Binding of U1A Protein Changes RNA Dynamics As Observed by ^{13}C NMR Relaxation Studies[†,‡]

Zahra Shajani,[§] Gary Drobny,[§] and Gabriele Varani*[,§,||]

Departments of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195-1700

Received December 28, 2006; Revised Manuscript Received March 20, 2007

ABSTRACT: Recognition of RNA by proteins and small molecules often involves large changes in RNA structure and dynamics, yet very few studies have so far characterized these motional changes. Here we extend to the protein-bound RNA recent ^{13}C relaxation studies of motions in the RNA recognized by human U1A protein, a well-known model for protein−RNA recognition. Changes in relaxation observed upon complex formation demonstrate that the protein-binding site becomes rigid in the complex, but the upper stem−loop that defines the secondary structure of this RNA experiences unexpected motional freedom. By using a helix elongation strategy, we observe that the upper stem−loop moves independently of the remainder of the structure also in the absence of U1A. Surprisingly, RNA residues making important intermolecular contacts in the structure of the complex exhibit increased flexibility in the presence of the protein. Both of these results support the hypothesis that RNA-binding proteins select a structure that optimizes intermolecular contacts in the manifold of conformations sampled by the free RNA and that protein binding quenches these motions. Together with previous studies of the RNA-bound protein, they also demonstrate that protein−RNA interfaces experience complex motions that modulate the strength of individual interactions.

A detailed understanding of the molecular basis for RNA−protein recognition is essential to understand a wide range of biological processes that depend on the specific binding of RNA-binding proteins to RNA. The three-dimensional structures of protein−nucleic acid complexes have provided very valuable insight into these recognition processes. However, these structures are not sufficient to fully understand molecular function, because conformational changes that are observed in many complexes have made structural flexibility and induced fit very common themes in RNA−protein recognition (*1*, *2*).

Numerous studies have shown that RNA−protein recognition often involves large changes in the both the structure and the dynamics of the binding partners, yet surprisingly few studies have been dedicated to the characterization of RNA motions. In fact, to the best of our knowledge, only one study has systematically characterized changes in RNA dynamics that occur upon binding of a ligand at a residue-by-residue level (*3*). We are not aware of any study where RNA motional properties were compared in the presence and absence of a protein, although a few studies have been dedicated to changes in protein dynamics that occur upon RNA binding (*4*, *5*).

Human U1A protein provides a paradigmatic case where conformational rearrangements and biological function are closely connected, as well as a model system to understand the molecular basis of protein−RNA recognition. This protein contains two copies of an RRM-type[1] RNA-binding domain and binds to hairpin II of U1 snRNA or to the 3′ untranslated region (3′UTR) (*6*−*8*) of its own pre-mRNA. The binding site in the 3′UTR consists of two asymmetric internal loops, each with seven unpaired bases with the consensus sequence AUUGCAC that is also contained in the stem−loop of U1 snRNA. When two U1A proteins bind to nearby RNA elements within the 3′UTR, polyadenylation is inhibited by direct interaction between U1A and poly(A) polymerases leading to repression of protein synthesis (*9*). Extensive structural studies by both crystallography (*10*−*12*) and NMR (*8*, *13*, *14*) have shown how U1A protein binds to its RNA targets. These studies demonstrated that the RNA structure is significantly altered upon binding to U1A (*15*), and the structural analysis suggested, but not demonstrated, that protein binding reduces RNA conformational dynamics significantly. Both the protein and RNA-binding regions were flexible in their unbound states but became ordered upon binding (*5*, *13*, *15*). Because this RNA-ordering process is a general feature of RRM−RNA recognition, it is important to investigate these changes in dynamics, in order to determine how these processes contribute to molecular recognition.

Protein dynamics for U1A in the presence and absence of RNA has been investigated by studying ^{15}N and ^{2}H NMR

[†] This work was supported by NIH Grant EB003152.

[‡] Assignments for the free 3′UTR RNA and 3′UTR RNA−U1A protein complex were previously published in ref *15*.

* To whom correspondence should be addressed. E-mail: varani@chem.washington.edu. Telephone: 206-543-7113. Fax: 206-685-8665.

[§] Department of Chemistry.

[||] Department of Biochemistry.

[1] Abbreviations: CSA, chemical shielding anisotropy; HSQC, heteronuclear single-quantum correlation spectroscopy; ppm, parts per million; T_1, longitudinal relaxation time; $T_{1\rho}$, rotating frame relaxation time; RRM, RNA recognition motif; NOE, nuclear Overhauser effect; DI, domain I; DII, domain II; BR, binding region.

10.1021/bi602658x CCC: $37.00
Published on Web 05/01/2007

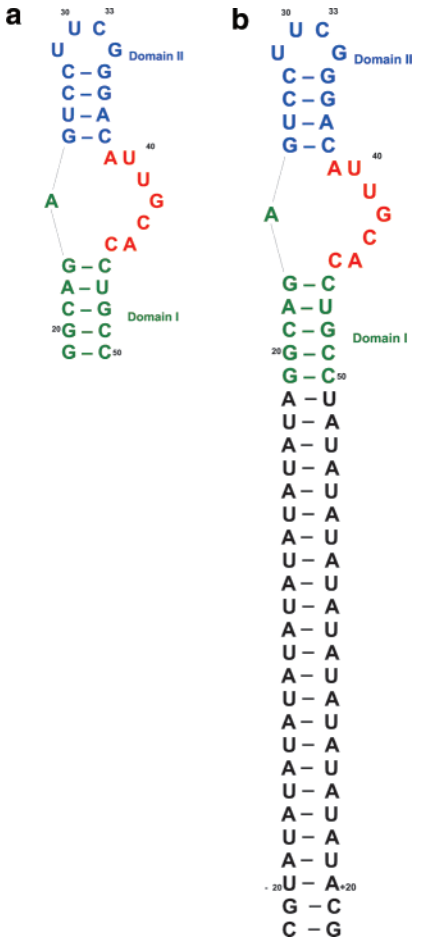

FIGURE 1: (a) Secondary structure of the 30 nucleotide RNA used in this study (*15*, *17*). (b) Secondary structure of the elongated RNA used in this study. Residues in green comprise the lower helical stem and are part of domain I, as defined in the text; residues in blue comprise the upper stem−loop and are part of domain II, while residues in red comprise the U1A binding region.

relaxation (*5*, *16*), and we also recently studied the dynamics of the free 3′UTR RNA (Figure 1a) by ^{13}C NMR relaxation (*17*). In the present work, we have extended our studies to the protein-bound RNA and characterized its dynamics by NMR. We observe that residues in the binding region that experienced complex motions on the picosecond to nanosecond and microsecond to millisecond time scale become rigid upon protein binding. In addition, the upper and lower helical stems experience different motions upon protein binding. By changing the RNA hydrodynamic profile by extending the lower stem through the elongation strategy (*18*) (Figure 1b), we have established that the upper stem−loop experiences these significant domain motions also in the absence of the protein. Thus, it appears that motion of helical domains is common to many RNAs and that protein binding does not always quench these conformational exchanges. Remarkably, several RNA residues that close the internal loop and make important protein contacts exhibit flexibility even in the presence of the protein. Thus, RNA−protein interfaces retain complex dynamic properties that are not uniformly distributed across an interface.

EXPERIMENTAL PROCEDURES

RNA Sample Preparation. The U1A binding site from the 3′UTR is 45 nucleotides long and contains two binding sites for the protein (*8*, *19*); however, a 30-mer construct recapitulates the essential features of the high-affinity recognition by U1A of its RNA target (Figure 1a) (*13*, *15*) although not cooperativity of binding and biological regulation (*8*, *9*). The first two base pairs were altered from CG to GC to allow for efficient initiation of transcription, and the loop was changed to a UUCG tetraloop for added stability and to prevent alternative conformers from forming (*15*); the lower helix was lengthened by one base pair to stabilize the secondary structure. Several RNA samples were prepared enzymatically as described (*20*) using uniformly $^{13}C/^{15}N$-labeled nucleotides obtained from Silantes. Briefly, DNA templates were ordered from IDT with 2′-*O*-methyl groups attached to the last two residues on the 5′ ends in order to inhibit the addition of extra nucleotides (*21*). Relaxation data were collected on three samples, one fully labeled and two partially AC and GU labeled, at concentrations ranging from 0.5 to 0.7 mM. Elongated samples were labeled with either AU or GC and elongated with either GC or AU, respectively, as described (*18*), at concentrations ranging from 0.3 to 0.6 mM. RNA was purified by denaturing PAGE (8 M urea), 20% for the 30-mer construct, and 10% for elongated samples and then electroeluted and ethanol precipitated before microdialysis in 10 mM phosphate buffer with 0.01 mM EDTA at pH 6.0.

Protein Expression. The N-terminal RNA-binding domain of U1A, consisting of amino acids 2−102, was cloned into a pET vector, overexpressed in *Escherichia coli*, and purified as described (*10*, *11*). Cells were grown in M9 minimal media supplemented with $^{15}NH_4Cl$ as the sole source of nitrogen at 37 °C and induced with IPTG at A_{600} ~0.60 and harvested by centrifugation $3^1/_2$ h after induction. Buffer conditions were the same as for the RNA.

NMR Relaxation Experiments. Numerous sites on each nucleotide can be accessed by ^{13}C relaxation studies in nucleic acids, but data analysis is complicated because the chemical shift anisotropies of the base C−H are asymmetric and non-collinear with the CH bond and most spin systems are not isolated. Our previous work on the free U1A RNA (*17*) details the various complications associated with each of these points and their solutions and estimates the errors caused by neglecting the anisotropic character of CSA of the base carbons. We also demonstrated under what conditions accurate relaxation data can be obtained from uniformly ^{13}C labeled RNA. The same methods were employed for this work; the reader is referred to that report (*17*) for experimental details.

Data collection was executed on a Bruker Avance-500 on a TXI triple resonance HCN probe in 99.9% D_2O at 25 °C. T_1, $T_{1\rho}$, and Het-NOE experiments were recorded as a series of 2D NMR spectra, in which the relaxation delay τ is parametrically increased. All experiments were performed in the constant time mode, essentially as described (*22*). The constant time delays were set at different values (7.15 and 12.5 ms, respectively) when recording base or sugar relaxation properties, corresponding to $^{13}C-^{13}C$ coupling constants of approximately 70 Hz (bases) and 40 Hz (sugars) (*23*). The relaxation delay between scans was set to 3.0 s to allow for complete relaxation of the complex. Spectra were recorded with 76−86 complex points in the indirect dimension. The total collection time for each set of experiments was ~40 h for $T_1/T_{1\rho}$ measurements. Selective excitation was

accomplished by application of ^{13}C 180° IBURP-shaped pulses (*24*) during the first INEPT transfer in order to reduce the effects of cross-correlated relaxation (*25*). Each spectral region required collection of an independent set of $T_1/T_{1\rho}$ experiments. Using these selective excitation pulses, ^{13}C relaxation data were separately collected for C8, C6, and C5 in the bases and for C1′ in the sugar with ^{13}C carriers (in ppm) set to 139.5 (C8, C6), 95 (C5), and 89.5 (C1′). Delays of 10, 40, 100, 160, 320, 400, 500, and 800 ms and 4, 8, 12, 16, 24, 32, 48 ms, respectively, were used for T_1 and $T_{1\rho}$ experiments of base and sugar ^{13}C resonances. $T_{1\rho}$ experiments were executed at a spin-lock field of 2 kHz. One of the experimental points at random was repeated for each run to evaluate the reproducibility of the measurements. For heteronuclear NOE measurements, a pair of spectra were recorded, one with initial proton saturation and one without, which was achieved by the application of 120° ^{1}H pulses every 5 ms. Spectra recorded with proton saturation utilized a relaxation delay of 2.5 s followed by a 2.5 s period of saturation. Spectra recorded in the absence of saturation employed a recycle delay of 5 s. Processing and analysis were performed as previously described (*17*).

Data Analysis. As previously described (*17*, *22*, *25*), ^{13}C relaxation studies are complicated by the fact that dipolar couplings between adjacent carbons contribute significantly to the observed relaxation behavior in uniformly labeled samples. This interference increases with the square of the correlation time (i.e., with molecular weight). These studies concluded that in uniformly ^{13}C-enriched sugars the transverse relaxation rates measured are similar to those expected for an isolated spin system, regardless of molecular weight, as long as Hartmann−Hahn magnetization transfer is minimized (*22*, *25*). Concerning longitudinal relaxation, recent studies compared partially (2′, 4′) and uniformly labeled AMP to determine the errors associated with $^{13}C-^{13}C$ auto- and cross-relaxation terms (*26*). Errors in T_1 were 7−14%, 19−25%, and 53−55%, respectively, for correlation times of 5.4, 11.8, and 16.9 ns, similar to those found in an earlier study (*22*). Selective excitation significantly reduces errors from cross-correlation relaxation for small molecules, but this effect cannot be completely suppressed in larger molecules. In addition, errors in the heteronuclear NOE's range from 3% for a 6 ns correlation time to 16% at 16.7 ns. We anticipate the correlation time for this complex to be 10−12 ns based on its molecular weight; furthermore, when studying dynamics of the RNA-bound U1A protein, we observed residual aggregation of the complex, making a quantitative model-free analysis impossible (*5*). For these reasons, a quantitative analysis for the complex using ModelFree would yield results that are likely to be unreliable. Therefore, the data analysis is based primarily on the comparison of relaxation times in the protein-free and -bound RNA. We demonstrated earlier that considerable information on the recognition mechanism can be obtained from this analysis (*5*).

Elongated RNA. The $^{1}H-^{13}C$ resonance intensities for C6H6 and C8H8 cross-peaks were measured by constant time HSQC experiments (*23*) and normalized for each nucleotide (A, C, G, U) independently. HSQC spectra were recorded for fully and partially labeled 30-mer RNA, as well as for the partially labeled elongated constructs (AU and GC). When calculating the average values, only residues for which data were available for both the 30-mer and elongated constructs were considered; resonances known to undergo chemical exchange were also excluded from the analysis. The relevant C−C, C−H, and H−H dipolar and ^{13}C CSA orientations are all directed almost perpendicular to the helix axis of domain I (*18*). This situation should lead to maximum intensities for base C−H resonances in domain I, while base resonances outside of domain I should have smaller intensities. Thus, any increase in intensities can be attributed to internal or collective motions experienced by residues in the upper stem−loop or in the internal loop.

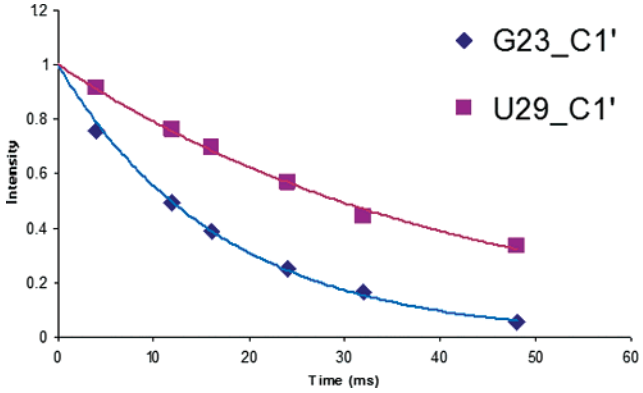

FIGURE 2: Representative relaxation decay curves for residues in domain I (G23, blue diamonds) and domain II (U29, purple squares) recorded at 11.7 T as described in the Experimental Procedures section.

RESULTS

Relaxation of U1A-Bound 3′UTR RNA. Spectral assignments for the U1A-bound 3′UTR RNA were available from previous studies (*15*). In analyzing relaxation data for the complex, we excluded resonances that were partially or completely overlapped with other peaks or that were broadened by conformational exchange and therefore had noisy relaxation curves. As is inevitable, the decrease in $T_{1\rho}$ due to the increase in correlation time upon complex formation resulted in significantly broader peaks and, thus, reduced signal intensity accounting for the increase in error compared to the free RNA (*17*). Errors in relaxation times are about 7% for T_1, 4% for $T_{1\rho}$, and 2% for NOE measurements. The ^{13}C relaxation data were collected twice, first on a fully labeled sample and then on two partially labeled samples where only A's and C's or G's and U's, respectively, were labeled, in order to resolve more resonances. The results reported here represent average values of the two sets of data that differed by less than 5% between them (data not shown). Typical T_1 and $T_{1\rho}$ relaxation curves for some representative residues (G23 and U29) are shown in Figure 2. The observed relaxation times obtained from the analysis of these decay curves and the measured heteronuclear NOE's are shown in Figures 3 and 4.

According to the solution structure of the U1A complex (*13*, *15*), the two double helical regions form essentially ideal A-form helices. Consistent with this structural rigidity, the relaxation data collected on the free RNA show that nearly all of the nucleotides in both helices exhibit similar longitudinal and transverse relaxation as well as heteronuclear NOE values (*17*). Surprisingly, the relaxation data for the bound RNA displays remarkably different relaxation rates

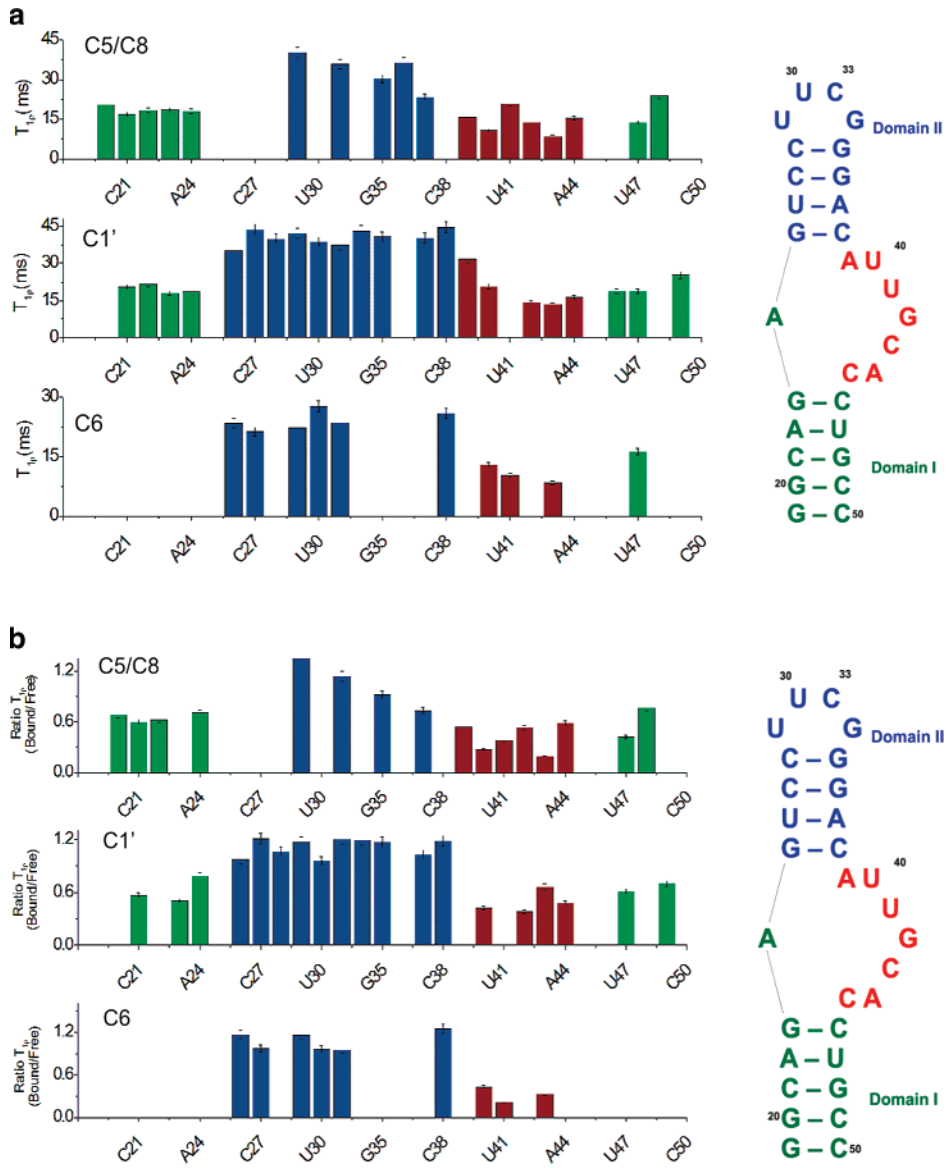

FIGURE 3: (a) ^{13}C $T_{1\rho}$ relaxation times for the RNA bound to U1A protein. (b) Ratio of ^{13}C $T_{1\rho}$ [$T_{1\rho}$(bound)/$T_{1\rho}$(free)]. Top to bottom: C5 and C8 spins; C1′ spins; C6 spins. Residues are color coded as in Figure 1; C6 are shown separately from C5 and C8 because the different CSA lead to large differences in relaxation rates even when motional processes are identical (*17*, *33*). Residues that had noisy relaxation curves, or that were partially or completely overlapped, were excluded from the analysis.

for the upper helical stem compared to the lower stem (Figures 3 and 4). In fact, the upper helical stem has T_1 relaxation times that are about half those observed in the lower stem and $T_{1\rho}$ values that are twice those of the lower stem. These relaxation rates indicate that the upper stem is somehow tumbling much faster than the rest of the molecule.

We did not conduct a model-free analysis of the protein−RNA complex due to its size. As discussed in Experimental Procedures, the results would contain large systematic uncertainty. Instead, we analyzed the data semiquantitatively by comparing relaxation rates in the presence and absence of U1A proteins, as well as their ratios. The variations in relaxation times described in the previous paragraph and shown in Figures 3 and 4 indicate that the complex can be divided between three different regions that exhibit different dynamics, as evidenced most clearly from the average values of the primary relaxation times. Thus, for the purpose of the subsequent analysis and discussion, resonances are grouped into three regions: domain I (D1), corresponding to the lower double helical stem, domain II (DII) and the apical stem−loop, and the binding region (BR), as color coded in Figures 1, 3, and 4.

A first striking observation is that average $T_{1\rho}$'s for both base and sugar resonances are similar for domain I and the binding region, but they are approximately twice as large in domain II. Furthermore, the T_1 values for the base residues in domain II are also approximately half those observed for domain I and for the binding region. For the sugar residues, the T_1's within domain II are about 62−68% of those observed in domain I and in the binding region. However, the NOE's do not vary much between the different regions. The summaries provided in Tables 1 and 2 demonstrate that the apical stem−loop experiences considerably different dynamics compared to the lower helical stem and the binding loop.

Approximate rotational correlation times can be calculated from the relaxation times and heteronuclear NOE's in the absence of significant internal motion (*27*). However, the differences in the observed relaxation rates described above indicate that each region of the RNA moves differently,

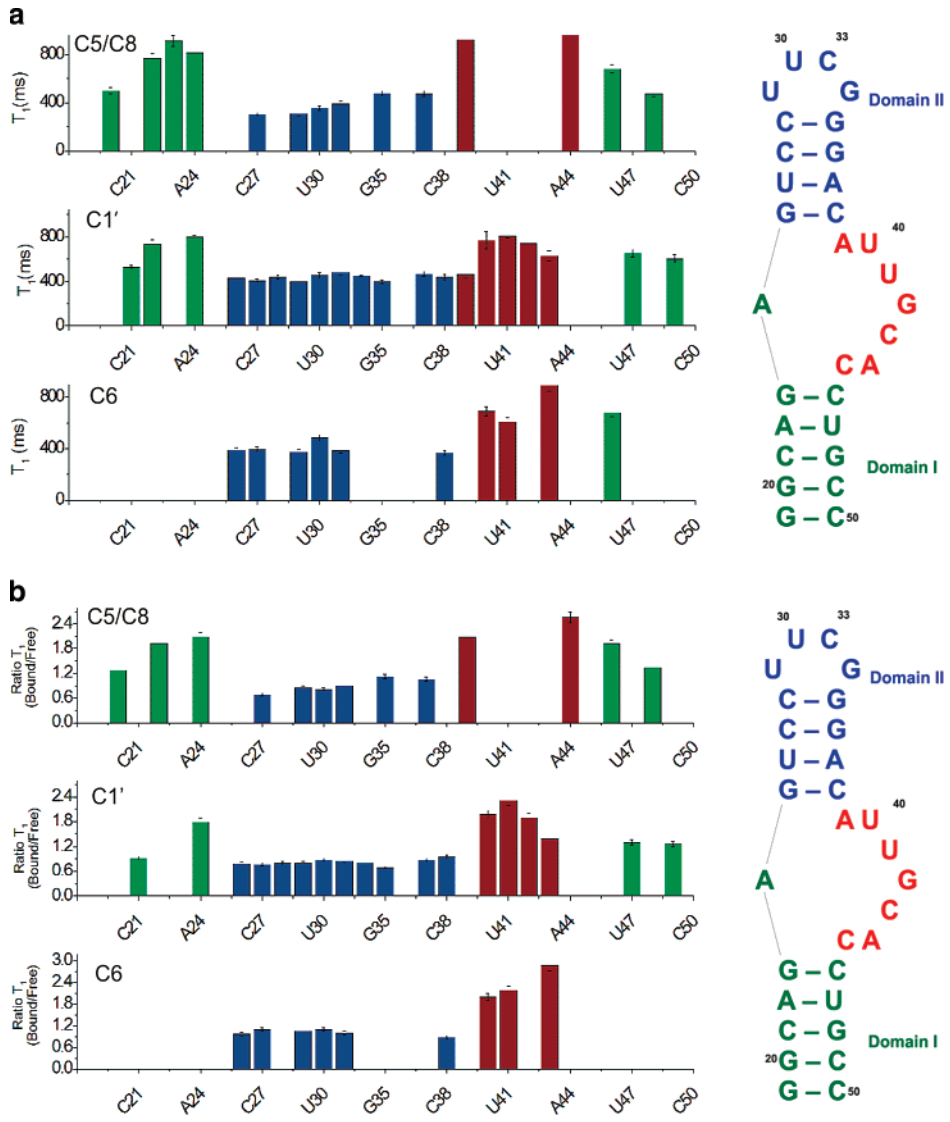

FIGURE 4: (a) ^{13}C T_1 relaxation times for RNA bound to U1A. (b) Ratios of ^{13}C T_1 [T_1(bound)/T_1(free)]. Top to bottom: C5 and C8 spins; C1′ spins; C6 spins. Residues are color coded as in Figure 1; C6 are shown separately from C5 and C8 because the different CSA lead to large differences in relaxation rates even when motional processes are identical (*17*, *33*). Residues that had noisy relaxation curves, or that were partially or completely overlapped, were excluded from the analysis.

Table 1: Average $T_{1\rho}$ Values and Ratios (Bound/Free) for Each Domain (As Defined in Figure 1) for the Protein-Bound U1A RNA

	average $T_{1\rho}$ (ms)			average $T_{1\rho}$ ratio (bound/free)		
	domain I	domain II	binding region	domain I	domain II	binding region
C5/C8	17	33	13	0.62	0.93	0.44
C6	N/A	24	11	N/A	1.08	0.33
C1′	20	40	16	0.56	1.12	0.43

Table 2: Average T_1 Values and Ratios (Bound/Free) for Each Domain (As Defined in Figure 1) for the Protein-Bound U1A RNA

	average T_1 (ms)			average T_1 ratio (bound/free)		
	domain I	domain II	binding region	domain I	domain II	binding region
C5/C8	789	386	N/A	1.95	0.91	N/A
C6	N/A	399	730	N/A	1.02	2.35
C1′	622	450	737	1.45	0.83	1.90

making it very difficult to extract a single correlation time for the entire molecule. It is nonetheless instructive to determine approximate correlational times within each domain extracted from the $T_{1\rho}/T_1$ ratios of the C1′ resonances (*28*). The C1′ relaxation values were used for several reasons. First, although ^{13}C CSA tensors in sugars are not collinear with the $^{13}C-^{1}H$ bond vector, the CSA is small; thus, the C1′ relaxation is dominated by the dipolar interaction to the directly attached proton to a much greater extent than for base resonances. Second, variations in CSA tensor values due to secondary structure are negligible for the sugars, so that accurate motional parameters can be extracted (*17*, *25*, *29*).

The very large differences in relaxation times shown in Figures 3 and 4 are reflected in equivalent striking differences in the apparent correlation times between residues in domains I and II. Residues in domain I have an apparent correlation time of 10.8 ns; an essentially identical value (10.5 ns) is observed for residues in the binding region. These values are as expected for a 20−25 kDa complex and demonstrate that the lower helix and the protein-binding region tumble at the rate expected for this protein−RNA complex. Remark-

ably, the apparent correlation time for domain II is much shorter; at 5.2 ns, it is very close to what we measured for the free RNA (*17*) (6 ns) and would correspond to a molecule with a molecular mass of 9−10 kDa. The similar values observed for both the binding region and domain I indicate that motion within the binding region is quenched upon binding, so that the loop where U1A binds becomes as ordered as domain I in the complex. The large discrepancy between the apparent correlation times of the two domains implies instead that domain II moves as a rigid body with respect to the rest of the RNA with a rate comparable to the motions observed in the free RNA in the absence of U1A protein.

When considering specific residues, only relatively few experience local variations within each domain in the relaxation times reflecting local changes in dynamic properties on the picosecond to nanosecond time scale. This result is very different from those we observed in the free RNA, where very large changes in relaxation times reflecting complex motional processes on a range of time scales were observed across the sequence (*17*). Among residues with relaxation rates deviating from average, G20, G48, and C49 are located near the ends of the molecule and exhibit short longitudinal and/or large transverse relaxation times, implying that they retain considerable flexibility due to base fraying even in the protein-bound form. Residue A39 is the first residue in the heptameric loop that is recognized by the U1A protein, and it is connected to domain II by strong stacking interaction with the stem-closing GC base pair formed by C38 and G25 (Figure 1) and by hydrogen bonds with Arg52 in the protein. Arg52 is a critical residue in U1A and other RRM proteins (*30*); if mutated to Gln, binding is essentially abolished. In the free RNA, the base resonance of A39 was relatively rigid in comparison to the rest of the binding region, because of its structural coupling with the apical stem−loop. In the protein-bound form, the base and anomeric resonances of A39 experience different motions. The base carbons exhibit relaxation times similar to those observed in the binding region, while the ribose experiences much faster motion, similar to those observed in domain II. These results suggest that residue A39 is the hinge point between the two domains.

Changes in Relaxation upon Complex Formation. Further insight into changes in dynamics is provided by directly comparing relaxation times observed for the free and protein-bound forms of the RNA. The average changes for each domain are summarized in Tables 1 and 2. For most residues, changes in relaxation rates upon protein binding simply reflect the increase in size and therefore in correlation time. This is clearly the case for resonances located in the lower helix, domain I. The binding region exhibits the largest amount of conformational flexibility in the free RNA; it also exhibits the greatest degree of change—very large average increases in T_1 (210−290% of the value observed in the free RNA) and large decreases in $T_{1\rho}$ (18−49% of the value for the free RNA). Thus, in this part of the RNA, there is considerable quenching of motions occurring on the picosecond to nanosecond time scale upon protein binding. For domain II (the upper stem−loop), however, the values of both T_1 and $T_{1\rho}$ remain the same, or even decrease slightly (T_1) or slightly increase ($T_{1\rho}$), despite the obvious increase in size of the complex. Thus, motions other than overall tumbling influence the relaxation rates for this part of the structure. Since these changes occur over the entire upper stem and tetraloop without large sequence-dependent variations, it is likely that domain II is moving collectively as a single rigid unit at a rate much faster than expected for the rigid tumbling of a complex of this size.

Table 3: Average Normalized Cross-Peak Intensities of Base Resonances for Each Domain (As Defined in Figure 1)

	domain I	domain II	binding region
3′UTR	0.12	0.13	0.29[a]
elongated 3′UTR	0.13	0.24	0.38[a]

[a] Ignoring residues that undergo chemical exchange (A24, G34, G42, and A44) as determined by Shajani and Varani (*17*).

When examining the molecular origin of these differences, we considered two possibilities: either the motions of domain II become uncoupled from the rest of the molecule as a result of protein binding or the two helices experience different collective motions in the free RNA. In this second case, we would have to assume that these collective domain motions would not be detectable in the free RNA because they occur at a rate comparable to rigid rotational diffusion. The addition of protein would alter the hydrodynamic profile of the RNA, separating the time scale for domain motions from those of rigid body tumbling. In order to distinguish these two possibilities, we adopted the elongation approach (*18*) by adding about 20 Watson−Crick base pairs to domain I (Figure 1b). This way, the global motions of the RNA were slowed down, allowing for the time scale of reorientation of the two helical segments to be separated from global tumbling (*18*). When we compared the two constructs, the chemical shifts of the elongated and nonelongated RNAs were in very good agreement for the nucleotides in common, demonstrating that extension of the lower helix does not change the structure of the protein-binding site and upper helical domain.

In order to analyze the results, we compared the intensities of the ^{13}C resonance in HSQC spectra between the RNA of Figure 1a and elongated molecules in their protein-free form (Figure 1b). Variations in resonance intensities (when chemical exchange is not present) report on the net dynamics of a particular site relative to the applied magnetic field (*18*). By excluding residues that experience chemical exchange from this study, resonances with higher intensities identify residues experiencing internal motions that are faster than overall tumbling. In an earlier study of the relaxation properties of the free RNA, we determined that base residues A24, G34, G42, and A44 experienced conformational exchange; thus, these residues were excluded from this analysis.

In the free RNA, the resonances of the Watson−Crick base pairs in domain I and domain II have very similar normalized average intensities (0.12 and 0.13, respectively). This result is consistent with the rigid tumbling of the two domains without any significant internal motion, with the exception of U30 that is known to experience considerable motion in the UUCG tetraloop structure. Resonances within the binding region (0.29) have higher intensities instead (Table 3). For this region, internal motions faster than overall tumbling occur on a picosecond to nanosecond time scale, consistent with relaxation experiments (*17*), leading to increased apparent intensities due to the sharper lines (Figure 5).

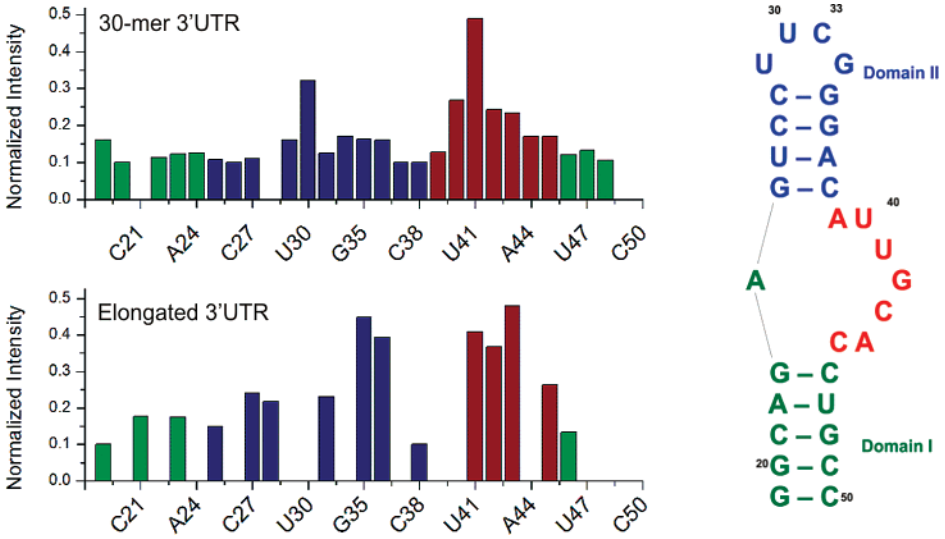

FIGURE 5: Normalized intensities of cross-peaks for base resonances (C8 on purines and C6 on pyrimidines) in $^{13}C-^{1}H$ HSQC spectra reported versus sequence for the original 30 nucleotide construct of the 3′UTR RNA of the U1A mRNA (top), for which the relaxation studies were executed, and for an elongated construct (bottom) of the same RNA containing an additional 20 base pairs within the lower double helical stem. The unlabeled elongated stem is not shown; domain I is shown in green, domain II in blue, and the binding region in red, as in Figure 1. Residues that were partially or completely overlapped were excluded from this analysis.

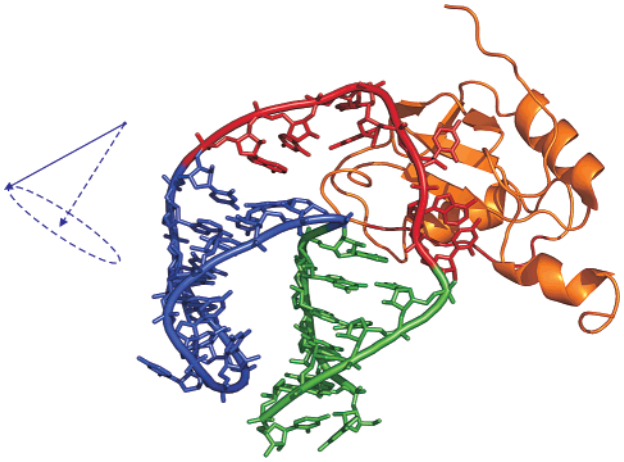

FIGURE 6: Structure of protein-bound 3′UTR. The arrows depict collective domain motions experienced by the upper stem−loop.

In stark contrast, when we analyzed the elongated sample, we observed very large differences in the relative intensities of resonances within Watson−Crick base-paired residues belonging to domain I and domain II (Table 3). The lowest relative intensities occur in the Watson−Crick base pairs of the extended stem (0.13), whereas residues within the upper stem (domain II) have much higher intensities, indicative of collective motion. These results suggest that the upper stem of 3′UTR RNA reorients as a rigid unit, moving at a rate that is very different from the rest of the molecule (Figure 6). Although these motions exist in the free RNA, they are not noticeable in conventional relaxation measurements, because the collective domain motions involving the entire upper stem−loop occur on the same time scale as global tumbling (~6 ns) and are therefore not discernible in the correlation function that determines relaxation measurements. Only when overall tumbling is slowed down, either by protein binding (leading to a correlation time of 10−12 ns) or by elongating the lower stem (leading to a correlation time of 18−19 ns), can domain motions be detected.

Protein Binding Increases the Flexibility of Part of the RNA Structure. As predicted in the structural analysis (*13*, *15*), protein binding quenches the dynamics in the binding region of the free RNA. However, we were also very surprised to observe that protein binding also induces increased motion in the G25-C38 base pair that closes the upper stem−loop and is a key recognition site for U1A protein. Indeed, the relative intensities of the C38 and G25 base resonances in the elongated 3′UTR (Figure 5) are much lower than the intensities of residues in domain II and much closer to those of domain I. Thus, in the free RNA, residue C38 does not experience the same domain motion as the rest of the upper stem, to which it would apparently belong. However, in the protein complex, the relaxation rates of C38 are close to those of the upper stem. Thus, it appears that protein binding decouples the motions of C38 from the lower stem−loop, even if this base and the neighboring G25 and A39 make considerable interactions with the protein, with Arg52 in particular (Figure 7).

DISCUSSION

In the present work, we have studied changes in motion that occur in the RNA-binding site for human U1A protein by measuring ^{13}C NMR relaxation and comparing these data with the results observed for the protein-free RNA (*17*). In studying dynamics of the protein-free RNA, we reported a complex set of motions on multiple time scales and proposed that the protein binds to this RNA by conformational capture. In other words, we proposed that the complex motions we observe allow the RNA to sample multiple conformations, one of which is captured by the protein during formation of the complex. Consistent with this suggestion and with the structure of the RNA free and in the complex (*15*), our results demonstrate that the binding region becomes much more ordered in the presence of the protein. Thus, protein binding quenches motions in the RNA substantially, as we also reported when studying motions in the backbone of U1A protein (*5*).

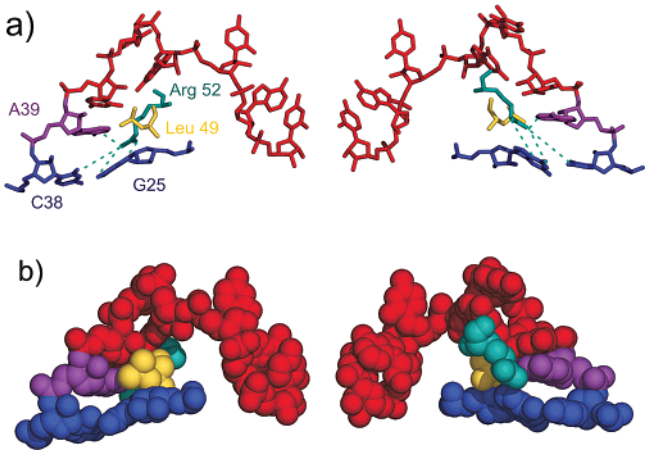

FIGURE 7: (a) Stick and (b) space-filling views of the contacts between the $\beta 2-\beta 3$ loop region of U1A and residues G25 (blue), C38 (blue), and A39 (purple) at the junction between the internal loop and the apical stem−loop of the RNA. The rest of the binding region is shown in red. Leu49 is shown in yellow, and Arg52 is in cyan. The remaining protein side chains are omitted. Hydrogen bonds are indicated by etched lines. Leu49 makes hydrophobic and electrostatic interactions with A39 and G25, while Arg52 hydrogen bonds to G25, C38, and A39.

The most surprising result was the observation of remarkably different relaxation rates for the upper helical stem compared to the rest of the RNA. These results indicate that this part of the structure tumbles as a rigid domain on a rate that is much faster than the rest of the molecule. Thus, while the lower stem and binding regions behave as expected for a 20−25 kDa molecule, the upper helix behaves as a 9−10 kDa molecule. It is possible for the upper stem−loop to become motionally decoupled from the rest of the RNA as a result of protein binding. However, by comparing the intensities of $^{13}C-^{1}H$ cross-peaks in the HSQC spectra for the protein-free RNA in the usual 30 nucleotide construct and in a domain-elongated version, we concluded that the two domains experience domain motions comparable to those observed in the complex even in the absence of U1A protein. These motions are not visible until the hydrodynamic profile is altered, either by elongation or by protein binding, because they occur on the same scale as global rotational diffusion for the complete RNA. Once overall tumbling is slowed down, these domain motions can be detected, implying that they occur at a rate of 5−10 ns. This is now the third RNA (*18*) where entire helices experience considerable and unexpected conformational freedom. Thus, domain motions could very well be a general feature of RNAs that are composed of multiple helices separated by single-stranded linkers.

Structural studies have shown that protein binding alters the overall structure of the binding region, but the two double helical regions remain unperturbed upon protein binding (*11*, *13*, *15*). The exceptions are the base pairs closing the internal loop (G23-C46 and G25-C38) and the U47-A22 base pair which make direct or electrostatic contacts with the protein side chains or phosphodiester backbone. The residues that are adjacent to the binding region are particularly important for U1A recognition, and the protein makes numerous contacts with them (*11*); disrupting the base-paired stem results in decreased levels of protein binding (*19*, *31*). Our results present a model where protein binding alters the profile of the RNA to such an extent that collective domain motion (spanning all residues between the apical loop and the U26-A37 base pair in the free RNA) extends to the C38-G25 base pair and to the ribose of residue A39, the first residue in the single-stranded loop extensively recognized by U1A protein.

It is entirely counterintuitive that these residues would have increased motion in the presence of U1A, since all three make extensive contacts with Arg52 (*13*), and residues A39 and G25 make contacts with Leu49 as well (Figure 7). These amino acids are a particularly important determinant of affinity and specificity: their mutation and disruption of the structure of the stem−loop junction and mutation of A39 all reduce the affinity of U1A substantially (*5*); mutations of G25 and C38 also reduce the stability of the complex (*32*).

When examining why the double helical region opposite the apical stem−loop moves instead as a rigid body with the single-stranded region, we notice that this part of the RNA makes several electrostatic interactions with Arg and Lys side chains located in two critical protein loops and at the C-terminus of the protein (*34*). Mutations of some of these residues (for example, Lys96) lead to considerable reductions in affinity. We propose that formation of these electrostatic contacts dispersed throughout the protein (Lys23, Arg47, and Lys96) may clamp the lower double helical region.

These highly unexpected results demonstrate that, even for residues that form multiple inter- and intramolecular interactions in a protein−RNA complex, residual conformation flexibility is observed or even induced by the protein binding. This observation builds upon our previous result on protein dynamics in the complex, where some protein side chains (though not Leu49) also retained conformational flexibility in the complex, even if backbone motions in the same loop to which Arg52 and Leu49 belong were substantially quenched (*5*). Thus, it appears that protein−RNA interfaces have intrinsic dynamic properties that are important to the recognition mechanism but not easily discernible from the examination of their structure.

In summary, our results reinforce the notion that motions in RNA play a role in molecular recognition by RNA-binding proteins. They also support our earlier hypothesis that the RNA samples multiple conformations in the protein-free form, to allow a protein to select a structure that optimizes intermolecular contacts upon binding (*2*). While these motions are quenched upon formation of the complex, domain motions involving the entire upper stem−loop remain present despite the formation of apparently important hydrogen bonds with key protein residues.

ACKNOWLEDGMENT

We thank Prof. A. M. Al-Hashimi (University of Michigan) and Prof. H. Schwalbe (University of Frankfurt) for sharing preprints of their work. We also thank Priti Deka, Dr. Neil Dobson, and Dr. Karen Lo for help in preparing the U1A protein.

SUPPORTING INFORMATION AVAILABLE

Relaxation data for C5/C8/C6 and C1′ resonances. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Williamson, J. R. (2000) Induced-Fit in RNA-Protein Recognition, *Nat. Struct. Biol. 7*, 834−837.
2. Leulliot, N., and Varani, G. (2001) Current Topics in RNA-Protein Recognition: Control of Specificity and Biological Function through Induced Fit and Conformational Capture, *Biochemistry 40*, 7947−7956.
3. Dayie, K. T., Brodsky, A. S., and Williamson, J. R. (2002) Base Flexibility in HIV-2 TAR RNA Mapped by Solution ^{15}N, ^{13}C NMR Relaxation, *J. Mol. Biol. 317*, 263−278.
4. Deka, P., Paranj, P. K., Perez-Canadillas, J. M., and Varani, G. (2005) Protein and RNA Dynamics Play Key Roles in Determining the Specific Recognition of GU-rich Polydenylation Regulatory Elements by Human Cstf-64 Protein, *J. Mol. Biol. 347*, 719−733.
5. Mittermaier, A., Varani, L., Muhandiram, D. R., Kay, L. E., and Varani, G. (1999) Changes in Sidechain and Backbone Dynamics Identify Determinants of Specificity in RNA Recognition by Human U1A Protein, *J. Mol. Biol. 294*, 967−979.
6. Boelens, W. C., Jansen, E. J. R., van Venrooij, W. J., Stripecke, R., Mattaj, I. W., and Gunderson, S. I. (1993) The Human U1 snRNP-Specific U1A Protein Inhibits Polyadenylation of its Own Pre-mRNA, *Cell 72*, 881−892.
7. Gunderson, S. I., Beyer, K., Martin, G., Keller, W., Boelens, W. C., and Mattaj, I. W. (1994) The Human U1A snRNP Protein Regulates Polyadenylation via a Direct Interaction with Poly(A) Polymerase, *Cell 76*, 531−541.
8. Varani, L., Gunderson, S., Kay, L. E., Neuhaus, D., Mattaj, I., and Varani, G. (2000) The NMR Structure of the 38 kDa RNA-Protein Complex Reveals the Basis for Cooperativity in Inhibition of Polyadenylation by Human U1A Protein, *Nat. Struct. Biol. 7*, 329−335.
9. Gunderson, S. I., Polycarpou-Schwarz, M., and Mattaj, I. W. (1998) U1 snRNP Inhibits Pre-mRNA Polyadenylation Through a Direct Interaction between U1 70K and Poly(A) Polymerase, *Mol. Cell 1*, 255−264.
10. Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) Structure of the RNA-Binding Domain of the U1 Small Nuclear Ribonucleoprotein A, *Nature 348*, 515−520.
11. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) Crystal Structure at 1.92 Å Resolution of the RNA-Binding Domain of the U1A Spliceosomal Protein Complexed with an RNA Hairpin, *Nature 372*, 432−438.
12. Price, S. R., Evans, P. R., and Nagai, K. (1998) Crystal Structure of the Spliceosomal U2B″-U2A′ Protein Complex Bound to a Fragment of U2 Small Nuclear RNA, *Nature 394*, 645−650.
13. Allain, F.-H. T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) Specificity of Ribonucleoprotein Interaction Determined by RNA Folding during Complex Formation, *Nature 380*, 646−650.
14. Avis, J., Allain, F. H.-T., Howe, P. W. A., Varani, G., Neuhaus, D., and Nagai, K. (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*, 398−411.
15. Gubser, C. C., and 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*, 2253−2267.
16. Kranz, J. K., and Hall, K. B. (1999) RNA Recognition by the Human U1A Protein is Mediated by a Network of Local Cooperative Interactions that Create the Optimal Binding Surface, *J. Mol. Biol. 285*, 215−231.
17. Shajani, Z., and Varani, G. (2005) ^{13}C NMR Relaxation Studies of RNA Base and Ribose Nuclei Reveal a Complex Pattern of Motions in the RNA Binding Site for Human U1A Protein, *J. Mol. Biol. 349*, 699−715.
18. Zhang, Q., Sun, X., Watt, E., and Al-Hashimi, H. M. (2006) Resolving Complex Motional Modes that code for RNA Adaptation, *Science 311*, 653−656.
19. van Gelder, C. W. G., Gunderson, S. I., Jansen, E. J. R., Boelens, W. C., Polycarpou-Schwartz, M., Mattaj, I. W., and van Venrooij, W. J. (1993) A Complex Secondary Structure in U1A pre-mRNA that Binds Two Molecules of U1A Protein is Required for Regulation of Polyadenylation, *EMBO J. 12*, 5191−5200.
20. Price, S. R., Oubridge, C., Varani, G., and Nagai, K. (1998) Preparation of RNA-Protein Complexes for X-Ray Crystallography and NMR, in *RNA-Protein Interaction: Practical Approach* (Smith, C., Ed.) pp 37−74, Oxford University Press, Oxford.
21. Kao, C., Zheng, M., and Rudisser, S. (1999) A Simple and Efficient Method to Reduce Nontemplated Nucleotide Addition at the 3′ Ends of RNAs Transcribed by T7 RNA Polymerase, *RNA 5*, 1268−1272.
22. Yamazaki, T., Muhandiram, R., and Kay, L. E. (1994) NMR Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly, *J. Am. Chem. Soc. 116*, 8266−8278.
23. Santoro, J., and King, G. C. (1993) A Constant-Time 2D Overbodenhausen Experiment for Inverse Corelation of Isotopically Enriched Species, *J. Magn. Reson. 97*, 202−207.
24. Geen, H., and Freeman, R. (1991) Band-Selective Radiofrequency Pulses, *J. Magn. Reson. B 93*, 93−141.
25. Boisbouvier, J., Wu, Z., Ono, A., Kainosho, M., and Bax, A. (2003) Rotational Diffusion Tensor of Nucleic Acids from ^{13}C NMR Relaxation, *J. Biomol. NMR 27*, 133−142.
26. Johnson, J., Julien, K., and Hoogstraten, C. (2006) *J. Biomol. NMR 35*, 261−274.
27. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone Dynamics of Proteins As Studied by ^{15}N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease, *Biochemistry 28*, 8972−8979.
28. Mandel, A. M., Akke, M., and Palmer, A. G. I. (1995) Backbone Dynamics of *Escherichia coli* Ribonuclease HI: Correlations with Structure and Function in an Active Enzyme, *J. Mol. Biol. 246*, 144−163.
29. Blad, H., Reiter, N. J., Abildgaard, F., Markley, J. L., and Butcher, S. E. (2005) Dynamics and Metal Ion Binding in the U6 RNA Intramolecular Stem-Loop as Analyzed by NMR, *J. Mol. Biol. 353*, 540−555.
30. Maris, C., Dominguez, C., and Allain, F.-H. T. (2005) The RNA Recognition Motif, a Plastic RNA Binding Platform To Regulate Posttranscriptional Gene Expression, *FEBS J. 272*, 2118−2131.
31. Hall, K. B. (1994) Interaction of RNA Hairpins with the Human U1A N-Terminal RNA Binding Domain, *Biochemistry 33*, 10076−10088.
32. Law, M. J., Rice, A. J., Lin, P., and Laird-Offringa, I. A. (2006) The Role of RNA Structure in the Interaction of U1A Protein with U1 Hairpin II RNA, *RNA 12*, 1168−1178.
33. Duchardt, E., and Schwalbe, H. (2005) Residue Specific Ribose and Nucleobase Dynamics of the cUUCGg RNA Tetraloop Motif by NMR ^{13}C Relaxation, *J. Biomol. NMR 32*, 295−308.
34. Allain, F. H.-T., Howe, P. W. A., Neuhaus, D. and Varani, G. (1997) Structural Basis of the RNA Binding Specificity of Human U1A Protein, *EMBO J. 16*, 5764−5774.

BI602658X