Global and Local Dynamics of the U1A Polyadenylation Inhibition Element (PIE) RNA and PIE RNA−U1A Complexes

Caroline Clerte and Kathleen B. Hall*

Department of Biochemistry and Molecular Biophysics, Box 8231, Washington University School of Medicine, St. Louis, Missouri 63110

Received April 30, 2004; Revised Manuscript Received August 13, 2004

ABSTRACT: The structure and dynamics of the polyadenylation inhibition element (PIE) RNA, free and bound to the U1A protein, have been examined using time-resolved FRET and 2-aminopurine (2AP) fluorescence. This regulatory RNA, located at the 3′ end of the U1A pre-mRNA, adopts a U-shaped structure, with binding sites for a single U1A protein at each bend (box 1 and box 2). The distance between the termini of the arms of the RNA is sensitive to its three-dimensional structure. Using Cy3/Cy5 FRET efficiency to monitor binding of Mg^{2+}, we show that the PIE RNA binds two Mg^{2+} ions, which results in a restriction of its distance distribution of conformations. Local RNA structure probing using 2AP fluorescence shows that the structure of box 2 changes in response to Mg^{2+} binding, thus tentatively locating the ion binding sites. Steady-state FRET data show that the distance $\langle R \rangle$ between the termini of the PIE RNA stems decreases from 66 Å in the free RNA, to 58 Å when N-terminal RNA binding domains (RBD1) of U1A are bound, and to 53 Å when U1A proteins bind. However, anisotropy measurements indicate that both Cy3 and Cy5 stack on the ends of the RNA. To examine the consequences of the restricted motion of the fluorophores, FRET data are analyzed using two different models of motion and then compared to analogous data from the Cy3/fluorescein FRET pair. We conclude that the error introduced into distance calculations by stacking of the dyes is within the error of our measurements. Distance distributions of the RNA structures show that the intramolecular distance between the arms of the PIE RNA varies on the time scale of the fluorescence measurements; the mean distance is dependent on protein binding, but the breadth of the distributions indicates that the RNA retains structural heterogeneity.

The human U1A protein is a component of the eukaryotic pre-mRNA processing machinery (*1*). It is involved in pre-mRNA splicing in that it is one of three specific proteins associated with the U1 snRNP: it binds specifically to stem loop II of the U1 snRNA (*2*). U1A contains three distinct domains: the N-terminal RNA binding domain (RBD) RBD1 (amino acids 1−102) (*3*), the C-terminal RBD2 (amino acids 195−282) (*4*), and the interdomain linker of ∼100 amino acids. Only the RBD1 domain of U1A is required for specific binding to its RNA targets (*2*). RBD2 does not bind to any RNA (*5*); its biological function remains unknown. The linker domain has been shown to be extremely flexible (*6*).

U1A is also involved in pre-mRNA 3′-end processing where it autoregulates its expression by binding to the 3′ untranslated region (UTR) of its own pre-mRNA to prevent polyadenylation (*7*−*9*). The 3′UTR of the U1A pre-mRNA contains a regulatory sequence of about 50 nucleotides located at a conserved distance upstream of the cleavage/polyadenylation signal (*7*). There are two binding sites for U1A protein in the 3′ UTR, and two full-length U1A proteins associated with the RNA are required for regulation of polyadenylation; RBD1 alone is not effective (*7*, *8*). Each binding site consists of an asymmetric internal loop, box 1 and box 2, and is comprised of seven unpaired bases [AUUGCAC (box 2) or AUUGUAC (box 1)] and a single unpaired base (A or C) on the opposite strand (*10*). The two internal loops are separated by a four base pair stem (stem 2). This structure is referred to as the polyadenylation inhibition element (PIE) RNA (Figure 1).

The topology of the PIE RNA is uniquely suited to its function. The asymmetry of the two internal loops in the RNA introduces two severe kinks which give rise to a global U-shape, juxtaposing the two U1A binding sites on the same face of the RNA (*11*). This unusual geometry of the PIE RNA was first deduced using polyacrylamide gel electrophoresis (*11*), where it was observed that migration of the RNA was retarded in the gel similar to that of RNA containing internal A_n bulges, leading to the description of PIE RNA axial bending. A more quantitative characterization of the PIE RNA U-shaped geometry was accomplished using fluorescence resonance energy transfer (FRET). For these experiments, the RNA was end-labeled with the fluorescein/Cy3 FRET pair, and the steady-state FRET efficiency was determined (*12*). The global bend of the free RNA was described by an interfluorophore distance of 64 Å.

Upon U1A RBD1 binding to fluorescein/Cy3-labeled PIE RNA, the apparent distance between the fluorophores is reduced to 54 Å, indicating that the global geometry of the RNA has been altered (*12*). An NMR structure of RBD1−

* Corresponding author. E-mail: kbh@wanda.wustl.edu. Phone: 314-362-4196. Fax: 314-362-7183.

10.1021/bi049117g CCC: $27.50
Published on Web 09/29/2004

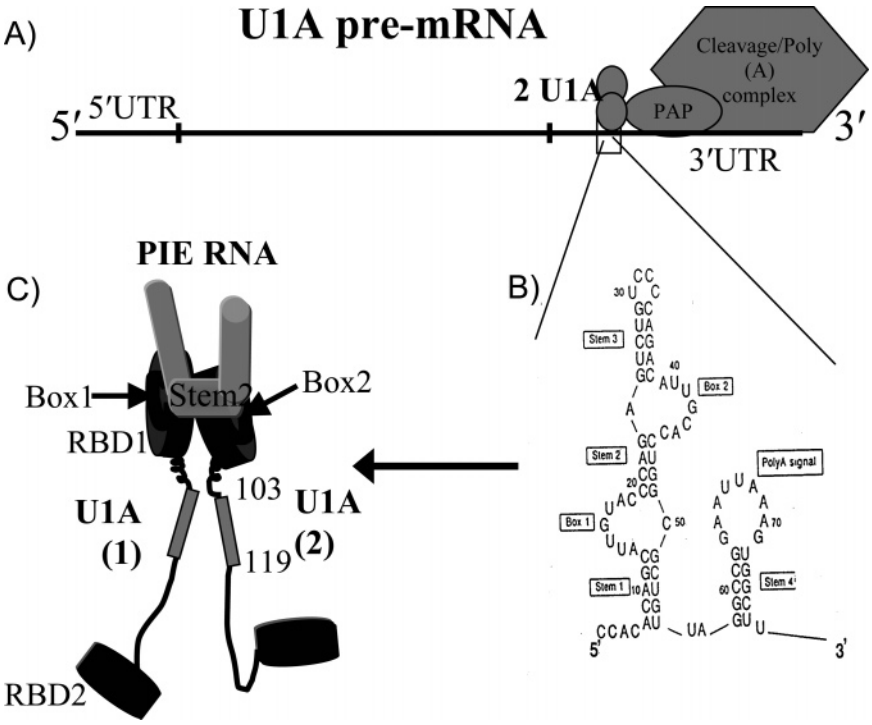

FIGURE 1: Schematic representation of the autoregulation mechanism of the U1A protein. (A) Two U1A proteins bind to the 3′UTR of its own pre-mRNA to inhibit the polyadenylation reaction by the 3′ end pre-mRNA processing machinery where PAP represents the poly(A) polymerase enzyme. (B) The secondary structure of the regulatory sequence within the 3′UTR, to which the two U1A proteins bind, is shown (*10*). (C) Structural model of the PIE RNA−U1A (1:2) complex. The PIE RNA forms a U-shape (*12*). Two U1A proteins are bound to the RNA via RBD1−box 1 and RBD1−box 2. The charged sequences (103−119) (represented by squares) are proposed to form a dimer (*9*); RBD2 domains move independently from the core of the complex (*16*).

box 2 (*13*) shows that box 2 adopts a stable structure through its interactions with the protein. An NMR study of box 2 alone showed that, in solution (5 mM sodium phosphate, pH 5.5), there was considerable flexibility of the bases in the internal loop (*14*); thus the interaction with RBD1 locks box 2 into a specific conformation. A similar result will occur with the box 1−RBD1 interaction (although its U → C mutation will reduce affinity).

RBD1 alone is not sufficient to inhibit polyadenylation by poly(A) polymerase (PAP): the entire U1A protein was required (*8*). It was possible that the conformational change induced in the PIE RNA when the full-length U1A protein is bound differs from that caused by the small RBD1. Our results show that there is indeed a difference.

In the PIE RNA−U1A protein (1:2) complex, the positively charged linker sequences (103−119) of the two U1A proteins are proposed to form a dimer which interacts directly with the C-terminal domain of poly(A) polymerase (PAP) to prevent polyadenylation (*7*−*9*, *15*). Structural characterization of the PIE−U1A complex showed that the charged sequences (103−119) in the linker adopt a compact conformation and are in proximity to each other (*16*). However, the remainder of the interdomain linker of the two U1A proteins was flexible, with the consequence that the two RBD2 domains tumble independently from the core of the complex (Figure 1).

Previous steady-state fluorescence experiments that described the PIE RNA produced an average static picture of its global structure. Such a picture is useful but may be misleading, since molecular flexibility can be an important part of its function. For example, perhaps the downstream stem loop containing the polyadenylation AAUUAAA sequence is coaxially stacked on stem 1 of the PIE RNA (see Figure 1). When two U1A proteins are bound to the PIE element, that complex may need to change its conformation in order to accommodate the large PAP complex bound in the vicinity of the poly(A) signal. More information about the intrinsic flexibility of the PIE−U1A complex will facilitate construction of a more accurate functional model.

Here, we use time-resolved fluorescence resonance energy transfer (TR-FRET) to obtain a description of the global dynamics of the free PIE RNA and PIE−protein complexes. For these TR-FRET experiments, three proteins were bound to the RNA: (i) RBD1 (1−101), (ii) full-length U1A (WT) protein, and (iii) full-length U1A (IL) protein in which the charged-linker sequence 101−121 and the sequence 122−140 have been swapped (Figure 2). We found that, in all PIE−protein (1:2) complexes, the mean distance between fluorophores became shorter, as indeed noted previously (*12*), but the distance distributions remain broad in the presence of the proteins, indicating that a heterogeneous population of structures is present.

In addition to examination of the global properties of the PIE RNA, its local structural and dynamic properties were probed by means of a single 2-aminopurine base specifically introduced either in stem 2 or in box 2. Its fluorescence properties were measured using time-resolved fluorescence and time-resolved anisotropy. On the basis of the fluorescence data, 2AP in stem 2 is stacked in the free RNA and remains stacked in the PIE−protein complexes. The 2AP in box 2 is stacked in the free RNA but becomes more unconstrained in the complexes. The combination of FRET and 2AP fluorescence measurements showed that two Mg^{2+} ions bind to PIE RNA, presumably in box 1 and box 2.

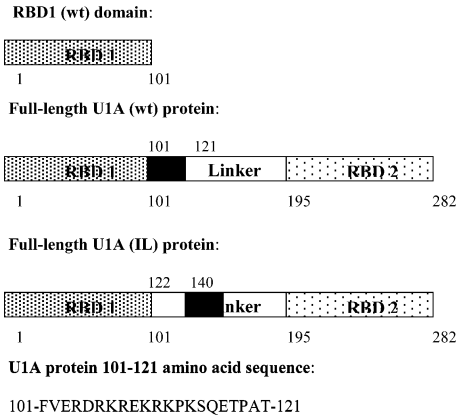

FIGURE 2: Schematic representation of the different recombinant proteins: RBD1 (1−101), U1A (WT) (1−282) protein, and U1A (IL) protein with the linker-charged sequence 101−121 (designated by the black square) swapped with the sequence 122−140. Amino acid sequence 101−121 in the linker domain is shown.

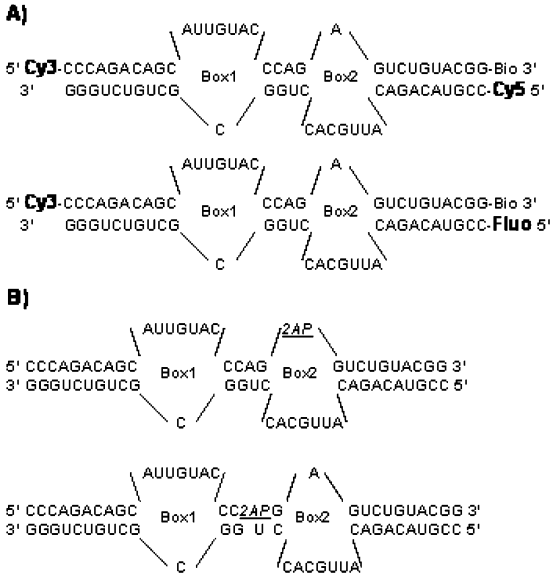

FIGURE 3: RNA molecules synthesized. (A) The 5′-end-labeled PIE RNA strand with Cy3, Cy5, or fluorescein (Fluo) for FRET experiments. (B) The internally labeled PIE RNA strand with 2-aminopurine (2AP) either in stem 2 or in box 2.

MATERIALS AND METHODS

RNA Fluorescent Sample Preparation. The PIE RNA was constructed from two strands of RNA, so that FRET pairs could be attached to the termini of stem 1 and stem 3 (Figure 3). RNA strands labeled at the 5′ end with Cy3, fluorescein, or Cy5 or internally with 2AP were purchased from Dharmacon Research Inc. (Boulder, CO). Strand labeling efficiency was verified by absorption spectroscopy (A_{260nm}/A_{550nm} for Cy3, A_{260nm}/A_{650nm} for Cy5, and A_{260nm}/A_{492nm} for fluorescein). The Cy3 RNA strand, Cy5 RNA strand, and fluorescein strand labeling was 100%, 75%, and 60%, respectively. Identical strands without dyes were also synthesized: PIE-Cy3 strand, $^{5'}$Cy3-CCCAGACAGCAUUGUACCCAGAGUCUGUACGG-Bio; PIE-Cy5 strand, $^{3'}$GGGUCUGUCGCGGUCCACGUUACAGACAUGCC-Cy5; PIE-fluorescein strand, $^{3'}$GGGUCUGUCGCGGUCCACGUUACAGACAUGCC-fluorescein; PIE-2AP-stem 2 strand, $^{5'}$CCCAGACAGCAUUGUACCC2APGAGUCUGUACGG; PIE-2AP-box 2 strand, $^{5'}$CCCAGACAGCAUUGUACCCAG2APGUCUGUACGG; PIE complementary strand, $^{3'}$GGGUCUGUCGCGGUCACGUUACAGACAUGCC.

(A) Cy3-Cy5 FRET Sample Preparation. PIE-Cy3 RNA was mixed with an excess of PIE-Cy5 RNA strand in milliQ water and heated to 95 °C for 5 min. Samples were cooled to 65 °C for 5 min, and buffer was added to give a concentration of 10 mM Tris (pH 8), 100 mM NaCl, and ±1 mM $MgCl_2$ in 1 mL final volume. Samples were cooled slowly to room temperature. The annealed sample was added to 1× buffer [10 mM Tris (pH 8), 100 mM NaCl, 75 μg/mL BSA, ±1 mM $MgCl_2$] in the fluorescence cuvette to give a final concentration of 50 and 70 nM for PIE-Cy3 RNA and the PIE-Cy5 RNA complementary strand, respectively.

(B) Fluorescein-Cy3 FRET Sample Preparation. The same conditions were used for the sample preparation except that final concentrations of 50 and 52 nM for the PIE-fluorescein RNA strand and PIE RNA complementary strand were used, respectively.

(C) 2AP RNA Fluorescence Sample Preparation. The same annealing protocol was followed. The solution contained 10 mM KPO_4 (pH 7), 50 mM NaCl, and 1 mM $MgCl_2$. The final RNA concentration was 4 and 4.2 μM for the PIE-2AP strand and PIE strand, respectively.

(D) Protein Construction. Site-directed mutagenesis standard protocols were used to generate all of the RBD1 and U1A mutants (*17*). All mutant DNAs were cloned into a plasmid under tac promoter control, and their sequences were verified by sequencing. All proteins were expressed in *Escherichia coli* BL21. The U1A (IL) mutant has the charged sequence (residues 101−121) of the linker domain and the sequence 121−141 inverted (Figure 2). The U1A (IL) cloning was done by ligation of three PCR fragments: (i) 1−101 PCR fragment, (ii) 101−140 fragment obtained by chemically synthesized DNA oligos, and (iii) 140−282 PCR fragment.

Protein Purification. (A) RBD1 Protein. BL21 cells were grown at 37 °C in LB media and induced in mid-log phase ($OD_{600nm} \cong 0.7$) with 1 mM isopropyl β-thiogalactoside (IPTG). The cells were allowed to grow for an additional 4 h after induction, centrifuged, washed with 20 mM Tris, 20 mM NaCl, and 2 mM ethylendiaminetetraacetic acid (EDTA), pH 7.5 (buffer G1), and resuspended in 20 mM sodium acetate, 200 mM NaCl, 2 mM EDTA, pH 5.3 (buffer B), containing pepstatin, and protease inhibitor cocktail (Sigma P2714) containing 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF), *trans*-epoxysuccinyl-L-leucylamino(4-guanidino)-butane (E-64), bestatin, leupeptin, aprotinin, and EDTA. Cells were lysed by French pressing. The lysate was incubated on ice for 30 min with DNase II and then centrifuged at 30000*g* for 30 min at 4 °C to remove the cell debris. The supernatant was dialyzed against 1 L of buffer B containing phenylmethanesulfonyl fluoride (PMSF) at 4 °C for about 2 h. The dialysate was loaded on a SP-Sepharose column equilibrated with 50 mM Tris and 20 μg/mL PMSF, pH 8. The flow-through contains the RBD1 protein which was then purified on a CM-Sepharose column equilibrated with 50 mM Tris, 100 mM NaCl, and 20 μg/mL PMSF, pH 7.5. RBD1 protein was eluted off the column in a gradient from 100 to 500 mM NaCl in buffer C containing PMSF. Fractions containing RBD1 protein were collected and concentrated using Vivaspin 20, 3000 MWCO. Finally, the concentrated protein was loaded on a G50 column equilibrated with 50 mM PO_4, pH

7. Fractions containing the RBD1 protein were concentrated using the Vivaspin.

(B) U1A Protein. U1A proteins were expressed and purified following standard protocols (*6*).

All proteins were checked for purity by silver-staining samples electrophoresed in SDS−PAGE gels. All proteins concentrations were determined spectrophotometrically at 280 nm using an extinction coefficient based on the Tyr/Phe content. Circular dichroism spectra and filter binding assays were performed on all proteins to verify their secondary structure and their primary RNA binding activity.

Steady-State Fluorescence Measurements. (A) Fluorescence Resonance Energy Transfer (FRET) Measurements for the Cy3Cy5 FRET Pair. FRET experiments were performed on an SLM 8100 spectrofluorometer with an excitation wavelength of 514 nm (bandwidth 2 nm) scanning from 525 to 700 nm with an emission bandwidth of 4 nm. All experiments were done at a constant temperature of 25 °C. PIE-Cy3 strand RNA at 50 nM and PIE-Cy5 strand RNA at 70 nM in 10 mM Tris (pH 8), 100 mM NaCl, and 75 μg/mL BSA were used for all experiments. RBD1, U1A (WT), or U1A (IL) was added to the fluorescent PIE RNA samples to give a final 1:2 RNA:protein stoichiometry and to obtain a final protein concentration equal to 120 nM.

The Mg^{2+} ion titration experiment was done by adding small aliquots of a stock solution of $MgCl_2$ to cover a concentration range from 10 nM to 10 mM. Samples were allowed to equilibrate for 15 min before spectrum acquisition. Each spectrum was collected at least twice to ensure fluorescence signal stability over time. Fluorescence intensities were corrected for buffer background and dilution. Emission spectra were corrected for the lamp fluctuation and wavelength detection variation using a reference channel. Polarization artifacts were avoided by acquisition at the magic angle (54.7°).

FRET is a distance-dependent transfer of excited-state energy from a donor to an acceptor as a result of through-space coupling between the fluorophore transition dipoles. The energy transfer efficiency (E) depends on the inverse sixth power of the distance (r) separating the two fluorophores:

$$E = R_0^6/(r^6 + R_0^6)$$

where R_0 is the Förster distance at which the efficiency of FRET is equal to 0.5. R_0 is a characteristic of the FRET pair used and depends on the spectral properties of the donor and acceptor and the relative orientation of their transition dipoles:

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

where J is the overlap between donor emission and acceptor absorption, ϕ_D is the quantum yield of donor fluorescence, n^{-4} is the refractive index of the medium between the donor and acceptor, and κ^2 is the orientation factor.

The efficiency of energy transfer (E) determined by donor (Cy3) quenching was calculated according to

$$E = 1 - (F_{DA}/F_D) = 1 - (\tau_{DA}/\tau_D)$$

where F_D is the integrated fluorescence intensity and τ_D is the amplitude-weighted average lifetime of the donor in the absence of the acceptor, and F_{DA} is the integrated fluorescence intensity and τ_{DA} is the amplitude-weighted average lifetime of the donor in the presence of the acceptor ($\tau = \sum\alpha_i\tau_i$).

Several fluorescence measurements were made to characterize the RNA constructs. The fluorescence intensities of the PIE-Cy3 single strand and PIE-Cy3 double strand were found to be identical, as were those of the PIE-Cy5 single strand and double strand. Thus the duplex construct does not quench the fluorescence of either dye. Fluorescence of the PIE-Cy3 double strand in the presence and absence of the different proteins showed a dependence on the protein construct. Cy3 fluorescence intensity increased by 14% ± 4% upon both full-length U1A proteins binding, respectively. There was no change in the Cy3 fluorescence intensity upon RBD1 binding. Therefore, for the PIE RNA−protein complexes, the FRET efficiency was calculated as

$$E = 1 - (F_{DA+protein}/F_{D+protein}) = 1 - (\tau_{DA+protein}/\tau_{D+protein})$$

In contrast, there was no change in Cy5 fluorescence intensity when the PIE-Cy5 double strand was bound to protein. Finally, the Cy5 double strand fluorescence background, when exciting at 514 nm, is negligible.

Excitation spectra of PIE-Cy5 (acceptor only) and PIE-Cy3 (donor only) free and bound to the different proteins were also recorded to verify that the absorption band shape and maximum of Cy5 and Cy3 were not disturbed upon protein binding. Cy5 dye tends to stack and "aggregate" when the local concentration is high enough (*18*), giving rise to an enhancement of the bluest absorption band and resulting in quenching of Cy5 fluorescence. Cy3 and Cy5 excitation spectra did not shift upon protein binding.

(B) Fluorescence Control Experiments for the Fluorescein-Cy3 FRET Pair. Measurements were performed on an SLM 8100 spectrofluorometer with an excitation wavelength of 490 nm (bandwidth 2 nm) scanning from 500 to 650 nm with an emission bandwidth of 4 nm. All experiments were done at a constant temperature of 25 °C. PIE-fluorescein strand RNA at 50 nM and PIE complementary strand RNA at 52 nM in 10 mM Tris (pH8), 100 mM NaCl, and 75 μg/mL BSA were used for all experiments. Proteins were added to give a final 1:2 RNA:protein stoichiometric ratio and to obtain a final protein concentration equal to 120 nM. Samples were allowed to equilibrate for 15 min before spectrum acquisition. These control experiments showed that the fluorescence intensity of fluorescein is constant upon protein binding.

(C) Fluorescence Measurements for 2AP. Fluorescence spectra of the RNA labeled with 2AP were recorded from 330 to 450 nm on a PTI spectrofluorometer with an excitation wavelength of 308 nm (bandwidth 2 nm) and an emission bandwidth of 4 nm. Measurements were done at constant room temperature. The PIE-2AP RNA final concentration was 4 μM in 10 mM potassium phosphate (pH 7) and 50 mM NaCl. Small aliquots of protein (RBD1 or full-length U1A) or $MgCl_2$ were added to the RNA solutions to give a stoichiometry of 1:2 RNA:protein or a final concentration of 1 mM $MgCl_2$. Samples were allowed to equilibrate for at least 15 min before each emission spectrum acquisition. Each

emission spectrum acquisition was repeated at least twice in order to verify the stability of the fluorescence signal.

(D) Anisotropy Measurements. Steady-state anisotropies were measured using the L-format method on a SLM 8100 spectrofluorometer. Anisotropy values were determined using the equation:

$$r = (I_{||} - GI_{\perp})/(I_{||} + 2GI_{\perp})$$

where $I_{||}$ and $I_{\perp}$ are the intensities of the fluorescence components polarized parallel and perpendicular, respectively, to the vertically polarized excitation beam. The G factor corrects for the polarization-dependent properties of the detection system.

Anisotropy of Cy3-labeled RNA used an excitation wavelength of 514 nm (bandwidth 4 nm); fluorescence polarization was collected at 570 nm (bandwidth 8 nm). For the Cy5-labeled RNA, the excitation wavelength was 640 nm (bandwidth 4 nm), and the polarization fluorescence was measured at 675 nm (bandwidth 8 nm). For fluorescein RNAs, the anisotropy was measured with 490 nm excitation (bandwidth 2 nm), and the polarization fluorescence was measured at 525 nm (bandwidth 4 nm). Before each acquisition, samples were allowed to equilibrate for 15 min. Each anisotropy measurement was repeated until the standard deviation of the measurement was <0.001.

Time-Resolved Fluorescence Measurements. (A) Cy3-Cy5 RNA. TR-FRET experiments were acquired in the time domain with a time-correlated single photon counting (TCSPC) instrument. The output from the mode-locked titanium−sapphire laser (Coherent) tuned at 950 nm was doubled to 475 nm using a frequency doubler (Coherent). The instrument response function at this wavelength typically has a full width at half-height (fwhh) = 230 ps.

(B) Fluorescein-Cy3 RNA. The output from the mode-locked titanium−sapphire laser tuned at 930 nm was doubled to 465 nm where Cy3 absorption is minimal. The long fluorescence lifetime of fluorescein required a Neos pulse picker to obtain a 7.6 MHz repetition rate. The instrument response function had a fwhh = 195 ps. The fluorescence of PIE-Cy3 RNA in the absence and presence of different protein samples was corrected for Cy3 leakage through the fluorescein emission band-pass filter.

(C) Time-Resolved FRET (TR-FRET) Data Analysis. The donor-only fluorescence decays exponentially with time:

$$I_{\mathrm{D}}(t) = \sum_{i}^{n} \alpha_i \exp\left(-\frac{t}{\tau_{\mathrm{Di}}}\right)$$

where τ_{Di} is the individual donor lifetime associated with the amplitude α_i.

In the presence of the acceptor at different distances r, the donor fluorescence decay becomes

$$I_{\mathrm{DA}}(t) = \sum_{k}^{m} A_k \int_0^{\infty} P_k(r) \sum_{i}^{n} \alpha_{\mathrm{Di}} \exp\left[-\frac{t}{\tau_{\mathrm{Di}}}\left(1 + \left(\frac{R_0}{r}\right)^6\right)\right] \mathrm{d}r$$

where $(1/\tau_{\mathrm{Di}})(R_0/r)^6$ is equal to the energy transfer rate constant (k_t), A_k is the amplitude of the kth population, and $P_k(r)$ is the corresponding donor−acceptor distance distribution. Fluorescence decays were processed with the parameter estimation program CFS_LS (*19*). The goodness of fit was determined by examination of the shape and the boundaries of the support plane of each parameter and by the value of the reduced χ^2.

(D) Distance Distribution Analysis. To analyze donor decays in donor/acceptor samples as a distribution of distances separating the two fluorophores, several assumptions were made: (i) the multiexponential decay of the donor only is not due to excited-state reactions on the time scale of the measurement but due only to ground-state heterogeneity. (ii) The relative orientation between the donor and acceptor is random so that $R_0 = 60$ Å (Cy3/Cy5) or 56 Å (fluorescein/Cy3). (iii) The critical distance (or Förster distance) R_0 of each exponential is the same so that $R_{0i} = R_0$. (iv) Intramolecular diffusion between the donor and acceptor during the fluorescence lifetime of the donor excited state can be ignored. (v) The distance distributions $P(r)$ were represented by a weighted Gaussian distribution (*20*−*22*).

$$P(r) = \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left(-\frac{(r - \overline{r_k})^2}{2{\sigma_k}^2}\right)$$

The α_i and τ_i from donor-only decay were fixed in the analysis of the donor/acceptor decays as a function of distance distribution. Adjustable parameters are A_k (amplitude of the kth population), $\overline{r_k}$ (mean distance of the kth population), and σ_k, which is related to fwhh = $2.355\sigma_k$. The effects of the energy transfer on donor decays were fit by models that described the donor−acceptor distance distribution by either one Gaussian or two Gaussian functions. The goodness of fit was judged by the shape and the boundaries of the support plane and the value of χ^2_{R}.

The assumption that the distance distribution can be described by a Gaussian function is not necessarily a good one (*20*−*22*). The fitting program CFS_LS gives the choice of a Gaussian or a Lorentzian distribution function: Lorentzian distance distribution analysis gives rise to narrower distributions than Gaussian distributions (*20*). The distinction of the different forms of distribution is generally minimal due to the large distribution resulting from multiexponential fluorescence decays of our system (*23*).

In fitting the time-resolved FRET data to a distance distribution, the resulting mean distance ($\bar{r}$) depends in part on the labeling efficiency of the RNAs. This value is an input parameter in the CFS_GAUDIS fitting routine, but it can also be varied by the routine to improve the χ^2_{R}. Our independent spectroscopic determination of the fractional labeling showed that 75% of the RNA strands contained Cy5, while Cy3 labeling was 100% efficient, and fluorescein-RNA was 60% labeled.

Calculation of the distance distribution of PIE RNA in the RNA−protein complexes was carried out using several models. First, assuming 100% labeling efficiency, single Gaussian distributions are very broad and give a mean distance that agrees with the steady-state $\langle R \rangle$ for PIE−RBD1 and PIE−U1A (WT) complexes. If again 100% labeling efficiency is assumed, but the data are fit with two Gaussian distributions, then one distance distribution is centered around 90−99 Å, while the other is unique to the specific protein bound. Our independent determination of fractional labeling showed that Cy5 labeling efficiency was 75%; therefore, we

can assign the long distance distribution to that population of RNAs with donor but no acceptor, which in these RNAs is 25%. In practice, then, analysis specified a single Gaussian distribution, using the labeling efficiency as a variable in CFS_GAUDIS. The labeling efficiency variable parameter consistently converged to a value of 0.80 ± 0.05, confirming the spectral determination.

(E) Time-Resolved Fluorescence Measurements for 2-AP. PIE-2AP time-resolved fluorescence decays were acquired using the TCSPC instrument. The laser was tuned at 900 nm, and the output light was tripled to 300 nm with a UOplaz tripler. The long fluorescence lifetime of 2AP required a Neos pulse picker to obtain a 7.6 MHz repetition rate. The instrument response function of the PMT was 150 ps fwhh.

2AP fluorescence decays were processed using FluoFit (PicoQuant) and were best fit by four exponentials. The goodness of fit was estimated by the shape and boundaries of the support planes of each parameter and the value of χ^2_R. However, the limitation of the instrument (IRF fwhh ≅ 150 ps; resolution is 32.5 ps per channel) often gave rise to unbounded support plane analysis for the amplitude or the lifetime of the very short component (25−50 ps).

(F) Time-Resolved Anisotropy Measurements. For anisotropy decays, the emission polarizer was alternated between the vertical (parallel) and horizontal (perpendicular) position with respect to the incident vertically polarized light. For both orientations, the polarization components of the decays were collected for equal time. Anisotropy decays curves were calculated from the equation:

$$r(t) = \frac{\mathrm{vv}(t) - G\mathrm{vh}(t)}{\mathrm{vv}(t) + 2G\mathrm{vh}(t)}$$

where vv(t) is the parallel component and vh(t) is the perpendicular component. The G factor was calculated by tail matching in the Fluofit software and was determined to be close to 1 for the 2AP measurements and close to 2 for the Cy3 measurements. The anisotropy decays are usually described by two exponential decays as a function of time:

$$r(t) = \beta_f \exp\left(-\frac{t}{\phi_f}\right) + \beta_s \exp\left(-\frac{t}{\phi_s}\right)$$

where β_f is the amplitude associated with the fast anisotropy component ϕ_f and β_s is the amplitude associated with the slow anisotropy component ϕ_s.

RESULTS

To describe the distance and its variation between the two arms of the PIE RNA, two FRET pairs were used. The Cy3/Cy5 pair has the technical advantage that the donor/acceptor excitation/emission wavelengths are cleanly separated. The fluorescein/Cy3 pair was used by Grainger et al. (*12*) in their steady-state FRET study of PIE RNA, and it is used here as a control to address the problem of orientation of the dyes on the RNA.

PIE RNA Global Parameters. Fluorescence intensities of the Cy3-, Cy5-, and Cy3/Cy5-labeled RNAs are shown in Figure 4, exciting Cy3 at 514 nm. In conditions without $MgCl_2$, the apparent distance between the two fluorophores is $\langle R\rangle$ = 66 ± 1 Å, with an efficiency E = 35% ± 2%. Upon addition of 1 mM $MgCl_2$, the FRET efficiency increases to 40% ± 3% and $\langle R\rangle$ = 64 ± 2 Å, in agreement with previous data of Grainger et al. (12), who obtained a value of 64 Å using the fluorescein/Cy3 FRET pair in similar solution conditions with $MgCl_2$. The data indicate that, upon binding of Mg^{2+} ions, the RNA global geometry becomes more restricted.

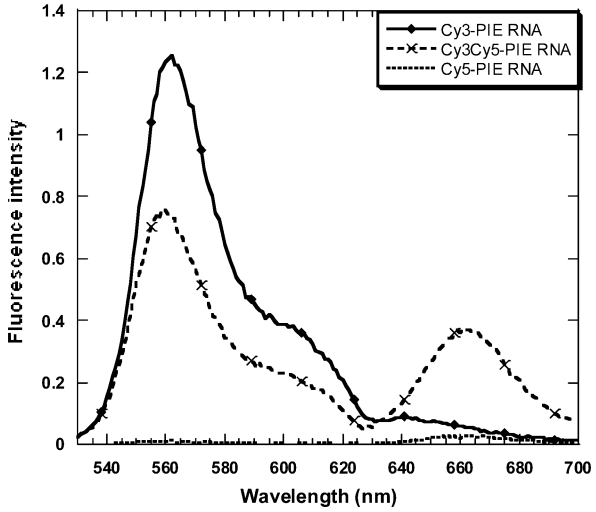

FIGURE 4: Steady-state FRET experiments with the Cy3Cy5 FRET pair. Conditions: excitation wavelength, 514 nm; bandwidth, 2 nm; emission bandwidth, 4 nm. Buffer: 10 mM Tris (pH 8), 100 mM NaCl, 75 μg/mL BSA, and 1 mM $MgCl_2$, at 25 °C. [Cy3-labeled RNA strands] = 50 nM; [Cy5-labeled RNA strands] = 70 nM.

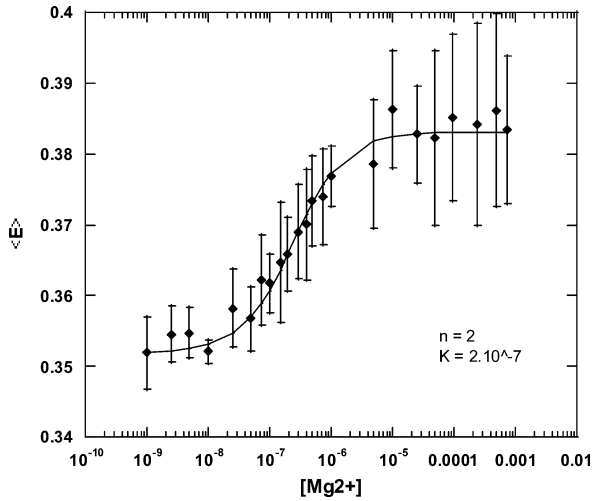

FIGURE 5: Ion-induced stabilization of the free PIE RNA studied by the change in FRET efficiency (E) as a function of Mg^{2+} concentration. [RNA] = 50 nM. Buffer: 10 mM Tris (pH8), 100 mM NaCl, and 75 μg/mL BSA, at 20 °C. Data were fit assuming identical and noninteracting binding sites; the curve shown corresponds to two sites with an apparent dissociation constant (K) = 2 × 10^{-7} M. Error bars are obtained by standard deviation calculation based on five different and independent measurements.

To further probe for divalent ion association with the PIE RNA, the FRET efficiency (E) was used as a reporter in $MgCl_2$ titration experiments, shown in Figure 5. The data were fit to a model in which the ligand binding sites are identical and noninteracting, assuming that the change in fluorescence (or here the FRET efficiency) is additive upon addition of Mg^{2+} ligands (24):

$$[Mg^{2+}]_T = [Mg^{2+}]_F(1 + K_A([RNA]_F)^n)$$

$$[\mathrm{RNA}]_T = [\mathrm{RNA}]_F + [\mathrm{Mg}^{2+}]_T[(nK_A[\mathrm{RNA}]_F)/(1 + K_A[\mathrm{RNA}]_F)]$$

$$(E - E_0)/E_0 = (\Delta E_{max} nK_A[\mathrm{RNA}]_F)/(1 + K_A[\mathrm{RNA}]_F)$$

$[\mathrm{RNA}]_F$ is the free PIE RNA concentration, K_A the association constant for ligand binding, n the number of ligands bound, and ΔE_{max} the difference in FRET efficiency from the initial value E_0 in no $[MgCl_2]$ to the final value. The best fit to the data was found when $n = 2 \pm 1$; there appears to be a finite number of ions required for the conformational change.

(A) Distance Distribution in Solution. Measurement of steady-state FRET efficiencies provides a static average picture of the RNA, and it was possible that the change in $\langle R \rangle$ was not due to a global conformation change but to a shift in the distribution of conformational states as the RNA moved through a range of possible structures. To determine the distance distribution of structures in solution, time-resolved FRET experiments were done in the absence and presence of a 1 mM final concentration of Mg^{2+} ion. Analysis finds a mean distance ($\bar{r}$) of 54 Å, independent of $MgCl_2$ (Figure 6). (Note that the mean distance is shorter than $\langle R \rangle$ calculated from steady-state FRET.) The difference between the structures is seen in the width of the Gaussian distance distribution, which substantially decreases upon $MgCl_2$ addition.

(B) Orientation Factor κ^2 Assumption. The FRET data analysis assumed that the orientation factor $\kappa^2 = {}^2/_3$, which corresponds to the dynamic average model where the transition dipoles of the donor and acceptor are randomly oriented on the time scale of the FRET process. To evaluate this assumption for the PIE RNA Cy3/Cy5 FRET pair, the FRET data were investigated by both theoretical and experimental approaches.

Steady-state and time-resolved anisotropy measurements for PIE-Cy3 RNA and PIE-Cy5 RNA showed that the values of the anisotropy were significant (PIE-Cy3, $r = 0.285 \pm 0.005$; PIE-Cy3 + proteins, $r = 0.305 \pm 0.004$; PIE-Cy5, $r = 0.28 \pm 0.007$; PIE-Cy5 + proteins, $r = 0.295 \pm 0.004$). Both dyes appear to be stacked on the ends of the duplexes and, therefore, could not be described as randomly reorienting in solution. Instead, the distribution of the orientation and the distance between the donor and acceptor are randomized in the isotropic static model (*24*). In this situation, $\kappa^2 = 0.476$. This model may be more appropriate for the PIE RNA, given that it has a distribution of structures in solution. The apparent distance $\langle R \rangle$ was calculated for both the dynamic average limit ($\langle R \rangle_{2/3}$) and the static average limit ($\langle R \rangle_{0.476}$) to assess its dependence on a model. As given in Table 1, the isotropic static model gives rise to a decrease of the apparent distances by 3 Å, well within experimental uncertainty.

In the previous study by Grainger et al. (*12*) PIE RNA was labeled with the fluorescein/Cy3 FRET pair. To compare the Cy3/Cy5 data with those data of Grainger et al., PIE RNAs labeled with fluorescein were also used in our experiments. The anisotropy of fluorescein in the RNA is small ($r = 0.064 \pm 0.001$), reflecting its free rotation. Although Cy3 anisotropy is high, indicating stacking of the fluorophore, the fluorescein behavior justifies the assumption of $\kappa^2 = {}^2/_3$ (*20*). Because our fluorescein labeling was 60% efficient, only time-resolved FRET experiments were possible, which measure the presence of a short decay component of the fluorescein fluorescence resulting from energy transfer to the acceptor.

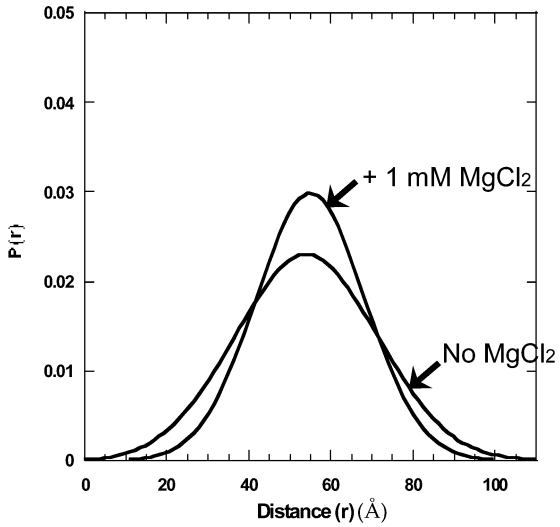

FIGURE 6: Analysis of the distance distributions by time-resolved FRET measurements as a function of Mg^{2+} ions for the free PIE RNA. PIE RNA, no Mg^{2+}: $\bar{r} = 54$ Å and fwhh = 39 Å. PIE RNA + Mg^{2+}: $\bar{r} = 55$ Å and fwhh = 29 Å.

Table 1: Steady-State FRET Efficiency (E) Measured and Apparent Distances $\langle R \rangle$ Calculated[a]

	E (%)	$\langle R \rangle_{2/3}$ (Å)	$\langle R \rangle_{0.476}$ (Å)
molecules/complexes, Cy3/Cy5 labeled			
free PIE RNA/no Mg^{2+}	35 (±2)	66.5 (±1)	63 (±1)
free PIE RNA/+Mg^{2+}	40 (±3)	64 (±2)	60.5 (±2)
PIE RNA−RBD1	55 (±5)	58 (±2)	55 (±2)
PIE RNA−U1A (WT)	68 (±3)	53 (±1)	50 (±1)
PIE RNA−U1A (IL)	68 (±2)	53 (±0.5)	50 (±0.5)
molecules/complexes, fluorescein/Cy3 labeled			
free PIE RNA/+Mg^{2+}	37.5	61	
PIE RNA−U1A (WT)	58	53	

[a] $\langle R \rangle_{2/3}$ is the apparent distance calculated for the measured efficiency of FRET, assuming the orientation factor $\kappa^2 = {}^2/_3$. $\langle R \rangle_{0.476}$ is the apparent distance calculated by the equation $\langle R \rangle_{0.476} = [\kappa^2/({}^2/_3)]^{1/6}\langle R \rangle_{2/3}$, assuming the orientation factor $\kappa^2 = 0.476$. The values reported for the efficiencies of FRET and the apparent distances represent the arithmetic mean of multiple measurements. The errors are the deviation from the maximum and minimum values from the means.

For the fluorescein/Cy3 PIE RNA, apparent distances were calculated using the equation $E = 1 - (\tau_{DA}/\tau_D) = R_0^6/(R_0^6 + \langle R \rangle^6)$, where $R_0 = 56$ Å. $\langle R \rangle$ values vary at most by 3 Å compared to the Cy3/Cy5 FRET results, demonstrating that the $\kappa^2 = {}^2/_3$ assumption for the $\langle R \rangle$ determination introduces little error into the calculation (Table 1). Distance distribution analyses for the Cy3/Cy5 time-resolved FRET data were compared to those for the fluorescein/Cy3 FRET pair; results are in good agreement (Figure 7).

There are several characteristics of the labeled RNA that ameliorate the assumption of the average value of κ^2. In the PIE RNA, time-resolved Cy3 fluorescence decays are highly heterogeneous, behavior which is often interpreted as an indication of multiple Cy3 conformations. This heterogeneity likely represents a different orientation of Cy3 with respect to the acceptor, and thus it decreases the effect of the κ^2

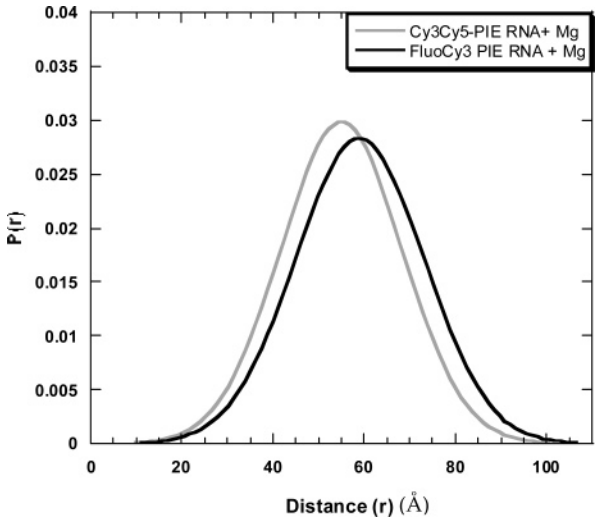

FIGURE 7: Distance distributions as a function of the FRET pair. Free PIE RNA-Cy3/Cy5 (gray line): $\bar{r} = 55$ Å and fwhh = 29 Å. Free PIE RNA-fluorescein/Cy3 (black line): $\bar{r} = 59$ Å and fwhh = 31 Å.

assumption on the distance calculation. Also, the R_0 values of both FRET pairs are close to the actual distances calculated for the PIE RNA, placing the experiments in the regime where large changes in the FRET efficiency give rise to small changes in the distance (*25*, *26*). Finally, the sixth root of κ^2 is used for distance calculations, reducing its contribution to the calculations. Therefore, the $\kappa^2 = {}^2/_3$ assumption will introduce minimal errors in the measurement of distances for the PIE RNA and PIE−protein complexes.

Local Structure and Dynamics. To study the local structure and dynamics of free PIE RNA, 2-aminopurine (2AP), a fluorescent analogue of guanosine and adenosine, was used as a site-specific probe. Two PIE RNA molecules were synthesized in which a single naturally occurring adenosine was replaced by 2AP: in stem 2, separating the box 1 and box 2 binding sites (PIE-2AP/St2), or in box 2 (PIE-2AP/box 2) (Figure 3).

(A) Local Structure and Dynamics in Stem 2. The free PIE-2AP/St2 fluorescence quantum yield is small (∼0.03) compared to that of free 2AP (0.65), as expected for a stacked base. Its fluorescence intensity does not change upon addition of 1 mM $MgCl_2$.

Although the fluorescence decay of free 2AP in solution is single exponential (∼10−11 ns), in the context of nucleic acid molecules, highly heterogeneous decays are typically observed (*27*). As anticipated, fluorescence decays of 2AP in stem 2 (Table 2, Supporting Information) were best fit to four exponentials with ∼70% of a very short component (∼50 ps) and only ∼4% of an ∼9 ns component. (We note that our instrument detection resolution does not permit a precise determination of a 50 ps component.) The 2AP fluorescence average lifetime is relatively short in every experimental condition, supporting the observation that 2AP fluorescence in stem 2 is quenched. Time-resolved and steady-state results are in good agreement.

Time-resolved anisotropy provides an indicator of motion of the fluorophore. The fitted initial anisotropy values vary between $r_0 = 0.32-0.35$, which is lower than the theoretical maximum value of 0.4. This very fast initial depolarization has been observed for other fluorophores (e.g., tryptophan $r_0 = 0.31$) and is attributed to internal photophysical processes and solvent relaxation that are too fast to be resolved by our instrument. Anisotropy decay curves were fit by two components: ∼70% of a short component, often attributed to segmental motion, and ∼30% of a longer component (Table 3, Supporting Information). This common interpretation, applied to this RNA, would indicate that 2AP in stem 2 has a very flexible conformation. However, the fluorescence intensity decays reveal that 2AP in stem 2 is highly quenched. 2AP stacking with other bases influences its electronic structure, resulting in changes in the internal photophysical process (internal conversion) and/or in charge delocalization (*28*). Those phenomena can also account for rapid fluctuations of the 2AP dipole moment orientation, giving rise to rapid depolarization. Therefore, in these molecules it seems more likely that most of the rapid depolarization is due to 2AP "wobbling" rather than free rotation (segmental motion). It is important to note that the short component (0.3 ns) is identical for all the PIE RNA-2AP/stem 2 samples.

The long anisotropy decay time is often attributed to the global tumbling of the molecule. In the PIE RNA, interpretation of 2AP anisotropy decays is complicated by the highly heterogeneous fluorescence decays and the global hydrodynamic behavior of the molecule. The observed 9 ns anisotropy decay, limited by the lifetime of 2AP fluorescence, is too short to reflect the tumbling of the entire molecule.

(B) Local Structure and Dynamics in Box 2. Box 2 comprises seven unpaired bases (AUUGCAC) and a single unpaired base (A) on the opposite strand, which in our experiments has been replaced by 2AP. Steady-state fluorescence experiments show that its quantum yield is 0.038 for the free RNA, indicating that 2AP fluorescence in box 2 is significantly quenched. A further 30% loss of fluorescence intensity is observed when 1 mM $MgCl_2$ is added to the RNA sample (quantum yield = 0.03).

Time-resolved fluorescence decays were fit with four exponentials revealing highly heterogeneous conformations for 2AP in box 2 (Table 4, Supporting Information). In the free PIE RNA, the decay times range from 26 ps to 9 ns, with 84% of the fluorescence decay associated with the very short components and only 3% with the longest component. The two intermediate time components of 0.78 and 4.2 ns account for 13% of the decay. This result suggests that a large majority of the 2AP adopts a stacked conformation. The quench of 2AP upon addition of Mg^{2+} is reflected by the increase of percentage (from 84% to 89%) of 2AP fluorescence associated with the very short decay component; all of the other lifetime decay components remain similar. The quenching observed upon Mg^{2+} addition calculated from the time-resolved fluorescence data (28%) and the steady-state data (30%) is in good agreement.

The time-resolved anisotropy data were fit to two exponentials (Table 5, Supporting Information), again with the caveat that the two correlation times determined should be interpreted rather as mean values of two unknown distributions. For the free RNA, the proportion of the anisotropy decay is about 70% of fast components and 30% of slower components (identical to 2AP in stem 2). Upon addition of Mg^{2+} the anisotropy decay does not noticeably change.

Our interpretation of the fluorescence data is that Mg^{2+} ion binding produces a more confined local structure for the

2AP in box 2, the implication being that Mg^{2+} ions bind within the PIE RNA internal loops.

Conformation and Dynamics of the PIE RNA−Protein Complexes: Global Conformation and Dynamics. Three different U1A constructs were bound to the PIE RNA, both to assess how the RNA global structure and dynamics changed and to look for any contribution of amino acids 101−121 (proposed to dimerize) to the RNA properties. RBD1 alone (amino acids 1−101) has been used in previous studies of PIE−protein interactions (*12*−*14*); it binds to the RNA but is inactive in regulation of polyadenylation (*8*). Two different full-length proteins were also bound: U1A (WT) and U1A (IL) which has the charged sequence (residues 101−121) and the sequence 121−140 inverted within the linker domain (Figure 2).

A complication for interpretation of FRET experiments was observed in control experiments using Cy3-PIE RNA in complexes with the proteins. Unexpectedly, Cy3 fluorescence intensity increased by 14% ± 4% when either full-length U1A protein was bound; no change occurred upon RBD1 binding. Either Cy3 directly interacts with a region of U1A outside the RBD1 domain or the protein could interact with the RNA duplex and displace Cy3 from its stacking interaction on the end of the stem. We note that the increase in fluorescence intensity could result from an increase in the quantum yield of Cy3 and so lead to an increase in the value of R_0. This effect is certain to be very small, however, since the contribution to R_0 varies as $(\Phi_{Cy3})^{1/6}$.

FRET efficiencies (E) and corresponding apparent distances ($\langle R \rangle$) are reported in Table 1. The FRET efficiency increases from 40% for the free RNA to 55% upon RBD1 binding, reducing the average distance between fluorophores from 64 to 58 ± 2 Å. A further increase in FRET efficiency to 68% occurs upon full-length U1A protein binding, reducing $\langle R \rangle$ to 53 ±1 Å. As previously shown by Grainger et al. (*12*), the PIE RNA undergoes significant global conformational changes upon RBD1 binding. The greater conformational change induced by the full-length U1A binding shows that other regions of the protein, not directly involved in RNA binding, can contribute to RNA structural changes.

On the basis of the suggestion of Gunderson et al. (*9*), it was possible that intermolecular dimer formation of the U1A sequences 103−119 was involved in the observed reduction in apparent distance between the ends of the RNA stems. Since this region of the U1A protein is just adjacent to the body of the complex, it was possible that it could further restrict the PIE RNA conformation and dynamic range of motion. However, a comparison of the PIE−U1A (WT) and PIE−U1A (IL) FRET data shows no difference in the measured FRET efficiencies. When an equimolar mixture of U1A (WT) and U1A (IL) was added to the PIE RNA, giving a 1:1:2 ratio of complexes containing U1A $(WT)_2$−U1A $(IL)_2$−[U1A (WT)/U1A (IL)], the measured FRET efficiency was identical to that of the other PIE−U1A complexes. These data suggest that if dimerization of amino acids 103−119 is occurring, it has no effect on the structure of the RNA.

The dynamics of the PIE−protein (1:2) complexes can be best described by a single distance distribution, derived from TR-FRET experiments (Figure 8, Table 7). For the PIE−RBD1 (1:2) complex, the Gaussian distribution has a mean distance at 40 Å and fwhh of 40 Å. For PIE−U1A (1:2) complexes, the mean distance of the distribution is shorter (34 Å); however, the fwhh becomes 46 Å for U1A (WT) and 31 Å for U1A (IL) (Figure 9). The mean distances are smaller than those determined from steady-state FRET data, although the trends are maintained. It is not clear if the difference in fwhh for the two U1A constructs is significant, due in part to experimental uncertainty and in part to the apparent interaction of Cy3 with U1A.

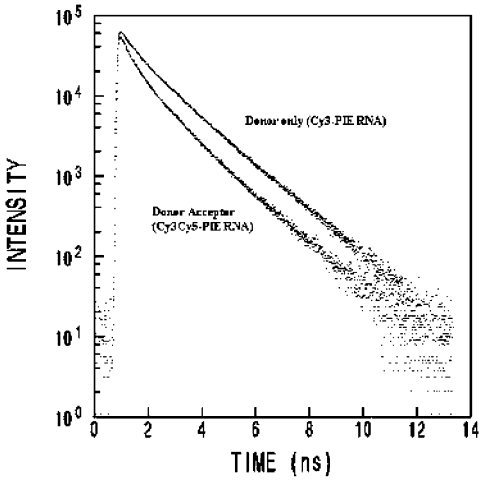

FIGURE 8: Time-resolved FRET experiments with the Cy3Cy5 FRET pair. Cy3 fluorescence decays for the donor-only (Cy3 RNA strand) and the donor−acceptor (Cy3Cy5 RNA strand) are shown. Buffer: 10 mM Tris (pH 8), 100 mM NaCl, 75 μg/mL BSA, and 1 mM $MgCl_2$, at 25 °C. [Cy3 RNA strand] = 50 nM; [Cy5 RNA strand] = 70 nM.

Local Conformation and Dynamics. The local structure and dynamics of PIE RNA bound to the different proteins have also been studied using 2AP.

(A) Local Structure and Dynamics in Stem 2. Neither the free PIE-2AP/St2 fluorescence quantum yield nor its fluorescence decay changes upon protein binding, in agreement with the steady-state fluorescence data. However, the 2AP emission maximum shifts two nanometers toward shorter wavelength when RBD1 domains are bound (data not shown). This subtle blue shift is not observed upon full-length U1A (WT) or U1A (IL) protein binding.

The time-resolved anisotropy data were fit to two exponentials (Table 3, Supporting Information). Every anisotropy decay shows ∼70% of a short component and ∼30% of a longer component. The short component (0.3 ns) remains constant for all of the PIE RNA molecules and complexes, indicating that the local fluctuations (wobbling of the stacked 2AP) are not perturbed upon protein binding. The long anisotropy decay component significantly increases upon RBD1 (10 ns) and full-length U1A protein (12 ns) binding, reflecting the increased size of both complexes. However, these values are too short to be a measure of the global tumbling time of the complex (the PIE−RBD1 complex is ∼55 kDa, and the PIE−U1A complexes are ∼100 kDa).

(B) Local Structure and Dynamics in Box 2. Steady-state fluorescence experiments give a quantum yield of 0.03 for the free PIE RNA-2AP in box 2. The 2AP fluorescence intensity increases by a factor of 1.4 ± 0.2 when RBD1 or

Table 7: Distance Distribution Parameters for Cy3Cy5-PIE RNA Free and Bound to RBD1, U1A (WT), and U1A (IL) Proteins[a]

molecules	$\bar{r}$ (Å)	fwhh (Å)	labeling efficiency	χ_R^2
Cy3Cy5-PIE, no Mg^{2+}	54 (53.8−54.1)	39 (38.7−39.5)	0.8 ± 0.05	1.6
Cy3Cy5-PIE + Mg^{2+}	55 (54.3−55.7)	29 (28.9−29.1)	0.8 ± 0.05	1.3
Cy3Cy5-PIE + RBD1 + Mg^{2+}	40 (39.6−40.7)	40 (39.3−40.6)	0.8 ± 0.05	1.5
Cy3Cy5-PIE + U1A (WT) + Mg^{2+}	34 (33.8−34.2)	46 (45.2−47)	0.8 ± 0.05	1.2
Cy3Cy5-PIE + U1A (IL) + Mg^{2+}	34 (32.6−34.4)	31 (30.7−31.2)	0.8 ± 0.05	1.2

[a] $\bar{r}$ = mean radius; fwhh = full width half-height. Analyses were performed assuming R_0 = 60 Å and $\kappa^2 = {}^2/_3$.

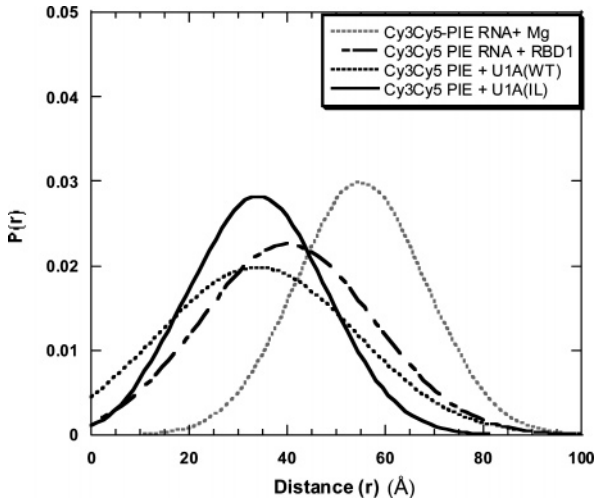

FIGURE 9: Analysis of the Gaussian distance distributions by time-resolved FRET measurements as a function of protein binding. Free PIE RNA (gray dotted line): $\bar{r}$ = 55 Å and fwhh = 29 Å. PIE RNA−RBD1 complex (black dashed line): $\bar{r}$ = 40 Å and fwhh = 40 Å. PIE RNA−U1A (WT) complex (black dotted line): $\bar{r}$ = 34 Å and fwhh = 46 Å. PIE RNA−U1A (IL) complex (black solid line): $\bar{r}$ = 34 Å and fwhh = 31 Å. Buffer: 10 mM Tris (pH 8), 100 mM NaCl, 75 μg/mL BSA, and 1 mM $MgCl_2$, at 25 °C. [Cy3 RNA strand] = 50 nM, [Cy5 RNA strand] = 70 nM, and [protein] = 120 nM.

full-length U1A (WT and IL) proteins bind to the RNA, indicating that its context has changed.

Time-resolved fluorescence decays of PIE-2AP/box 2 RNA with and without protein differ primarily in the amplitude and lifetime of the long component (Table 4, Supporting Information). The 2AP intensity-weighted average fluorescence lifetime increases from 6.1 ns to 6.67 and 7.4 ns when RBD1 and full-length U1A proteins are bound, respectively.

Time-resolved anisotropy data are most sensitive to formation of the complex (Table 5, Supporting Information). For the free RNA, the proportion of the anisotropy decay is about 70% of fast components and 30% of slower components. Upon protein binding the proportions become equal (50% each), suggesting that the 2AP in box 2 becomes less stacked upon complex formation.

DISCUSSION

Although other structural and fluorescence studies of the PIE RNA have been reported (*12*−*14*), the work described here uses time-resolved fluorescence measurements to obtain a description of the dynamics of the RNA with and without Mg^{2+} and with and without protein. Unlike those previous studies, the proteins used here include the full-length (282 amino acid) U1A protein, which is the biologically active species that forms the in vivo PIE−U1A complex.

The Free PIE RNA Structure. The presence of Mg^{2+} ions causes the PIE RNA to become more constrained. The Mg^{2+} ion titration data were fit by a two-state ligand binding model that assumes that ions have specific binding sites within the RNA and that only these binding events give rise to the signal change observed. Chelated ions satisfy this model since they are making specific contacts with the RNA (*29*). The number of sites determined from this model (2 ± 1) could correspond to the two internal loops, box 1 and box 2. This hypothesis is supported by the local RNA conformation study that shows that 2AP fluorescence in box 2 is sensitive to the presence of Mg^{2+} ions.

The bound Mg^{2+} could produce a conformation of the box that facilitates binding of the protein. Alternatively, Mg^{2+} could compete for protein binding. Since U1A RBD1 binds with nanomolar affinity to box 2 in conditions of 150 mM NaCl and 1 mM $MgCl_2$, pH 7 (*30*), it is likely that it evicts the Mg^{2+}. However, higher concentrations of Mg^{2+} substantially decrease the affinity of the protein for the RNA (*31*), suggesting that the ion can inhibit association.

If the Mg^{2+} ions are bound in a specific conformation formed by box 1 and box 2, it should be possible to probe these sites more quantitatively. For example, 8-hydroxyquinoline-5-sulfonic acid (HQS), a fluorescent probe used to quantify bound Mg^{2+} (D. E. Draper, personal communication), could more precisely determine the number of interacting Mg^{2+} ions. Electron paramagnetic resonance (EPR) spectroscopic methods could be used to examine the metal ions bound to their specific sites (*32*, *33*).

PIE−Protein Interactions. With the use of the three protein constructs in the fluorescence experiments, we have the opportunity to see how, if at all, the full-length U1A protein contributes to the structure of the RNA. Since the full-length protein is required for the inhibition of polyadenylation, the study of the PIE−U1A complex here probes this functional association.

(A) PIE−RBD1 Complex. Previous NMR structural studies used the PIE RNA together with RBD1 (amino acids 2−102). The NMR data indicated that the two helix C regions (amino acids 93−98) of the juxtaposed RBD1 proteins were in very close proximity to each other, forming a protein−protein interface (*34*). Those results also indicated that the C-terminal tails, including helix C, were in proximity to stem 2 and could form electrostatic interactions between side chains of amino acids and the phosphate backbone of the stem. Our 2AP fluorescence data provide another picture of stem 2 in this complex.

The emission spectrum of 2AP in stem 2 shows a 2 nm blue shift when RBD1 binds, which is absent when the full-length proteins are bound. The small but significant blue shift indicates that the RBD1 proteins change the local environment of the 2AP base; a blue shift is typically indicative of

a more hydrophobic environment. NMR data (*13*) place the basic side chains of Lys23 (helix A) and Arg47 (loop 3) in contact with A22 (2AP) and G23 in stem 2 in the PIE box 2−RBD1 complex, interactions that could be maintained in the complete PIE RNA−RBD1 complex. In the NMR structure of the PIE−RBD1 complex, the C-terminal tails (helix C) of the proteins are shown to interact with each other, via a hydrophobic interface. This interprotein interaction leaves the charged face of the helix pointing toward the RNA, in the vicinity of stem 2 (*34*). If this juxtaposition does account for the blue shift observed in our experiments, then its absence when the full-length proteins are bound suggests that there has been a change in the position of the two helix C sequences.

The 2AP in box 2 reports on any changes in the structure of this asymmetric loop upon ligand binding. NMR data indicate that, in the absence of Mg^{2+}, the bases in the loop are stacked but flexible (*14*). The adenosine on the strand that spans the loop nucleotides was shown to be stacked on the adjacent C•G base pair; this is the adenosine that we replaced with 2AP. Fluorescence results support this structure but also indicate that when Mg^{2+} is bound, there is a change in the conformation of this unpaired adenosine, likely an increase of its stacking. When RBD1 binds to box 2, NMR data (*13*) show that this spanning adenosine loses stacking with one G•C pair in an adjacent stem, while maintaining stacking with the other G•C pair. Again, the fluorescence data support this observation, based on the change in amplitudes of 2AP lifetimes.

(B) PIE−U1A Complexes. U1A protein binding reduces $\langle R \rangle$ and $\bar{r}$ between the ends of the PIE RNA arms by 5−6 Å beyond what was observed with RBD1 alone. The geometry of these bound complexes cannot be unambiguously described, however, since either or both the dihedral angle between arms and the bending angles within the boxes can be altered. Changes in the width of the distance distributions differ and depend on whether U1A (WT) or U1A (IL) is bound, but both distributions encompass a wide range of distances accessible to the RNA in the complexes.

An unexpected result was the observed 14% increase in Cy3 fluorescence upon U1A binding. This significant increase is not observed when RBD1 is bound, thus implicating an interaction between Cy3 and the remainder of the U1A protein as the cause of the fluorescence change. Assuming that Cy3, a hydrophobic dye, is packing in a portion of the protein, the obvious question is what region of the protein would be most plausible. Previous fluorescence experiments have shown that, in the PIE−U1A complex, the two RBD2 domains tumble independently from the core of the complex (*16*). Although it is possible that the nonpolar β-sheet surface of RBD2 could be the site of interaction, the more likely site(s) is (are) somewhere in the linker domain of the U1A protein.

This apparent interaction between Cy3 and full-length U1A could change the structure and dynamics of the RNA in the complexes or give rise to an artifact in the analysis of the data. While we cannot rule out the possibility that the dye makes a stable interaction with a portion of the U1A protein, the presumed association of Cy3 with the protein results in a change in the relative proportions of the components of the anisotropy decays of the dye: the short decay time (0.18 ns) contributes about 20% of the decay amplitude in the free PIE RNA but 40% in the U1A-bound complex (data not shown). This increase in the proportion of the short component suggests that Cy3 undergoes more local motion when U1A is bound.

The presence of an interaction between Cy3 at the end of a PIE RNA arm and a site on the protein (presumably within the linker) fortuitously provides some information on the juxtaposition of the full-length U1A protein and the RNA. When the PIE RNA is present in its pre-mRNA, it is upstream by two nucleotides of the polyadenylation signal (a stem loop structure recognized by the polyadenylation apparatus). The Cy3−U1A interaction indicates that the protein must wrap around (the linker is flexible) to reach the termini of the arms. Assuming that indeed it is the linker regions of the proteins that are in proximity to the arms of the PIE RNA, then, in the context of the pre-mRNA, where the polyadenylation signal stem loop essentially extends from one duplex arm of the PIE structure, the U1A linker would become immediately proximal to the polyadenylation stem loop. This places the linker (and RBD2) in position to interact with any component of the polyadenylation machinery it encounters.

Summary. These experiments show that the stems of PIE RNA are flexible relative to each other, with a mean separation of 55 Å but with 90% of the population having a separation of 55 ±15 Å (the fwhh of a Gaussian distribution, or 2.345σ). When RBD1 binds, 90% of the RNA molecules have conformations described by a distance distribution of 40 ± 20 Å, and when U1A (WT) is bound, these statistics change to 34 ± 23 Å. Binding of U1A (IL) reduces the width of the distribution (34 ± 15 Å), possibly due to the altered position of the 101−121 amino acid sequence but possibly caused by uncharacterized photophysical properties of Cy3 as it interacts with the protein. We interpret the distributions in terms of the dynamic motion of the RNA structure.

ACKNOWLEDGMENT

We thank Tom Stump for help with TCSPC experiments and Dr. Samuel Gunderson (Rutgers University) for communication of the results for the polyadenylation and U1A−PIE RNA binding assays.

SUPPORTING INFORMATION AVAILABLE

Time-resolved fluorescence and anisotropy parameters for the various PIE RNA and PIE RNA−protein complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Sillikens, P. T. G., Habets, W. J., Beijer, R. P., and van Venrooij, W. J. (1987) cDNA cloning of the human U1 snRNA-associated A protein: Extensive homology between U1 and U2 snRNP-associated proteins, *EMBO J. 6*, 3841−3848.
2. Scherly, D., Boelens, W., van Venrooij, W. J., Dathan, N. A., and Mattaj, I. W. (1989) Identification of the RNA binding segment of human U1A protein and definition of its binding site on U1 snRNA, *EMBO J. 8*, 4163−4170.
3. Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A, *Nature 348*, 515−520.
4. Lu, J., and Hall, K. B. (1997) Tertiary structure of RBD2 and backbone dynamics of RBD1 and RBD2 of the human U1A protein determined by NMR spectroscopy, *Biochemistry 36*, 10393−10405.

5. Lu, J., and Hall, K. B. (1995) An RBD that does not bind RNA: NMR secondary structure determination and biochemical properties of the C-terminal RNA binding domain from the human U1A protein, *J. Mol. Biol. 247*, 739−752.
6. Jean, J. M., Clerte, C., and Hall, K. B. (1999) Global and local dynamics of the human U1A protein determined by tryptophan fluorescence, *Protein Sci. 8*, 2110−2120.
7. 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.
8. Gunderson, S. I., Beyer, K., Martin, G., Keller, W., Boelens, W. C., and Mattaj, I. A. (1994) The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase, *Cell 76*, 531−541.
9. Gunderson, S. I., Vagner, S., Polycarpou-Schwarz, M., and Mattaj, I. W. (1997) Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and the coupling of splicing and polyadenylation, *Genes Dev. 11*, 761−773.
10. van Gelder, C. W. G., Gunderson, S. I., Jansen, E. J. R., Boelens, W. C., Polycarpou-Schwarz, M., Mattaj, I. W., and van Venrooij, W. J. (1993) A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A is required for regulation of polyadenylation, *EMBO J. 12*, 5191−5200.
11. Grainger, R. J., Murchie, A. I. H., Norman, D. G., and Lilley, D. M. J. (1997) Severe axial bending of RNA introduced by the U1A binding element present in the 3′ untranslated region of the U1A mRNA, *J. Mol. Biol. 273*, 84−92.
12. Grainger, R. J., Norman, D. G., and Lilley, D. M. J. (1999) Binding of U1A protein to the 3′ untranslated region of its pre-mRNA, *J. Mol. Biol. 288*, 585−594.
13. Howe, P. W., Allain, F. H., Varani, G., and Neuhaus, D. (1998) Determination of the NMR structure of the complex between U1A protein and its RNA polyadenylation inhibition element, *J. Biomol. NMR 11*, 59−84.
14. Gubser, C. C., and Varani, G. (1996) Structure of polyadenylation regulatory element of the human U1A pre-mRNA 3′-untranslated region and interaction with the U1A protein, *Biochemistry 35*, 2253−2260.
15. Klein Gunnewiek, J. M., Hussein, R. I., van Aarssen, Y., Palacios, D., de Jong, R., van Venrooij, W. J., and Gunderson, S. I. (2000) Fourteen residues of the U1 snRNP-specific U1A protein are required for homodimerization, cooperative RNA binding, and inhibition of polyadenylation, *Mol. Cell. Biol. 20*, 2209−2217.
16. Clerte, C., and Hall, K. B. (2000) Spatial Orientation and dynamics of the U1A proteins in the U1A-UTR complex, *Biochemistry 39*, 7320−7329.
17. Shortle, D., DiMaio, D., and Nathans, D. (1981) Directed mutagenesis, *Annu. Rev. Genet. 15*, 265−294.
18. Gruber, H. J., Hahn, C. D., Kada, G., Riener, C. K., Harms, G. S., Ahrer, W., Dax, T. G., and Knaus, H. G. (2000) Anomalous fluorescence enhancement of Cy3 and cy3.5 versus anomalous fluorescence loss of Cy5 and Cy7 upon covalent linking to IgG and noncovalent binding to avidin, *Bioconjugate Chem. 11*, 696−704.
19. Johnson, M. L., and Faunt, L. M. (1992) Parameter estimation by least-squares methods, *Methods Enzymol. 210*, 1−37.
20. Lakowicz, J. R. (1988) *Principles of Fluorescence Spectroscopy*, 2nd ed., Plenum, New York.
21. Haas, E., Wilchek, M., Katchalski-Katzir, E., and Steinberg, I. Z. (1975) Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer, *Proc. Natl. Acad. Sci. U.S.A. 72*, 1807−1811.
22. Albaugh, S., Lan, J. Q., and Steiner, R. F. (1989) The effect of a distribution of separations upon intramolecular distances in biopolymers, as determined by radiationless energy transfer, *Biophys. Chem. 33*, 71−76.
23. Rice, K. G., Wu, R. G., Brand, L., and Lee, Y. C. (1991) Interterminal distance and flexibility of a triantennary glycopeptide as measured by resonance energy transfer, *Biochemistry 30*, 6646−6655.
24. Eftink, M. R. (1991) Fluorescence techniques for studying protein structure, *Methods Biochem. Anal. 35*, 127−205.
25. Cantor and Schimmel (1980) *Biophysical Chemistry, Part II: Techniques for the Study of Biological Structure and Function*, W. H. Freeman and Co., San Francisco.
26. Wu, P., Lee, K. B., Lee, Y. C., and Brand, L. (1991) Solution conformations of a biantennary glycopeptide and a series of its exoglycosidase products from sequential trimming of sugar residues, *J. Biol. Chem. 271*, 1470−1474.
27. Guest, C. R., Hochstrasser, R. A., Dupuy, C. G., Allen, D. J., Benkovic, S. J., and Millar, D. P. (1991) Interaction of DNA with the Klenow fragment of DNA polymerase I studied by time-resolved fluorescence spectroscopy, *Biochemistry 30*, 8759−8770.
28. Jean, J. M., and Hall, K. B. (2001) 2-Aminopurine fluorescence quenching and lifetimes: Role of bases stacking, *Proc. Natl. Acad. Sci. U.S.A. 98*, 37−41.
29. Draper, D. E. (2004) A guide to ions and RNA structure, *RNA 10*, 335−343.
30. Williams, D. J., and Hall, K. B. (1996) Thermodynamic comparison of the salt dependence of natural RNA hairpins and RNA hairpins with nonnucleotide spacers, *Biochemistry 35*, 14665−14670.
31. Hall, K. B., and Stump, W. T. (1992) Interaction of N-terminal domain of U1A protein with an RNA stem/loop, *Nucleic Acids Res. 20*, 4283−4290.
32. Morrissey, S. R., Horton, T. E., Grant, C. V., Hoogstraten, C. G., Britt, R. D., and DeRose, V. J. (1999) Mn^{2+}-nitrogen interactions in RNA probed by electron spin−echo envelope modulation spectroscopy: application to the hammerhead ribozyme, *J. Am. Chem. Soc. 121*, 9215−9218.
33. Hoogstraten, C. G., Grant, C. V., Horton, T. E., DeRose, V. J., and Britt, R. D. (2002) Structural analysis of metal ion ligation to nucleotides and nucleic acids using pulsed EPR spectroscopy, *J. Am. Chem. Soc. 124*, 834−842.
34. Varani, L., Gunderson, S. I., Mattaj, I. W., Kay, L. E., Neuhaus, D., and Varani, G. (2000) The NMR structure of the 38 kDa U1A protein−PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein, *Nat. Struct. Biol. 7*, 329−335.

BI049117G