biophys J 86(6):3444–3460