Spatial Orientation and Dynamics of the U1A Proteins in the U1A−UTR Complex†

Caroline Clerte and Kathleen B. Hall*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Saint Louis, Missouri 63110

Received February 16, 2000; Revised Manuscript Received April 20, 2000

ABSTRACT: The human U1A protein contains three distinct domains: the N-terminal RBD1 (amino acids 1−101), the C-terminal RBD2 (amino acids 195−282), and the linker region (amino acids 102−194). The RBD1 domains of two U1A proteins bind specifically to two internal loops in the 3′ untranslated region (3′ UTR) of its own pre-mRNA. Tryptophan fluorescence and fluorescence resonance energy transfer data show that the two RBD2 domains do not interact with any regions of the UTR complex and display an overall tumbling that is uncorrelated from the core of the complex (formed by RBD1−UTR), indicating that the linker regions of the two U1A proteins remain flexible. The two RBD2 domains are separated by an apparent distance greater than 74 Å in the UTR complex. The linker region adjacent to the RBD1 domain (103-ERDRKREKRKPKSQETP-119) is supposedly involved in protein−protein interactions (*12*). A single cysteine, introduced at position 101 or 121 of the U1A protein, was used as a specific attachment site for the fluorophore pair IAEDANS [*N*′-iodoacetyl-*N*′-(1-sulfo-5-*n*-naphthyl)ethylenediamine]/DABMI [4-(dimethylamino)-phenylazophenyl-4′-maleimide]. In the U1A−UTR complex (2:1), the dyes at the 101 position are separated by $\langle R \rangle = \sim 51$ Å, while the dyes at the 121 position are at an apparent distance $\langle R \rangle = \sim 58$ Å. The 101−121 crossed distance on adjacent U1A proteins averages to $\langle R \rangle = 55$ Å. These results suggest that the amino acid sequence 101−121 of the two U1A proteins in the complex are held in proximity to each other in a compact conformation.

RNA binding domains (RBDs)[1] constitute a large family of RNA binding proteins (*1*, *2*). The domains are predicted to adopt $\beta_1\alpha_1\beta_2$-$\beta_3\alpha_2\beta_4$ secondary structures and an α/β sandwich global fold (*3*), in which the four-stranded antiparallel β-sheet packs against the helices. Both β_1 and β_3 strands are conserved in RBDs, as are the amino acid residues that form the hydrophobic core of the domain. The full-length human U1A protein (282 residues) (*4*) contains three distinct domains: the N-terminal RBD1(1−101) (*5*), the C-terminal RBD2 (195−282) (*6*), and a linker of ∼100 amino acids (Figure 1). Only RBD1 is required to bind to its RNA targets (*7*). U1A is one of the three specific proteins associated with the U1 snRNP (*7*) and thus is a component of the eukaryotic pre-mRNA splicing apparatus (*8*). It binds to the stem-loop II of the U1 snRNA with a very high affinity (*9*).

In addition, U1A regulates its production by binding to the 3′ untranslated region (3′ UTR) of its own mRNA to prevent polyadenylation (*10*−*12*). In this complex, the RBD1 domains of two U1A proteins bind to the two asymmetric internal loops (box 1 and box 2) in the 3′ UTR (Figure 2). These binding sites comprise seven unpaired bases (AUUGUAC or AUUGCAC) and a single unpaired base (C or A) on the opposite strand (*13*). The two internal loops are contained in a stem-loop structure and are separated by four base pairs that form stem 2. The seven unpaired bases are located on opposite strands. The asymmetry of these bulges introduces a severe kink into the axis of the RNA (*14*). In addition, a significant global conformational change of the RNA is induced by the binding of the two RBD1 proteins (*15*). Since only the RBD1 domains are required for the specific binding to the UTR RNA, the interactions between these domains and the RNA have been extensively studied and thus are well-described (*16*−*20*).

In contrast, very little is known concerning the function of the rest of the U1A protein (linker + RBD2). The RBD2 domain in the U1 snRNP is a primary target of autoimmune antibodies, suggesting that this domain is accessible in the particle (*21*). However, the function of RBD2 is unknown, for it does not seem to bind to any RNA (*22*). The sequence of the linker region (102−194) of U1A protein is very rich in proline and lysine and highly sensitive to proteases. The secondary structure prediction of the linker by PHD (*23*−*25*) provides no specific structure for this sequence. Also, previous studies have shown that the linker is extremely flexible, giving rise to the uncorrelated tumbling of RBD domains in the free U1A protein (*26*). The 103−119 sequence (ERDRKREKRKPKSQETP) of the linker has been proposed to form a dimer in the UTR complex (*11*). This dimer seems to interact directly with the carboxyl terminal of the poly(A) polymerase, inducing its inhibition (*12*). Since only RBD1 domains are required for binding to the UTR RNA, the rest of the protein (linker + RBD2) remains available for any interactions or functions in the polyadenylation process. Therefore, the investigation of the dynamic and spatial proximity of the two entire U1A proteins in the UTR complex can provide significant structural information critical to understanding the biological function of the U1A protein in the polyadenylation apparatus.

† This work is supported by NIH R01-GM46318 (KBH).

* To whom correspondence should be addressed. Phone: (314) 362-4197. Fax: (314) 362-7183. E-mail: hall@bionmr3.wustl.edu.

[1] Abbreviations: RNA binding domains, RBDs; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; snRNA/P, small nuclear ribonucleic acid, ribonucleoprotein.

10.1021/bi000372k CCC: $19.00
Published on Web 05/24/2000

(1-62) MAVPETRPNHTIYINNLNEKIKKDELKKSLYAIFSQFGQILDILVSRSLKMRGQAFVIFKEV

(63-122) SSATNALRSMQGFPFYDKPMRIQYAKTDSDIIAKMKGTF*VERDRKREKRKPKSQETP*ATK
101 121

(123-182) *KAVQGGGATPVVGAVQGPVPGMPPMTQAPRIMHHMPGQPPYMPPPGMIPPPGLAPGQIPP*

(183-242) *GAMPPQQLMPGQ*MAPAQPLSENPPNHILFLTNLPEETNELMLSMLFNQFPGFKEVRLVPG

(243-282 or 287) RHDIAFVEFDNEVQAGAARDALQGFKITQNNAMKISFAKK
AAGCSLPR

FIGURE 1: Human U1A protein sequence: The charged sequence (103−119) of interest is shown underlined, bold characters. The linker region of U1A protein is written in italic characters, and RBD1 (1−101) and RBD2 (195−282) domains sequences are written in normal characters. Asterisks (*) show the positions 101 and 121 mutated to a cysteine residue for the labeling reactions with the fluorescent probes. The modified carboxyl terminal tail sequence of RBD2 domain is written below the corresponding wild-type sequence.

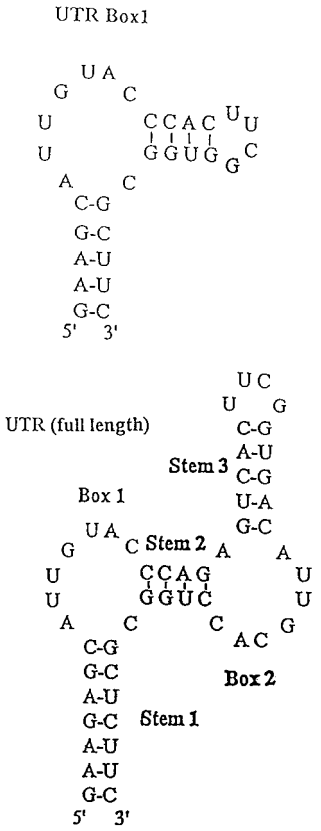


FIGURE 2: Sequences and secondary structures of the RNAs used in the experiments. UTR box 1 contains one binding site from the U1A mRNA 3′ UTR. UTR (full length) contains the two binding sites box 1 and box 2.

In an attempt to define the flexibility and the geometry of the two U1A proteins in the UTR complex, we have carried out steady-state fluorescence and fluorescence resonance energy transfer (FRET) studies on this complex. First, to study the conformational flexibility of the bound U1A proteins, we have used the fluorescence of a single tryptophan introduced at the specific position 248 (β_3 strand) on the β-sheet surface of RBD2. Since the β-sheet platform is nonpolar, it may be a favorable site for protein−protein interactions. Therefore, in addition to its use in probing the dynamic behavior of U1A protein, the tryptophan residue has been used as a reporter of its local environment. Then, the relative distance between the two U1A proteins in the UTR complex has been determined by FRET experiments. For these studies, a single cysteine residue has been introduced at different positions in the U1A protein and used as a specific attachment site for the different fluorophores. We have studied the relative distance and the orientation of the two RBD2 domains in the UTR complex using IAEDANS [*N*′-iodoacetyl-*N*′-(1-sulfo-5-*n*-naphthyl)ethylenediamine]/IAF [5-(iodoacetamido)fluorescein] as a donor/acceptor pair attached at the specific position 283 in the modified C-terminal tail of RBD2 domain. Then, since the 103−119 charged sequence of the linker is proposed to form a dimer in the UTR complex, the relative distance between the two regions has been investigated. For these FRET experiments, a single cysteine residue introduced at either amino acid 101 or 121, which brackets the charged sequence of interest, has been modified by the IAEDANS/DABMI [4-(dimethylamino)-phenylazophenyl-4′-maleimide] fluorophore pair.

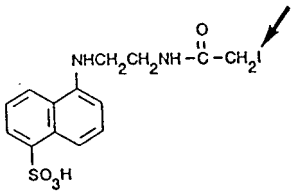
MATERIALS AND METHODS

Site-Directed Mutagenesis. All RBD1 and U1A mutations were generated using standard PCR site-directed mutagenesis methods (*27*). All mutant DNAs were cloned into a plasmid under tac promotor control (*9*) and verified by DNA sequencing. Mutants were transformed into *Escherichia coli* BL21 (*28*) using standard molecular biology techniques (*29*).

Protein Purification. The RBD and U1A mutants were expressed and purified following standard procedures (*9*, *26*). All cysteine protein mutants were purified in the presence of 2-mercaptoethanol (10 mM final concentration). Protein preparation purity was assayed by silver-staining samples subjected to SDS−PAGE. All protein concentrations were determined spectroscopically at 280 nm using extinction coefficients based on the Tyr/Trp/Phe/Cys content.

Covalent Modification of Proteins. (1) Labeling of Single Cysteine RBDs and U1As with 5-[2-(2-Iodoacetamido) ethylamino]-1-naphthalene Sulfonic Acid (IAEDANS) (Figure 3). The modification of single cysteine residues was carried out in storage buffer at pH 7.5 (20 mM Tris, 200 mM NaCl, and 2 mM EDTA). IAEDANS (Molecular Probes, Portland, OR), the fluorescence donor (D), was dissolved to a final concentration of 20 mM in dimethyl formamide (DMF). Protein samples were incubated with 10-fold molar excess of DTT for 30 min at 4 °C. Excess DTT was removed by gel filtration on a Bio-Gel P2 gel column (10 × 1.5 cm) at 4 °C. Fractions of 500 μL were collected. Fractions containing protein were pooled (usually 1.5−2 mL total volume), and protein concentration was determined (typically 50−70 μM). The protein sample was then labeled with constant stirring at room temperature for 2 h in the dark with 10-fold molar excess of IAEDANS. To stop the reaction, 10 mM DTT final concentration was added to the sample and stirred for 30 min at 4 °C. Unbound probe was separated from the labeled protein by gel filtration on a Bio-Gel P2 gel column (30 × 1.5 cm). One milliliter fractions were collected and stored at 4 °C protected from the light.

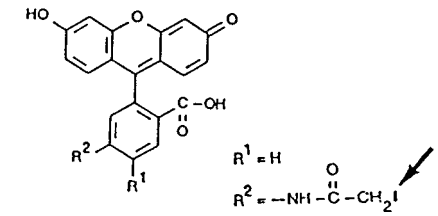
DONOR



1,5-IAEDANS

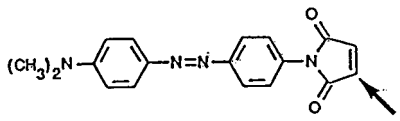
(5-((((2-(iodoacetyl)amino)ethyl)
amino)naphthalene-1-sulfonic acid)

ACCEPTORS



6-IAF

(6-iodoacetamidofluorescein)



DABMI

(4-dimethylaminophenyl
azophenyl-4'-malimide)

FIGURE 3: Structure of the different fluorescent probes used for the FRET experiments (Molecular Probes). Arrows show the reaction site of the probes with the SH group of the cysteine residues introduced by site-directed mutagenesis in the U1A protein.

(2) Labeling of U1A (Cys283) with 5-Iodoacetamidofluorescein (IAF) (Figure 3). Modification was carried out in the storage buffer at pH 8. IAF (Molecular Probes, Portland, OR), the fluorescence acceptor (A), was dissolved to a final concentration of 20 mM in DMF. The U1A (C283) protein was incubated with 10-fold molar excess DTT for 30 min at 4 °C. Excess DTT was removed by gel filtration, 2-fold molar excess of IAF was rapidly added to the protein sample, and the labeling reaction proceeded for 2 h at 4 °C in the dark with constant stirring. The reaction was then stopped by addition of 10 mM DTT final concentration and stirred at 4 °C for 30 min before loading the labeling mix onto a gel filtration column, where free dye was separated from the labeled protein.

(3) Labeling of Single Cysteine RBDs and U1As with 4-Dimethylamino phenylazophenyl-4′-maleimide (DABMI) (Figure 3). The reaction was carried out at pH 8 in storage buffer. DABMI (fluorescence acceptor, A) was dissolved to a final concentration of 20 mM in DMF. Protein samples were incubated with 10-fold molar excess of DTT for 30 min at 4 °C. Excess DTT was removed by gel filtration. Two-fold excess of DABMI was then quickly added to protein samples. The labeling reaction was incubated overnight with constant stirring at 4 °C protected from the light. The reaction was stopped by the addition of 10 mM DTT final concentration and left at 4 °C for 30 min. Unbound dye was separated from the labeled protein by gel filtration.

Labeling efficiency was calculated using the following absorption coefficients (1-cm path length): IAEDANS, 5700 cm^{-1} M^{-1} at 337 nm (Molecular Probes); IAF, 71 000 cm^{-1} M^{-1} at 492 nm; and DABMI, 24 800 cm^{-1} M^{-1} at 460 nm. IAF absorption at 280 nm is equal to 35% of the maximum absorption at 460 nm. The labeling ratio for IAF was determined using the following relation: [IAF]/[protein] = $[A_{(460\ nm)}/\epsilon_{\ IAF\ (460\ nm)}]/[A_{(280\ nm)} - (0.35A_{(492\ nm)})/\epsilon_{\ protein\ (280\ nm)}]$. After extensive dialyses of the labeled proteins against the storage buffer were performed, the labeling ratios did not change and were between 0.8 and 1. The purity of each sample was assayed by silver staining samples subjected to SDS−PAGE, which was also used to estimate protein concentration by comparing the labeled protein with unlabeled protein of known concentration.

RNA Synthesis. RNA molecules were enzymatically synthesized in vitro from DNA oligonucleotides using SP6 RNA polymerase as described previously (*9*, *30*). The secondary structures of the different RNA molecules used are shown in Figure 2.

Steady-State Fluorescence Measurements. *(1) Tryptophan Fluorescence*. Steady-state fluorescence spectra of U1A F248W were recorded from 310 to 380 nm on a SLM 81000 spectrofluorometer with an excitation wavelength of 300 nm (bandwidth of 2 nm) and an emission bandwidth of 8 nm. The temperature was kept constant at 25 °C, while the viscosity of the solutions was varied by the addition of sterile 50% (V/V) solution of glycerol. For the U1A F248W−UTR box 1 RNA complexes, samples were prepared by dilution of the protein stock solution to a final concentration of 5 μM into 20 mM Tris (pH 8), 200 mM NaCl, and 1 mM $MgCl_2$. For U1A F248W−UTR (full length) RNA complexes, samples were prepared by dilution of the protein stock solution to a final concentration of 10 μM into the same buffer. Small aliquots of RNA stock solution were added to the stirred protein solutions to a final concentration of 5 μM, and thus a stoichiometry of 1:1 for the U1A F248W−UTR box 1 RNA complex or a stoichiometry of 2:1 for the U1A F248W−UTR RNA complex was obtained. The samples were allowed to equilibrate for 15 min before each emission spectrum data acquisition. Emission spectra of each solution were measured before and after addition of RNA. Fluorescence intensities were corrected by subtraction of a buffer blank, for inner filter effect and dilution (*31*).

(2) Fluorescence Resonance Energy Transfer (FRET). The fluorescence energy transfer experiments were performed with an excitation wavelength of 350 and 360 nm (bandwidth of 2 nm) for IAEDANS/IAF pair and IAEDANS/DABMI pair respectively and an emission bandwidth of 4 nm. All the experiments were performed at 25 °C. U1A(C283) protein labeled with IAEDANS (D) and with IAF (A) samples were diluted to a final concentration of 0.5 μM in 20 mM Tris (pH 8), 200 mM NaCl, and 1 mM $MgCl_2$ buffer (buffer F). The final concentrations for proteins labeled with

IAEDANS (D) and DABMI (A) were 2 μM in buffer F. Small aliquots of RNA stock solution were added to stirred protein solutions (D + A) to obtain a final concentration of 0.5 μM or 2 μM RNA. Samples were allowed to equilibrate for 15 min before data acquisition. Fluorescence intensities were corrected by subtraction of a buffer blank.

The efficiency (E) of fluorescence energy transfer is related to the donor−acceptor distance R according to

$$E = R_0^6/(R_0^6 + R^6) \quad (1)$$

where R_0 is the Förster critical distance at which the energy transfer efficiency is 50%. R_0 is defined by the following relation:

$$R_0 = 9790(J\kappa^2\phi_D n^{-4})^{1/6}\ \text{Å} \quad (2)$$

where J is the spectral overlap integral of the dyes, ϕ_D is the quantum yield of the donor in the absence of the acceptor, n is the refraction index of the medium, and κ^2 is the orientation factor of the transition dipole moments. For total randomization of the relative donor−acceptor orientation, κ^2 is assumed to be equal to 2/3. (The low anisotropy values measured for proteins labeled with IAEDANS and IAF indicate that this is a valid approximation here) (*32*−*35*).

For every FRET experiment, the following control experiments were done: (i) Addition of the UTR RNA to D sample alone. This control allows verification that the fluorescence of the donor does not change upon binding to the RNA. (ii) Mixture of the D and A in the absence of RNA demonstrates that no energy transfer occurs between the two fluorophores when the U1A−UTR complex is not formed. In addition, these controls verify that the total absorption at the excitation wavelength used is minimal and thus that significant inner filter effects do not occur during the experiment.

The efficiency (E) of fluorescence energy transfer from D to A was calculated from the decrease (or quenching) of the fluorescence intensity of D. The samples are a mixture of D/D (25%), A/A (25%), and D/A (50%) UTR complexes. For the FRET experiments done with the IAEDANS/DABMI pair, A (DABMI) is not fluorescent. Therefore, the fluorescence signal comes from only the D/D and D/A complexes. Thus one-third of the total fluorescence signal comes from the D/D complexes, and so the true fluorescence signal coming from only D/A complexes is equal to

$$F_{(D/A)} = F_{(DA\ total)} - [(1/3)F_{(DA\ total)}] \quad (3)$$

where $F_{(DA\ total)}$ is the integration from 370 to 600 nm of the fluorescence intensities for the D + A samples.

The true fluorescence signal for the FRET experiment is equal to

$$F_{(complex)} = F_{(FRET)} - [(1/3)F_{(DA\ total)}] \quad (4)$$

where $F_{(FRET)}$ is the integration from 370 to 600 nm of the fluorescence intensities measured after addition of the UTR RNA to the D + A sample.

The efficiency (E) was calculated using the following relation:

$$E = 1 - [F_{(complex)}/F_{(D/A)}] \quad (5)$$

where $F_{(complex)}$ is the integration from 370 to 600 nm of the true (or corrected) fluorescence intensities for D in the presence of A in the UTR complex, and $F_{(D/A)}$ is the integration of the true (or corrected) fluorescence intensities for D in the presence of A without RNA present.

(3) Steady-State Anisotropy Measurements. Anisotropies were measured using the L-format method. Before each anisotropy data acquisition, samples were allowed to equilibrate for 30 min. Each anisotropy point is the average value of at least 40 measurements with a standard deviation <0.002. For each sample, the anisotropy was measured at the emission maximum and fit to the Perrin equation:

$$1/r = 1/r(0)\ [\ 1 + K(T/\eta)] \quad (6)$$

where $K = \kappa_B\tau/V$ with κ_B the Boltzmann constant, τ is the excited-state lifetime, V is the hydrodynamic volume of the molecule, η is the viscosity of the solution, and $r(0)$ is the initial anisotropy (*32*).

The anisotropy measurements for the U1AF28W−UTR complexes were recorded at the maximum emission wavelength (340 nm) with an excitation wavelength of 300 nm (bandwidth of 4 nm) and an 8-nm emission bandwidth. The experimental conditions were those used for the fluorescence spectra measurements previously described. For IAEDANS measurements, the excitation wavelength was 365 nm with a bandwidth of 4 nm, and the emission wavelength was 500 nm with a bandwidth of 8 nm. The anisotropies for each labeled protein were measured as a function of temperature over the range 8−30 °C [previous CD experiments show that the protein secondary structure is not disrupted over this range (*26*)]. Samples were prepared by dilution of the protein stock solution to a final concentration of 5 μM in 20 mM Tris (pH 8), 200 mM NaCl, and 1 mM $MgCl_2$.

RESULTS

Orientation and Position of the Two RBD2 Domains in the UTR Complex. (1) U1A F248W−UTR Complexes. Tryptophan is a very useful probe to report on its local environment and also to determine the solution properties of the proteins. Therefore site-directed mutagenesis was used to introduce a tryptophan residue in the RBD2 domain of the U1A protein. The tryptophan at position 248 is solvent-exposed on the β-sheet surface of RBD2. The hydrophobicity of this surface and the local concentration of RBD2 in the UTR complex should favor protein−protein interactions through the β-sheet surface. The comparison between the free and the bound protein fluorescence should provide information on the RBD2 domain dynamics and on the local environment of the probe. Both U1AF248W protein and RNAs were characterized for subsequent use in fluorescence experiments. The CD spectrum of the protein shows that the substitution of Phe248 with Trp does not perturb the secondary structure (*26*). In addition, nitrocellulose filter binding assays show that the dissociation constant (K_D) obtained for U1A F248W−RNA hairpin complex is equal to that of RBD1(wt)−RNA hairpin [2×10^{-11} M ($\pm 1 \times 10^{-11}$ M) in 10 mM sodium cacodylate, 150 mM NaCl, 1 mM $MgCl_2$, pH 7 and 22 °C; data not shown]. This result demonstrates that the mutation made in RBD2 does not disturb the binding function of the protein mutant. The binding efficiency (capability) of the UTR box 1 (wt) (1:1

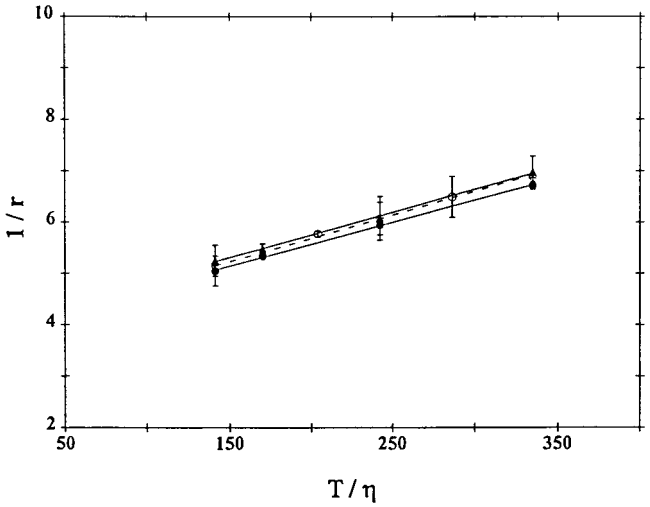


FIGURE 4: Perrin plots of the viscosity-dependent fluorescence anisotropy of the U1A F248W protein free and bound to the UTR box 1 or the UTR (full length) RNA. (○) The free U1A F248W protein; (●) the U1A F248W protein bound to the UTR (full length) RNA forming the U1A F248W−UTR complex; (▲) the U1A F248W protein bound to the UTR box 1 RNA forming the U1A F248W−UTR box 1 complex.

complex) and UTR (wt) (2:1 complex) RNAs was also characterized by monitoring the fluorescence quenching of a Trp at position 56 in RBD1 F56W protein mutant (data not shown). RBD1 F56W and RBD2 F248W have Trp introduced at analogous sites on the β-sheet surface. In both complexes, there is complete quenching of Trp 56, indicating a 100% binding efficiency. This tight binding and high specificity allow us to form very stable complexes for the following experiments.

(2) Steady-State Fluorescence and Anisotropy. The fluorescence spectrum (300-nm excitation) of the U1A F248W protein is identical to the fluorescence spectra of U1A F248W bound to UTR box 1 (1:1 complex) or bound to the UTR (full length) RNA (2:1 complex) (data not shown). The local environment of Trp248 in RBD2 does not change when the protein is free or bound to the UTR RNA. The Trp fluorescence insensitivity here indicates that the β-sheet of RBD2 is not involved in any interactions in the UTR complex. The steady-state anisotropies for the free U1A F248W protein, the U1A F248W−UTR box 1 and U1A F248W−UTR (full length) complexes were measured at the fluorescence maximum (340 nm) using a 300-nm excitation wavelength as a function of solution viscosity. The Perrin plots ($1/r$ versus T/η) for the free protein and both complexes are shown in Figure 4. The inverse anisotropy ($1/r$) values indicate basically no changes in the depolarization between the free protein and the protein bound to the UTR RNAs. This result indicates that RBD2 domain motion is uncorrelated from the slow overall tumbling of the large core of the complex formed by the RBD1 domains bound to the UTR. Previous results demonstrated that the 100-amino acid linker region between RBD1 and RBD2 of the U1A protein is flexible in the free protein, such that the tumbling of the two RBD domains is uncorrelated (*26*). It is significant that this result does not change when the U1A protein is bound to the UTR RNAs.

Relative Orientation and Average Distance between the Two RBD2 Domains in the U1A−UTR Complex. The previous fluorescence results indicate that there are no interactions between RBD2 domains and any regions of the protein or RNA in the UTR complex. Also, the two RBD2 domains are apparently free to tumble (independently from the core of the complex). But the relative position of these two domains in the complex remains unknown. Information about distances on the order of 10−100 Å can be provided by the FRET method. In an attempt to determine the relative proximity of the two RBD2 domains in the UTR complex, U1A (C283) was labeled with donor and acceptor fluorophores for FRET experiments. IAEDANS (donor)/fluorescein (acceptor) is the pair of fluorophores used here and is characterized by a Förster critical distance (R_0) equal to 45 Å, where R_0 represents the distance between the two probes at which the efficiency of energy transfer is 50%.

(1) Specificity and Effect of Labeling of U1A (C283). Site-directed mutagenesis was used to introduce a cysteine residue suitable for subsequent labeling with pairs of fluorescent thiol-reactive probes. The cysteine residue is introduced at position 283 in the extended carboxyl terminal tail of RBD2 (Figure 1). The choice of this site presents two advantages. First, the addition of the fluorescent dyes at the C-terminal tail will minimize possible perturbations of the RBD2 domain structure and increase the probability that covalently bound probes will rotate more freely. Second, the cysteine residue will likely be accessible for the reaction with the dye. The labeling efficiency was excellent, with values of 85 and 95% for IAEDANS and IAF, respectively. Moreover, control experiments using U1A F248W protein without C283 in the presence of the dye demonstrated that no nonspecific association of the dyes with the protein occurs during the reaction. (For future reference, attempts to introduce tetramethylrhodamine were not successful, since it stuck to the protein and could not be separated away). After the labeling reaction, the proteins were characterized by silver-stained SDS−PAGE and nitrocellulose filter binding assays using wild-type RNA hairpin. The binding ability of the UTR RNA was also characterized by monitoring the quenching of the Trp56 in RBD1 protein, as described previously.

The emission spectra of each free labeled U1A protein indicate that the probes are solvated. The maximum of the fluorescence is measured at 490 and 520 nm for U1A (C283) labeled with IAEDANS (U1A (C283)-D) and U1A (C283) labeled with IAF (U1A (C283)-A), respectively (data not shown). The fluorescence anisotropies of each labeled protein were used to estimate the free segmental motion of each dye and thus to determine whether the orientation factor (κ^2) in eq 1 can be replaced with its dynamically averaged value of 2/3. The low anisotropy values obtained for each of the labeled proteins suggest that the orientation factor is substantially dynamically averaged on the fluorescence timescale (nanoseconds) (*33*, *35*). The anisotropy values were between 0.05 and 0.06 for all the free proteins labeled with IAEDANS. The anisotropy increases slightly up to 0.07−0.09 when these proteins labeled with IAEDANS are bound in the UTR complex. It is important to note that the distance calculated will be an apparent distance, since it is a solution average over an undefined distribution of distances. However, given the relatively high concentration of RBD2 in the complex, the two domains could approach each other close enough to be detected by FRET.

(2) Energy Transfer Results. The fluorescence spectra are measured from 370 to 650 nm using a 360-nm excitation wavelength. The spectrum of the donor (U1A (C283)-IAEDANS) alone is identical to the spectrum of the donor in the presence of the UTR RNA, indicating that the RNA concentrations used are low enough to avoid any inner filter effects. The acceptor (U1A (C283)-IAF) alone excited at 360 nm (IAF minimum absorption) still gives a significant fluorescence signal. However, the spectrum of the donor in the presence of acceptor (donor/acceptor spectrum) is identical to the calculated [donor alone + acceptor alone] spectrum. Therefore, no energy transfer occurs in the absence of the UTR RNA.

Unfortunately, in the presence of UTR RNA, no evidence of energy transfer is observed between the donor and the acceptor. The donor/acceptor spectrum is identical to the donor/acceptor/UTR spectrum. The lowest efficiency of energy transfer that we can accurately detect (or observe) is 0.05 and corresponds to $\langle R \rangle = 74$ Å for this pair of fluorophores. Therefore, the absence of energy transfer between these two dyes indicates that the two RBD2 domains are separated by a distance greater than 74 Å in the U1A−UTR complex.

Relative Orientation and Average Distance between the Two (101−121) Linker Sequences in the UTR Complex. The interactions between the RBD1 domain and the RNA are known and well-described (*16*−*20*). On the basis of these structural data, several models have been proposed for the interactions and thus the orientation of the two RBD1 domains in the UTR complex (*15*, *36*). Given the proximity of the two binding sites in the UTR RNA, some regions of the two U1A proteins must become juxtaposed either transiently or statically. The regions that are close to the core include the α_3 helices and the adjacent linker amino acids. Also, according to the results of Gunderson et al. (1997), the two 103−119 sequences of the linker form a dimer which interacts directly with the C-terminal domain of the poly(A) polymerase inducing its inhibition. In an attempt to investigate the geometry of the complex close to its core and to probe the starting point of the flexibility of the linker region of the U1A protein, FRET experiments were done using two fluorescent dyes introduced at positions 101 and 121 of either RBD1 protein or the entire U1A protein (see U1A sequence in Figure 1). The 101 and 121 positions have been selected for mutation because they bracket the charged sequence of interest. The fluorescent dye pair used in these experiments is IAEDANS [donor(D)] and DABMI [acceptor(A)], which is characterized by an R_0 of 40 Å.

(1) Specificity and Effect of the Labeling of RBD F101C, U1A F101C, U1A T121C. For each protein, a single cysteine residue was introduced by site-directed mutagenesis for the subsequent reaction with the thiol-reactive dyes. Phenylalanine 101 and threonine 121 were selected for substitution: the 101 position is just after the α3 helix that ends at amino acid 98; the 121 position is in the linker region after the 103−119 charged sequence (Figure 1). The labeled ratios for every mutant with IAEDANS and DABMI were between 0.75 and 1. Both labeled proteins and UTR RNA were characterized before FRET experiments. As previously described, the binding ability of the UTR RNA was characterized; protein purity was assayed by silver-stained SDS PAGE, and nitrocellulose filter binding assays showed that every labeled protein bound the wild-type RNA hairpin with normal affinity (data not shown). The proteins labeled with IAEDANS (protein-D) have been also characterized by their fluorescence properties. The emission spectrum of each protein-D shows a maximum of fluorescence at 490 nm and therefore suggests that the probe is solvated. The fluorescence anisotropy of each protein-D is measured at the emission maximum using a 365-nm excitation wavelength. The low anisotropies observed imply that the covalently bound IAEDANS undergoes fast segmental local motion (data not shown). On the other hand, the acceptor (DABMI) is not fluorescent. Therefore, only the absorption spectra of each protein-A were recorded; the maximum of absorption was observed at 460 nm for all protein-A species (data not shown).

(2) Energy Transfer Results. The fluorescence resonance energy transfer experiments were recorded from 370 to 600 nm using a 350-nm excitation. For every FRET experiment, the following control experiments were done. (i) Addition of the UTR RNA to protein-D decreased the total fluorescence by less than 4%, showing that (a) there is no interaction between the fluorophore and the RNA upon protein binding, (b) there is no quenching of the donor fluorescence by proximal protein, and (c) the concentrations used are low enough to keep the total absorption at 350 nm small and thus avoid significant inner filter effects. (ii) The fluorescence spectra of protein-D alone and in the presence of protein-A are the same, indicating that there is no energy transfer between D and A in the absence of the UTR RNA (data not shown). Since DABMI is not fluorescent, the efficiency of energy transfer is calculated by the measure of the decrease (or quenching) of the donor fluorescence. E was calculated using the following relation:

$$E = 1 - [F_{(\text{complex})}/F_{(\text{D/A})}]$$

where $F_{(\text{complex})}$ is the integration from 370 to 600 nm of the true (or corrected) fluorescence intensities for D in the presence of A in the UTR complex, and $F_{(\text{D/A})}$ is the integration of the true (or corrected) fluorescence intensities for D in the presence of A without RNA present (cf. Materials and Methods).

The efficiencies and the corresponding apparent distances $\langle R \rangle$ of every complex are given in Table 1. For the case in which both donor and acceptor are attached at the amino acid 101, the distance between them is ∼50 Å, based on the efficiency of ∼20% (Figure 5). This value is independent of whether the protein is RBD1 or U1A. The fact that the presence of the rest of the U1A protein (the linker plus RBD2) does not modify the relative position between these two probes suggests that the 101 position is physically close to the core of the complex and so is subject to constraints on its orientation. A significantly larger $\langle R \rangle$ of 58.7 Å is observed between the two dyes attached at position 121 of U1A.

The distance between the dyes attached at amino acid 101 on one U1A protein and amino acid 121 on the other adjacent U1A protein in the U1A−UTR (2:1) complex, the "crossed distance", can also be measured. The apparent distances calculated for the RBD101-D−U1A121-A−UTR and the corresponding RBD101-A−U1A121-D−UTR complexes (where only one protein is full length; the other one is

Table 1: Energy Transfer for IAEDANS/DABMI (Donor/Acceptor) Pair in the UTR Complex[a]

donor site	acceptor site	E^b	$\langle R\rangle_{(2/3)}$(Å)[c]
RBD 101	RBD 101	0.22 (± 0.02)[d]	49.4 (± 0.9)[e]
U1A 101	U1A 101	0.18 (± 0.02)	51.5 (± 1)
U1A 121	U1A 121	0.09 (± 0.01)	58.7 (± 1)
RBD 121	U1A 101	0.17 (± 0.01)	51.7 (± 0.3)
RBD 101	U1A 121	0.168 (± 0.01)	52.2 (± 0.2)
U1A 121	U1A 101	0.18 (± 0.02)	51.4 (± 1.4)
U1A 101	U1A 121	0.078 (± 0.01)	60.3 (± 0.6)

[a] The RBD1 (1−121) protein was labeled only with the energy donor IEADANS. RBD1 (1−101), U1A (F101C), and U1A (T121C) proteins were labeled either with the energy donor (IAEDANS) or with the energy acceptor DABMI. All measurements were carried out at 25 °C in 20 mM Tris (pH 8), 200 mM NaCl, and 1 mM $MgCl_2$. [protein-D] = [protein-A] = [UTR RNA] = 2 μM. [b] E is the efficiency of fluorescence energy transfer. [c] $\langle R\rangle_{(2/3)}$ is the apparent distance calculated from the measured efficiency of fluorescence energy transfer, assuming the orientation factor $\kappa^2 = 2/3$ for the determination of the R_0 value. Under this assumption, the R_0 for the IAEDANS/DABMI fluorophores pair is equal to 40 Å and $E = R_0^6/(R_0^6 + R^6)$. [d,e] The values reported for the efficiencies and the distances represent the arithmetic mean of multiple measurements (2 −4). The errors reflect the deviations of the maximum and minimum values from the means.

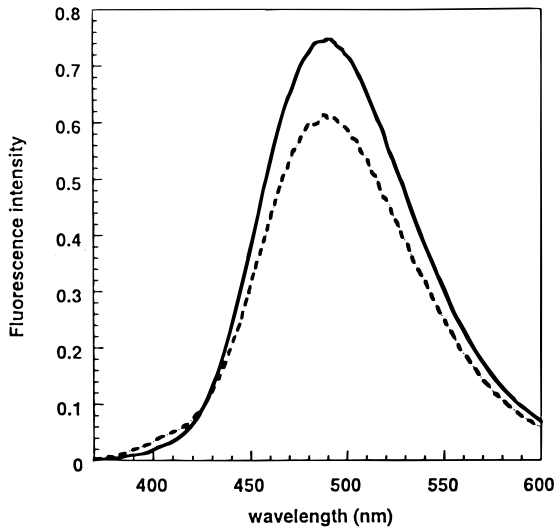


FIGURE 5: Steady-state FRET experiments for the U1A protein labeled at the position 101 with IAEDANS (U1A-D) and DABMI (U1A-A) at 25 °C. (—) Emission spectrum of the U1A-D/U1A-A mixture. (- - -) Emission spectrum of the U1A−UTR (full length) complex (U1A-D/U1A-A/UTR).

truncated at 101 are 52.2 ± 0.2 and 51.8 ± 0.3 Å, respectively. The $\langle R\rangle$ for the U1A101-D−U1A121-A−UTR and the corresponding U1A101-A−U1A121-D−UTR complexes (where now both proteins are full-length) are, however, significantly different: 51.4 ± 1 and 60.3 ± 0.5 Å, respectively. These results show that the location of the donor and the acceptor in the U1A−UTR complex influences their relative orientation, and several physical interpretations are feasible. One is that the local segmental motion of the dyes depends on their environment. Since the measured anisotropy of IAEDANS indicates that it undergoes fast segmental motion at both positions in U1A, it seems likely that DABMI is more rigidly attached to the protein. Indeed, the reaction with DABMI occurs through the maleimide group, and it has a much shorter tether as compared to that of IAEDANS (Figure 3). Thus, the theoretical assumption ($\kappa^2 = 2/3$) in the analysis gives rise to an underestimate for one direction and an overestimate for the reverse direction, resulting in lower and upper bounds for the actual distance. Another scenario is that in the presence of the complete U1A, the relative geometries of the two fluorophores are altered, perhaps through steric constraints that are specific to each dye. The average distance of 55 ± 5 Å indicates the uncertainty in this value.

The linker region of the U1A protein has been predicted by PHD (*25*) to have no particular secondary structure (data not shown). However, the sequence between 101 and 121 is very peculiar by the fact that it is rich in lysine and proline and it has a (− + − +++ − +++) charged sequence from 103 to 112. In an attempt to interpret the FRET data in terms of protein structure, the experimental results are compared to the two following extreme models: (i) the fully extended and rigid 20 amino acid long peptide and (ii) the totally flexible random coil peptide. For the first model, the two 101−121 regions can be represented as two parallel cylinders of diameter equal to 24 Å (lysine residue side chain is equal to 12 Å) and a length of 76 Å (assuming that the 20 amino acid sequence is fully extended with a $C_\alpha-C_\alpha$ distance of 3.8 Å). In this model, the maximum crossed distance 101−121 will be equal to 90 Å. Clearly, the FRET data do not support this model.

When only one protein contains the sequence 101−121 in the UTR complex, there is no opportunity for interstrand protein−protein interactions. Experimentally, this is the case for RBD101-D−U1A121-A−UTR and the corresponding RBD101-A−U1A121-D−UTR complexes that may be good experimental representations of the flexible peptide model. Since the apparent distance calculated for these complexes (∼52 Å) is not very different from the average distance calculated for the U1A−UTR complexes (∼55 Å), the flexible and unconstrained model might be the most accurate to describe the U1A−UTR complex. Thus, the apparent distances obtained from the FRET experiments suggest that the 101−121 region of the linker is unstructured (consistent with the PHD prediction) but that these two adjacent sequences might interact together to compact and keep the two 121 positions relatively close.

DISCUSSION

Structural and dynamic properties of the two U1A proteins in the UTR complex, investigated here using steady-state polarization of fluorescence and steady-state resonance energy transfer, have provided important information for the understanding of this system.

RBD2 Domains in the UTR Complex. Our data show that the linker region of the U1A protein remains extremely flexible in the UTR complex giving rise to a totally uncorrelated tumbling of the two RBD2 domains from the core of the complex. A large distance distribution between the two RBD2 domains results from this freedom of motion, and therefore, the probability that these two domains approach each other in the fluorescence energy transfer time scale (nanoseconds) is small. We found also that the nonpolar platform formed by the β-sheet of the RBD2 domain is not involved in any interactions with the rest of the protein or with the RNA. The fact that in the UTR complex the U1A proteins remain highly flexible and the two RBD2 domains have so much freedom of motion is undoubtedly a sine qua

non for the conformation changes and adaptations necessary to perform an efficient assembly of the polyadenylation or splicing machinery.

The RBD2 domain adopts an α/β global fold structure very similar to that of RBD1 (*6*). Despite its structural similarity to RBD1 and the presence of the two highly conserved RNP-1 and RNP-2 sequences in the β-sheet, RBD2 domain does not show any capability to bind RNAs (*22*). This domain is specifically recognized by autoimmune antibodies in the U1 snRNP indicating that it is exposed in this particle (*21*). Our results indicate that again both RBD2 domains are exposed in the UTR complex. The experiments performed here do not provide any information about the function of the RBD2 domain in the UTR complex. But since our experiments are done in vitro using only purified U1A proteins and the UTR RNA (containing only the two binding sites box 1 and box 2), these conditions do not account for other proteins and RNA sites (e.g., the polyadenylation site) that are required and involved in vivo in the polyadenylation process. We speculate that RBD2 might play an important function by interacting with other factors in this complex machinery.

The Linker Region of U1A Protein Adjacent to RBD. This FRET study of the linker region adjacent to RBD1 has given us structural and dynamic data describing the behavior of the U1A proteins close to the core of the complex. The C-terminal tail of the RBD1 domain is formed by the α_3 helix (from amino acid 92 to 98). This helix has been shown to interact weakly with the surface of the β-sheet surface in the free protein, but it is pushed away upon binding to permit the specific interactions between the RNA bases and the amino acids on the surface of the β-sheet (*16*, *26*, *37*). Previous data have shown that the α_3 helix helps to order the adjacent amino acids (Thr89Asp90Ser91) that make specific contacts with the RNA (*19*). Thus, in the UTR complex, the U1A protein probably adopts a rigid structure extending to the residue 98. In addition to stabilizing the specific interactions between RBD1 and the RNA, the α_3 helix may be used also as a spacer (or an arm) between the core of the complex and the flexible region of the linker. Our data show that the apparent distance between fluorophores at position 101 in the adjacent proteins in the UTR complex is independent of whether the protein is RBD1 or U1A, suggesting that the rest of the U1A protein (linker and RBD2) has no effect on the relative position of two 101 residues in the UTR complex. This is not surprising since the residue 101 is only three amino acids away from the end of the α_3 helix.

The positively charged sequence (ERDRKREKRKPK-SQETP) from residue 103 to 119 in the linker region adjacent to the core of the complex is predicted by PHD (*25*) to be unstructured. Our FRET data indicate that this region does not adopt a rigid extended conformation suggesting that it conserves a certain flexibility. Our FRET experiments also suggest that the 103−119 sequences of the two U1A proteins bound to the UTR RNA might interact together in a compact conformation that keeps the fluorophores at position 121 relatively close ($\langle R \rangle = 58$ Å). Indeed, the charged 103−119 sequences of the two U1A proteins bound to the UTR RNA have been proposed to form a dimer that interacts specifically with the C-terminal 20 amino acids of poly(A) polymerase (PAP) to enhance its inhibition (*12*). The dimerization

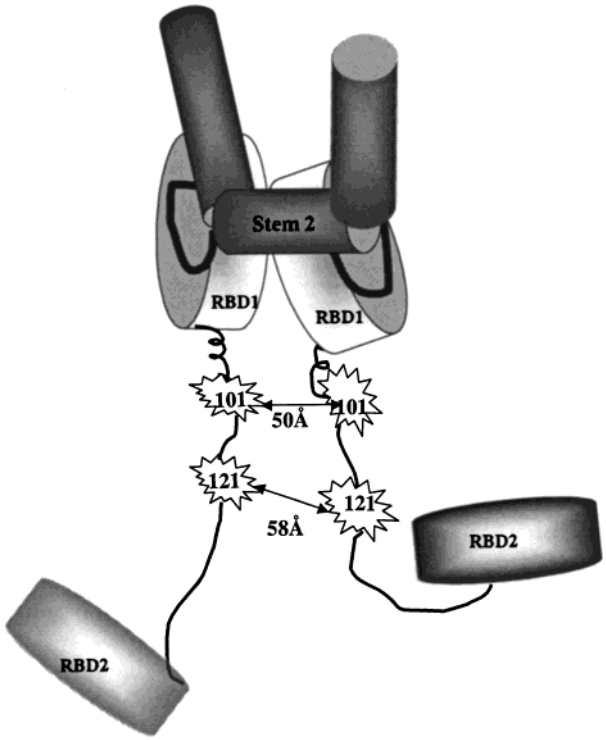


FIGURE 6: U1A−UTR structural model. UTR structure is modeled after Grainger et al. (1999) with the RNA bent to form a U shape. The two U1A proteins are shown schematically with RBD1 bound to the internal loop, the linker regions from 101 to 121 juxtaposed, and RBD2 randomly oriented away from the core of the complex, in agreement with the FRET data. Positions of FRET dyes in the linker are indicated.

hypothesis is further supported by the observation that two U1A proteins bound to the UTR are required to inhibit PAP activity. The models that have been proposed for the U1A−UTR complex (*14*, *15*, *36*) agree that the two U1A proteins lie on the same face of the RNA, favoring protein−protein interactions. These two binding sites are separated by only a four base pair stem of which the length is critical for the inhibition of the poly(A) polymerase activity, suggesting that both the spacing between the two binding sites and the relative position (or orientation) of the two U1A proteins are important (*36*).

The FRET experiments carried out on the U1A−UTR complex have provided structural data that are consistent with previous results (*11*, *14*, *18*, *30*) and which together lead to a model for the U1A−UTR complex shown in Figure 6. As suggested by the results of Grainger et al. (1999), the UTR RNA is severely bent with two 90° turns and adopts a U shape, positioning the two binding sites next to each other on the same face of the RNA. The two RBD1 domains of the U1A proteins bind to the two internal loops, leaving the protein linker regions flexible; the two RBD2 domains thus freely diffuse away from the core of the complex. Previous FRET experiments by the Lilley group (*15*) fluorescently labeled the RNA at the ends of the UTR stem 1 and stem 3. In the absence of protein, the average distance ($\langle R \rangle$) between the ends of the RNA was 64 Å. Upon binding of two RBD1 proteins, $\langle R \rangle$ was reduced to 55 Å, indicating a significant change in the global geometry of the RNA. The authors offered several physical explanations for this observation, including kinking of the UTR boxes or alteration of the dihedral angles between stem 1 and stem 3. On the basis of

their modeling, however, neither of the previously proposed models of the UTR−RBD1 complex (*16*, *36*) was able to accommodate such geometries.

The addition of the distance data from our FRET experiments may assist in the modeling of this structure. When the protein (RBD1 or U1A) is bound to the RNA, it seems likely that the structure of the body of the RBD is relatively constrained through contacts with the RNA. RBD1 α3 should be held away from the body of the RBD−RNA complex; contacts between the residues at the β4/α3 hinge and the RNA will fix its geometry. The attachment of the FRET pair at position 101 makes this position a convenient one for reporting on what should be a relatively fixed geometry between two bound proteins. Because the observed $\langle R \rangle$ of ∼50 Å between the adjacent 101-to-101 positions is independent of whether the protein is RBD1 or U1A, such a hypothesis seems reasonable, since the presence of the additional 20 kDa of U1A protein would be expected to exert some tension on the C-terminal tail of RBD1 to pull it into random orientations. This distance will be most useful in further modeling of the core of the UTR−U1A structure. The same probes now at position 121 in U1A show a 50% reduction in FRET efficiency, which corresponds to $\langle R \rangle =$ 59 Å. These results would be consistent with greater flexibility of the chain at this distance away (20 amino acids) from the (rigid) body of the complex. It would be possible to map the distances between respective sites on the linker by progressively moving the FRET pair down the length of the linker and so determine the point of divergence of the protein chains that leads to uncorrelated motion of RBD2. These experiments would require use of a FRET pair with a larger $\langle R_o \rangle$.

Implication for the Biological Activity of the U1A−UTR Complex. It is important and useful to relate this structural model to the biological function in vivo because many other factors are required in this 3′ UTR region. Indeed, the pre-mRNA is posttranscriptionally modified by the addition of the poly(A) tail at its 3′-end, which requires that the heterotetrameric complex cleavage and polyadenylation specificity factor (CPSF) binds to the conserved A(A/U)-UAAA sequence (*38*−*40*) and the heterotrimeric complex cleavage stimulatory factor (CStF) binds to the G/U-rich element (*41*, *42*). Also, the poly(A) polymerase and auxiliary proteins are recruited in this region, forming a large complex (*43*−*47*). The polyA signal (A(A/U)UAAA) stem-loop is separated by only two nucleotides from the UTR RNA, juxtaposing the two bound U1A proteins to the polyadenylation machinery.

Since the U1A protein autoregulates its production by binding to the 3′ UTR of its own pre-mRNA, the mechanics and the interactions between the proteins and the RNA sites in this large complex will be functional features of the regulation. If the U1A 103−119 sequence forms a dimer and interacts with the C-terminal domain of the PAP, the rest of the U1A protein (120−195 amino acids of the linker and the RBD2 domain) will have to be excluded (or pushed away) to permit this particular protein−protein interaction to occur. The presence of a very flexible sequence might be the best way to resolve this kind of conformational constraint problem. Indeed, the linker may be used as a flexible arm between two distinct domains (RBD1 and RBD2) to allow first, a great conformational adaptation of the U1A protein and also to allow the two RBD domains to reach their respective active sites. The dynamic features of the U1A protein are certainly a key for its biological functions in the polyadenylation or the splicing apparatus.

To pursue the characterization of this polyadenylation complex, it will be important to investigate the structural and dynamic properties of the U1A protein within a more complete polyadenylation apparatus containing the PAP and/or the polyA signal (A(A/U)UAAA) site. The study of this complex in conditions that include more components of the polyadenylation complex might even provide some information about the biological function of the RBD2 domain of U1A protein.

Note: Klein Gunnewiek et al. [Klein Gunnewiek, J. M. et al. (*2000*) *Mol. Cell. Biol. 20*, 2209] describe new results that indicate the importance of the U1A sequence 103−115 for inhibition of polyadenylation in the UTR complex.

ACKNOWLEDGMENT

We thank Dr. John Jean for many helpful discussions and Professor Tim Lohman for the use of his fluorimeter.

REFERENCES

1. Birney, E., Kumar, S., and Krainer, A. R. (1993) *Nucleic Acids Res. 21*, 5803−5816.
2. Burd, C. G., and Dreyfuss, G. (1994) *Science 265*, 615−621.
3. Ghetti, A., Padovani, C., Di Cesare, G., and Morandi, C. (1989) *FEBS Lett. 257*, 373−376.
4. Sillekens, P. T., Habets, W. J., Beijer, R. P., and van Venrooij, W. J. (1987) *EMBO J. 6*, 3841−3848.
5. Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) *Nature 348*, 515−520.
6. Lu, J., and Hall, K. B. (1997) *Biochemistry 36*, 10393−10405.
7. Scherly, D., Boelens, W., van Venrooij, W. J., Dathan, N. A., Hamm, J., and Mattaj, I. W. (1989) *EMBO J. 8*, 4163−4170.
8. Moore, M. J., Query, C. C., and Sharp, P. (1993). Splicing of Precursors to mRNA by the Spliceosome, in *The RNA World* (Gesteland, R. F., Atkins, J. F., Eds) pp 303−357, Cold Spring Harbor Laboratory Press, New York.
9. Hall, K. B., and Stump, W. T. (1992) *Nucleic Acids Res. 20*, 4283−4290.
10. Boelens, W. C., Jansen, E. J. R., van Venrooij, W. J., Stripecke, R., Mattaj, I. W., and Gunderson, S. I. (1993) *Cell 72*, 881−892.
11. Gunderson, S. I., Beyer, K., Martin, G., Keller, W., Boelens, W. C., and Mattaj, I. A. (1994) *Cell 76*, 531−541.
12. Gunderson, S. I., Vagner, S., Polycarpou-Schwartz, M., and Mattaj, I. W. (1997) *Genes Dev. 11*, 761−773.
13. van Gelder, C. W., Gunderson, S. I., Jansen, E. J., Boelens, W. C., Polycarpou-Schwarz, M., Mattaj, I. W., and van Venrooij, W. J. (1993) *EMBO J. 12*, 5191−5200.
14. Grainger, R. J., Murchie, A. I. H., Norman, D. G., and Lilley, D. M. J. (1997) *J. Mol. Biol. 273*, 84−92.
15. Grainger, R. J., Norman, D. G., and Lilley, D. M. J. (1999) *J. Mol. Biol. 228*, 585−594.
16. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., and Nagai, K. (1994) *Nature 372*, 432−438.
17. Hall, K. B. (1994) *Biochemistry 33*, 10076−10088.
18. Gubser, C. C., and Varani, G. (1996) *Biochem 35*, 2253−2260.
19. Kranz, J. K., and Hall, K. B. (1999) *J. Mol. Biol. 285*, 215−231.
20. Boelens, W., Scherly, D., Jansen, E. J. R., Kolen, K., Mattaj, I. W., and van Venrooij, W. J. (1991) *Nucleic Acids Res. 19*, 4611−4618.
21. Klein Gunnewiek, J. M. T., van de Putte, W., and van Venrooij, W. J. (1997) *Clin. Exp. Rheumatol. 15*, 549−560.
22. Lu, J., and Hall, K. B. (1995) *J. Mol. Biol. 247*, 739−752.
23. Rost, B., and Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.

24. Rost, B., Sander, C., and Schneider, R. (1994) *CABIOS, Comput. Appl. Biosci. 10*, 53−60.
25. Rost, B. and Sander, C. (1994) *Proteins 19*, 55−72.
26. Jean, J. M., Clerte, C., and Hall, K. B. (1999) *Protein Sci. 8*, 2110−2120.
27. Shortle, D. (1981) *Annu. Rev. Genet. 15*, 265−294.
28. Grodberg, J., and Dunn, D. (1988) *J. Bacteriol. 170*, 1245−1253.
29. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Chapter 14, Cold Spring Habor Laboratory Press, Cold Spring Harbor, NY.
30. Beck, D. L., Stump, W. T., and Hall, K. B. (1998) *RNA 4*, 331−339.
31. Lohman, T. M., and Mascotti, D. P. (1992) *Methods Enzymol. 212*, 424−458.
32. Lakowicz, J. R. (1983) Principles of Fluorescence Spectroscopy, Plenum, New York.
33. Eftink, M. R. (1991) *Methods Biochem. Anal. 35*, 127−195.
34. Selvin, R. P. (1995) *Methods Enzymol. 246*, 300−335.
35. Pengguang, W., and Brand, L. (1992) *Biochemistry 31*, 7939−7947.
36. Jovine, L., Oubridge, C., Avis, J. M., and Nagai, K. (1996) *Structure 4*, 621−631.
37. Avis, J. M., Alain, F. H−T., Howe, P. W. A., Varani, G., Nagai, K., and Neuhaus, D. (1996) *J. Mol. Biol. 257*, 398−411.
38. Bienroth, S., Wahle, E., Suter-Crazzolara, C., and Keller, W. (1991) *J. Biol. Chem. 266*, 19768−19776.
39. Keller, W., Bienroth, S., Lang, K. M., and Christofori, G. (1991) *EMBO J. 10*, 4241−4249.
40. Bardwell, V. J., Wickens, M., Bienroth, S., Keller, W., Sproat, B. S., and Lamond, A. I. (1991) *Cell 65*, 125−133.
41. Takagaki, Y., Manley, J. L., MacDonald, C. C., Wilusz, J., and Shenk, T. (1990) *Genes Dev. 4*, 2112−2120.
42. Murthy, K. G. K., and Manley, J. L. (1992) *J. Biol. Chem. 267*, 14804−14811.
43. Bardwell, V. J., Zarkower, D., Edmonds, M., and Wickens, M. (1990) *Mol. Cell. Biol. 10*, 846−849.
44. Takagaki, Y., Reyner, L. C., and Manley, J. L. (1989) *Genes Dev. 3*, 1711−1724.
45. Cristofori, G., and Keller, W. (1988) *Cell 54*, 875−889.
46. Gilmartin, G. M., and Nevins, J. R. (1989) *Genes Dev. 3*, 2180−2189.
47. Wahle, E. (1991) *Cell 66*, 759−768.

BI000372K