Origins of Binding Specificity of the A1 Heterogeneous Nuclear Ribonucleoprotein†

Norzehan Abdul-Manan,‡ Shawn M. O'Malley,‡ and Kenneth R. Williams*,§

Department of Molecular Biophysics and Biochemistry and Howard Hughes Medical Institute, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510-8024

Received September 26, 1995; Revised Manuscript Received January 5, 1996⊗

ABSTRACT: The A1 heterogeneous nuclear ribonucleoprotein (hnRNP) is the best studied of the "core" hnRNP proteins that are tightly associated with heterogeneous nuclear RNA (hnRNA) within eukaryotic nuclei. Previous studies suggested that hnRNP A1 preferentially binds (under nonequilibrium conditions) to the pyrimidine-rich span of sequence at the 3′-splice site of most introns [Swanson, M. S., & Dreyfuss, G. (1988) *EMBO J.* 11, 3519-3529; Buvoli *et al*. (1990) *Nucleic Acids Res.* 18, 6595-6600; Ishikawa *et al*. (1993) *Mol. Cell. Biol. 13,* 4301−4310]. Recently, Burd and Dreyfuss [(1994) *EMBO J. 13,* 1197−1204] used selection/amplification from pools of random sequence RNA to uncover an even higher-affinity A1 oligo that contained two copies of a high-affinity consensus sequence, UAGGGU/A. We have extended these studies by using a fluorescence assay to characterize the equilibrium binding properties of A1 to each of these oligonucleotides. By also characterizing the binding of A1 to sequence-randomized control oligonucleotides, we have been able to better evaluate the inherent "sequence-specific" binding properties of A1. Although these studies indicate that under equilibrium conditions A1 *cannot* specifically recognize the β-globin, 3′-splice site DNA oligo analogue studied by Buvoli *et al*. (1990), they confirmed the high-affinity binding to the "winner" 20-mer RNA that was uncovered via selection/amplification and that has the sequence UAUGAUAGGGACUUAGGGUG (Burd & Dreyfuss, 1994). In 0.1 M NaCl, we found that A1 has ∼100-fold higher affinity for this winner sequence than it does for either a randomized version of this sequence or a 20-mer oligo corresponding to an unrelated β-globin intron sequence. This winner RNA oligo aggregates in solution to form an apparent dimer that may represent a G-quartet resulting from dimerization of two Hoogsteen base-paired hairpins. On the basis of salt sensitivity studies carried out with various fragments of A1, the ability of A1 to discriminate the winner sequence from its randomized control results primarily from increased ionic interactions with the glycine-rich, COOH terminal domain of A1 that extends from residue 196 to 319. Nonetheless, most of the *overall* energy of binding for the A1 winner complex results from determinants that are resident within the first 195 residues of A1. The unique ability of the winner sequence (but not its sequence-randomized control) to form a higher-order aggregate, which may correspond to a G-tetrad, appears to facilitate the additional ionic interactions with the COOH terminal domain. Taken together, these data suggest the need to reevaluate possible and probable functions of A1 *in vivo.*

Proteins that bind pre-mRNA transcripts and are not stable components of other ribonucleoprotein complexes such as snRNPs are termed heterogeneous nuclear ribonucleoproteins (hnRNPs) (Dreyfuss *et al*., 1993). There are at least 20 different hnRNP and six of the more abundant migrate on sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) as three doublets named A1/A2, B1/B2, and C1/C2 (Beyer *et al*., 1977). Of these six "core" proteins, A1 hnRNP is the best characterized. Analysis of the sequence of A1 (Merrill *et al*., 1986) and limited proteolysis (Kumar *et al*., 1986) both suggested A1 contains two major domains, the 1−195 "UP1" region and the glycine-rich, 196−319 COOH terminal domain. Within the NH_2 terminal, two-thirds of A1 is a region of internal sequence homology such that, when residues 3−93 are aligned with 94−194, 32% of the residues are identical (Merrill *et al*., 1986). On the basis of the high degree of conservation of basic and aromatic residues in these ∼90-residue domains (62 and 80%, respectively), these two internal repeats were postulated to represent two independent nucleic acid binding domains (Merrill *et al*., 1986). Subsequently, homologous RNA binding domains were found in the yeast poly(A) binding protein (Adam *et al*., 1986) and now in more than 100 other eukaryotic RNA binding proteins (Birney *et al*., 1993). This ∼90-residue domain has now been referred to as the RNP RNA binding domain (RBD, Dreyfuss *et al*., 1988) or the RNA recognition motif (RRM, Query *et al*., 1989), and proteins have been found that have from one (type C hnRNP) to as many as four [nucleolin and the poly(A) binding protein] of these domains (Bandziulis *et al.*, 1989; Merrill & Williams, 1990; Kenan *et al*., 1991; Shamoo *et al.*, 1995). Although the extent of sequence identity among RBDs is low, there are two short consensus sequences, RNP1 and RNP2, that are more highly conserved and that are located about 30 residues apart (Dreyfuss *et al*., 1988). On the basis of X-ray crystallographic and NMR studies on RBDs from U1A and the type A and C hnRNP proteins (Nagai *et al*., 1990; Hoffmann *et al*., 1991; Wittekind *et al*., 1992), all RBDs contain a β1-α1-β2-β3-α2-β4 structure that results in a four-stranded β-sheet "platform" backed by two helices. The conserved RNP1 and RNP2 sequences are juxtaposed

† This study was supported by grants from the NIH (GM31539 to K.R.W.) and from the Swebelius Foundation (to S.M.O.).

‡ Department of Molecular Biophysics and Biochemistry.

§ Howard Hughes Medical Institute.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1996.

0006-2960/96/0435-3545$12.00/0

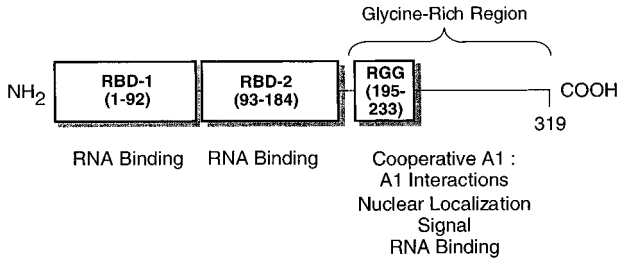

FIGURE 1: Schematic representation of structural and functional domains in the hnRNP 1 protein. The A1 fragments used in binding studies include 1−92 and 93−184, which correspond to RBD-1 and RBD-2, respectively; 1−184, which contains both RBDs; 1−195, which has also been referred to as UP1 (Herrick & Alberts, 1976) and which contains both RBDs as well as an 11-residue sequence that spans residues 185−195; and a COOH terminal synthetic peptide analogue which corresponds to residues 260−307 and which thus represents 35% of the COOH terminal domain extending from residue 185 to 319. Within the COOH terminal domain is an interesting region (labeled RGG) spanning residues 195−233 that contains four Arg-Gly-Gly repeats that together constitute an RGG box (Kiledjian & Dreyfuss, 1992; Dreyfuss *et al.*, 1993).

in the middle of the adjacent central antiparallel β-strands 3 and 1, respectively. The first data regarding the location of the RNA binding site in an RBD came from photo-cross-linking studies which demonstrated that phenylalanines 16 and 58 in A1 RBD-1 (and the corresponding phenylalanines in A1 RBD-2) are at the interface of the A1/nucleic acid complex (Merrill *et al*., 1988). These two phenylalanine residues are located in the A1 RNP2 and RNP1 sequences, respectively (Dreyfuss *et al*., 1988). The crystal structure of U1A RBD-1 complexed with its hairpin target RNA (stem-loop II of U1 snRNA) demonstrates that the U1 RNA loop is splayed across the surface of the β-sheet platform and that the loop does indeed interact extensively with RNP1 and RNP2 (Oubridge *et al*., 1995). The U1A residues (tyrosine 13 and phenylalanine 56) that correspond with the cross-linking sites in A1 RBD-1 are located next to each other on β-strands 1 and 3 (in RNP2 and RNP1, respectively) and are near the center of the β-sheet platform that forms a major part of the binding surface of the RBD. As predicted from the cross-linking studies on A1 hnRNP (Merrill *et al*., 1988), both tyrosine 13 and phenylalanine 56 in U1A are involved in base-stacking interactions with U1 RNA (Oubridge *et al*., 1995).

The COOH terminal one-third of A1 (residues 185−319) is not homologous with its NH_2 terminal two-thirds. The 135 residues in the COOH terminal domain of A1 have an unusually high glycine content (45%) and a reasonably well-conserved repeat sequence, GNF/YGGS/GRG, that provides a regular spacing of flexible, aromatic, and positively charged residues (Wilson *et al.*, 1987; Cobianchi *et al*., 1988). Within the COOH terminal domain (see Figure 1) is an interesting region spanning residues 195−233 that contains four Arg-Gly-Gly repeats that together constitute an RGG box (Kiledjian & Dreyfuss, 1992; Dreyfuss *et al.*, 1993). Numerous other RNA binding proteins, such as splicing factors, hnRNPs, and RNA helicases, contain RGG repeats (interspersed with aromatic amino acids) that have a spacing similar to that found in A1 (Birney *et al.*, 1993). On the basis of deletion studies, the 22-residue RGG motif in hnRNP U may represent the entire RNA binding domain of this hnRNP protein (Kiledjian & Dreyfuss, 1992). The COOH terminal domain of A1 also contains a nuclear localization signal that spans residues 259−288 (Weighardt *et al.*, 1995).

Each of the A1 RBDs as well as its COOH terminal domain bind nucleic acids (Merrill *et al.*, 1988; Kumar *et al.*, 1990; Nadler *et al.*, 1991; Shamoo *et al.*, 1994). Although the binding energies of the two A1 RBDs are not additive (Shamoo *et al*., 1994, 1995), together they contribute ~50% of the free energy of A1 binding to an extended single-stranded (ss) lattice (Shamoo *et al*., 1994). Under physiological salt concentrations, this corresponds to an affinity of only about 5×10^4 M^{-1} for poly[r(ϵA)] (Nadler *et al*., 1991), which is too low to be detected by most nonequilibrium binding assays. The remaining A1 binding energy derives from cooperative A1/A1 and direct A1/nucleic acid interactions that are resident in the glycine-rich COOH terminal domain (Kumar *et al*., 1990; Nadler *et al*., 1991; Shamoo *et al*., 1994).

In addition to its presumed role in pre-mRNA packaging and transport, A1 hnRNP appears to also have a number of other activities that might be of biological importance. Both *in vitro* and *in vivo* studies demonstrate that A1 has the potential to influence 5′-splice site selection in pre-mRNAs that contain multiple 5′-splice sites (Mayeda & Krainer, 1992; Cáceres *et al*., 1994; Yang *et al.*, 1994). Several reports (Pontius & Berg, 1990; Kumar & Wilson, 1990; Munroe & Dong, 1992) demonstrate that A1 promotes renaturation of complementary single-stranded nucleic acids. This strand-annealing activity of A1 is localized in its COOH terminal domain (Kumar & Wilson, 1990; Munroe & Dong, 1992) and appears to be modulated by phosphorylation of serine 198 within this domain (Cobianchi *et al*., 1993).

To understand the full range of A1 functions in pre-mRNA biogenesis, it is important that we quantitatively understand the ability of A1 to selectively bind high-affinity RNA targets and that we elucidate determinants of this specificity. Previous studies indicated that hnRNP A1 exhibits preferential binding to the pyrimidine-rich, 3′-splice site of most introns (Swanson & Dreyfuss, 1988; Buvoli *et al*., 1990; Ishikawa *et al*., 1993). Mutations of the conserved AG dinucleotide at the 3′-splice site appeared to significantly decrease A1 affinity (Burd and Dreyfuss, 1988; Buvoli *et al*. 1990). Recently, Burd and Dreyfuss (1994) utilized selection/amplification from pools of random sequence RNA to identify high-affinity target sequences for A1. They found that the sequence UAGGGU/A represents a high-affinity A1 consensus sequence and that this sequence resembles consensus sequences for vertebrate 5′- and 3′-splice sites. The highest-affinity 20-mer "winner" sequence identified for A1 contained a duplication of this 6-mer sequence, and it bound 300-fold more tightly than an unrelated β-globin intron sequence did to A1 (Burd & Dreyfuss, 1994).

We have extended previous studies on sequence-specific binding of A1 by quantitating (under equilibrium conditions) the degree of A1 specificity for the highest-affinity oligos identified by Buvoli *et al*. (1990) and Burd and Dreyfuss (1994). In the case of the winner sequence identified by Burd and Dreyfuss (1994), we have also begun to identify the protein and RNA determinants that account for this specificity.

MATERIALS AND METHODS

Purification of A1 hnRNP and Its Fragments. A1 was overexpressed and purified as described in Cobianchi *et al*.

(1988). The UP1 fragment of A1 (residues 1−195) was overexpressed and purified as described in Shamoo *et al*. (1994). A1 fragments corresponding to residues 1−184, 1−92, and 93−184 were made by partial cleavage of UP1 with *Staphylococcus aureus* V8 protease (Pierce) and were purified as described in Shamoo *et al*. (1994). The COOH terminal peptide corresponding to residues 260−307 in A1 was synthesized, purified, and characterized in the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University as described (Nadler *et al*., 1991).

Nucleic Acids. Oligos were synthesized in the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University and were purified via reverse phase and anion exchange high-performance liquid chromatography (HPLC) as described in Webster *et al*. (1991). Poly[r(A)] was from Pharmacia and was ethenylated as described by Nadler *et al*. (1991). Matrix-assisted laser desorption mass spectrometry was carried out by Mass Metrics at the University of Wisconsin at Madison.

Fluorescence Spectroscopy. Fluorescence spectroscopy was used to quantitate binding affinities. Generally, this approach can be carried out either by titration of the protein against a fixed concentration of nucleic acid (forward titration) or by addition of nucleic acid to a fixed concentration of protein (reverse titration) [see Kowalczykowski *et al*. (1986) for a review]. In the present work, fluorescence titrations were carried out in 2 mL, temperature-regulated, and continuously stirred cuvettes on an SLM 8000C spectrofluorometer interfaced to an HP Vectra computer. Three 10 s acquisitions were averaged for each data point. Forward titrations with poly[r(ϵA)] were performed with excitation and emission wavelengths of 315 and 400 nm, respectively. Corrections were made for background fluorescence and dilution effects due to addition of protein. Reverse titrations were carried out with excitation and emission wavelengths of 280 and 325 nm, respectively. Reverse titrations were corrected for background fluorescence, dilution, protein photobleaching, and inner filter effects resulting from the addition of nucleic acids. Inner filter corrections were calculated from third-order regression analysis of curves obtained from the titration of *N*-acetyltryptophanamide with the appropriate oligo. An 8 nm band-pass was used for the emission and excitation wavelengths for all titrations.

On the basis of previous studies (Nadler *et al*., 1991; Shamoo *et al*., 1994), the following occluded site sizes (n) were assumed: A1, $n = 14-28$ depending upon the salt concentration; UP1, $n = 10$; 1−184, $n = 14$; RBD-1 (1−92) and RBD-22 (93−184), $n = 6$; and COOH terminal peptide (260−307), $n = 4$. For those poly[r(ϵA)] titrations carried out in the presence of oligo competitors, it was assumed that the free protein concentration is identical at the same extent of fluorescence enhancement of poly[r(ϵA)] in the absence and presence of the oligo competitor. This assumption allows for the determination of the apparent affinity of the competing oligonucleotide (K_{comp}) using the expression

$$K_{comp} = [comp]_{bound} \times [poly[r(\epsilon A)]]_{free} \times K_{poly[r(\epsilon A)]}/[[comp]_{free} \times [poly[r(\epsilon A)]]_{bound}]$$

In this expression, the $[comp]_{bound}$ is calculated from the difference between $[protein]_{total}$ and the sum of $[protein]_{free}$ and $[poly[r(\epsilon A)]]_{bound}$ (calculated in terms of binding site concentration bound at that point in the titration). The $[protein]_{free}$ at that point of fluorescence enhancement is obtained from an identical titration carried out in the absence of competitor.

For reverse titrations, the maximum percent quenching ($\%Q_{max}$) for those oligo titrations that did not reach saturation was estimated using double reciprocal analysis. The y-intercept from a double reciprocal plot of $1/\Delta F$ versus $1/[oligo]_{total}$ gives $1/\Delta F_{max}$. Once the latter value has been established, the free ligand concentration can be calculated at each point in the titration and the data can be plotted as $1/\Delta F$ versus $1/[oligo]_{free}$. The apparent oligo affinities can then be calculated using the formula $K_{app} = 1/(\text{slope} \times \Delta F_{max})$ (Kelly *et al*., 1976).

UV Spectroscopy/Thermal Denaturation Studies. Thermal transitions for oligos were recorded on a Perkin-Elmer lamda 6 model spectrophotometer equipped with a thermally controlled cuvette holder in 20 mM potassium phosphate (pH 6.0). Absorbances were recorded at 260 nm from 15 to 80 °C using a 1 °C/min rate of heating and an acquisition time of 30 s. The absorbance values reported represent the differences between the values obtained for oligonucleotide in the buffer versus those for the buffer alone.

RESULTS

Under Equilibrium Conditions A1 Does Not Have Significant Binding Specificity for an Oligonucleotide Derived from the 3′-Splice Site of IVS1 of β-Globin Pre-mRNA. Using an assay based on irreversible photo-cross-linking, Buvoli *et al*. (1990) found that A1 hnRNP had an approximately 5-fold higher affinity for a DNA analogue of the 3′-splice site of IVS1 of β-globin pre-mRNA (oligo 4 in Table 1) than for several unrelated DNA oligos and that changing the invariant 3′-splice site from AG to GG decreased the affinity of A1 by 20-fold. Similar results had previously been obtained by Swanson and Dreyfuss (1988), who relied on a nonequilibrium assay based on immunopurification of A1 hnRNP bound to RNase T1-generated fragments. To quantify the sequence-specific binding properties of A1 under equilibrium conditions, we first examined the interaction of A1 with RNA and DNA oligonucleotides derived from the 3′-splice site of IVS1 of β-globin pre-mRNA (Buvoli *et al*., 1990). We initially carried out these studies using reverse titrations where oligo-induced quenching of intrinsic A1 fluorescence is used to monitor binding. As controls, we have used the same mutant oligonucleotide described previously (Buvoli *et al*., 1990) in which the invariant AG (underlined in oligos 1 and 4, Table 1) has been changed to GG (oligos 2 and 5, Table 1) as well as a sequence-randomized version (oligos 3 and 6, Table 1) of the wild type oligonucleotide. Parts A and B of Figure 2 show reverse titration curves for A1 binding to these three RNA oligos and their ssDNA analogues. The calculated K_{app} values for A1 binding to the RNA oligos in 100 mM NaCl were as follows: wild type (i.e. oligo 1) = 1.0×10^6 M^{-1}, mutant (i.e. oligo 2) = 6.5×10^5 M^{-1}, and random (i.e. oligo 3) = 1.6×10^6 M^{-1}. Similarly, no significant difference in A1 affinity was observed upon mutation or randomization of the sequence of the corresponding DNA analogues of these oligos (Figure 2B). To further verify these results, we employed a competition assay based on a forward fluorescence titration where the binding of A1 to poly[r(ϵA)] is

Table 1: Oligonucleotides Used for Fluorescence Binding Studies with A1 hnRNP and its Fragments

oligo	description	sequence	reference
1	β-globin pre-mRNA with the invariant AG at the 3′-splice site underlined	UCUAUUUUCCCACCCUUAG	this work
2	mutant version of oligo 1 in which the invariant 3′-splice site AG is changed to GG	UCUAUUUUCCCACCCUUGG	this work
3	sequence-randomized version of oligo 1	UCCACUAUUACCCGUUUCU	this work
4	DNA analogue of oligo 1	TCTATTTTCCCACCCTTAGGT	Buvoli *et al.* (1990)
5	mutant version of oligo 4 where the invariant 3′-splice site AG is changed to GG	TCTATTTTCCCACCCTTGGGT	Buvoli *et al.* (1990)
6	sequence-randomized version of oligo 4	TCCACTATGTACCCCTTTGTT	this work
7	highest-affinity A1 selected 20-mer with two consensus sequences underlined (i.e. winner sequence)	UAUGAUAGGGACUUAGGGUG	Burd and Dreyfuss (1994)
8	first sequence-randomized version of oligo 7 (i.e. random 1)	UGCUGAUGUUGAUGAGAGAG	this work
9	second sequence-randomized version of oligo 7 (i.e. random 2)	GGAACGUGUGUAUGAUGAUG	this work
10	third sequence-randomized version of oligo 7 (i.e. random 3)	UGAUGAAUGAUGGUGAUGGC	this work
11	20-mer from a span of β-globin intron pre-mRNA not containing a splice site	GAUCACUUGUGUCAACACAG	Burd and Dreyfuss (1994)
12	DNA analogue of oligo 7 (highest-affinity A1 selected 20-mer with two consensus sequences underlined)	TATGATAGGGACTTAGGGTG	this work
13	DNA analogue of oligo 8 (i.e. random 1)	TGCTGATGTTGATGAGAGAG	this work
14	DNA analogue of oligo 11 (20-mer from β-globin pre-mRNA)	GATCACTTGTGTCAACACAG	this work

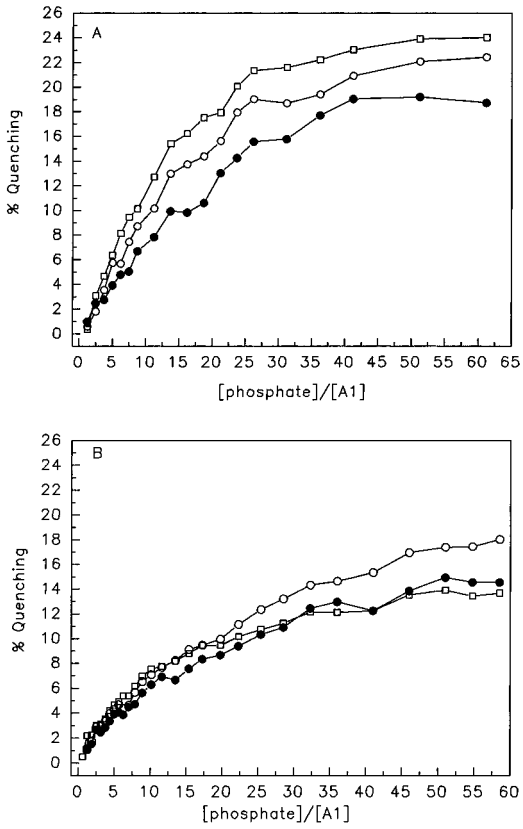

FIGURE 2: Reverse titrations of 1 μM A1 with increasing amounts of RNA (A) and corresponding DNA (B) oligos containing a β-globin 3′-splice site. All titrations were carried out in 10 mM Tris/HCl (pH 7.4) and 100 mM NaCl at 25 °C, and in each case, A1 was titrated with increasing amounts of wild type (open circles), mutant (closed circles), and sequence-randomized (open squares) oligonucleotides. The 19-mer RNAs used in panel A include oligos 1−3 in Table 1 with oligo 1 representing the wild type sequence that ends with the invariant AG at the 3′-splice site. The 21-mer DNAs used in panel B include oligos 4−6 in Table 1 with oligo 4 representing the same DNA analogue of the β-globin 3′-splice site used by Buvoli *et al.* (1990). Oligos 1 and 4 are the same except that oligo 4 extends two nucleotides beyond the invariant AG and oligo 4 is a DNA analogue of oligo 1.

competed for by an RNA oligonucleotide that is not ethenylated. The A1 apparent affinities determined under identical buffer conditions via this competition assay for the wild type (oligo 1), mutant (oligo 2), and random (oligo 3) oligos were similar to those obtained from reverse titrations, i.e. 1.1×10^6, 6.3×10^5, and 4.5×10^5 M^{-1}, respectively (data not shown). Since the approximately 2-fold range in apparent affinities observed for these oligos is within the normal variation in replicate determinations, we conclude that (under equilibrium conditions) A1 cannot differentiate between them. When similar binding studies were repeated with UP1 (residues 1−195), as the specificity determinants for the DNA analogue of this β-globin splice site sequence had been reported to reside within UP1 (Buvoli *et al.*, 1990), this A1 fragment was also found to be unable to discriminate between these β-globin-derived oligos (data not shown).

The High-Affinity Winner RNA Oligo Identified by Burd and Dreyfuss (1994) Aggregates in Solution. Anion exchange HPLC purification of the 20-mer winner RNA (oligo 7 in Table 1) was complicated by the apparent ability of this oligo to aggregate in solution. That is, the HPLC profile showed an early-eluting and a broad, late-eluting peak that we believe correspond to monomer and aggregate species, respectively. On the basis of relative peak area, there was approximately twice as much aggregate as monomer present. Under strongly denaturing conditions (i.e. 5 M urea and an ammonium phosphate gradient on a column heated to 65 °C), the relative size of the monomer peak increased at the expense of the aggregate, which still, however, accounted for ∼30% of the total eluted peak area. Matrix-assisted laser desorption mass spectral analysis demonstrated that both peaks had similar masses, that is, 6511 and 6523 for the monomer and aggregate, respectively (data not shown). Consistent with these data, MALDI analysis of the mixture of these species (prior to anion exchange HPLC) provided a single peak with a mass of 6495 as compared to the predicted (monoisotopic) mass of 6488. As expected, when the later-eluting fraction was diluted and reinjected onto the column, the profile showed a major early peak and a minor late peak, thus indicating the two species are in equilibria. Since it appeared that the late-eluting winner peak was an aggregate of the early-eluting peak, unless otherwise stated, the following studies were carried out on the mixture of the two species that existed prior to anion exchange HPLC.

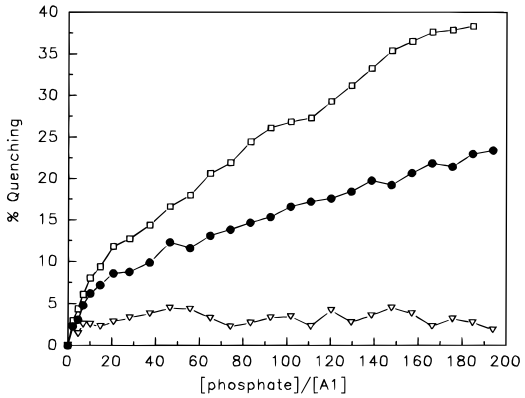

FIGURE 3: Reverse titrations of A1 with 20-mer RNA oligonucleotides derived from the high-affinity winner sequence identified by Burd and Dreyfuss (1994). All titrations were carried out in 10 mM Tris/HCl (pH 7.4) and 25 mM NaCl at 25 °C. A1 (0.2 μM) was titrated with increasing amounts of winner (oligo 7, open inverted triangles), sequence-randomized winner [i.e. random 1 (oligo 8, open squares)], or β-globin intron (oligo 11, closed circles) RNA oligonucleotides.

Initially, reverse titrations were employed to quantify the interaction of A1 with the winner sequence (oligo 7, Table 1) and with two control oligos: a sequence-randomized version of the winner RNA (i.e. oligo 8) and a 20-mer (unrelated) RNA sequence from a β-globin intron (oligo 11). As shown in Figure 3, the results of this analysis were surprising in that the randomized version of the winner RNA oligo appeared to bind A1 significantly more tightly than the actual winner sequence. In fact, even at relatively high concentrations, the winner oligo failed to give significant quenching. Given the tight binding of winner RNA to A1 that was previously observed via a nonequilibrium filter binding assay (Burd & Dreyfuss, 1994), we speculated that binding of winner RNA to A1 was fundamentally different from that of other oligos so far examined in that it did not result in quenching of intrinsic A1 fluorescence (which presumably results primarily from the single A1 tryptophan residue at position 36 in RBD-1). That the winner sequence does indeed bind tightly to A1 is demonstrated by the poly[r(ϵA)] competition assays shown in Figure 4A that were carried out under the same conditions as were the reverse phase titrations (i.e. 25 mM NaCl). Hence, a 50-fold excess of winner RNA (oligo 7) nearly eliminated binding of A1 to poly[r(ϵA)] as shown by the lack of fluorescence enhancement upon addition of A1. In contrast, the presence of a 50-fold molar excess of β-globin intron (oligo 11) results in a fluorescence binding isotherm that is only slightly lower than that for poly[r(ϵA)] in the absence of competitor. The binding isotherm obtained in the presence of a 50-fold excess of random 1 (oligo 8) appears intermediate between the curves corresponding to the winner and β-globin intron competitors. The estimated affinities for A1 binding to the winner, random, and β-globin intron (i.e. oligos 7, 8, and 11 in 0.1 M NaCl) were 4.0×10^8, 5.0×10^6, and 4.2×10^6 M^{-1}, respectively. These values are in good agreement with the affinities of 10^9 and 3.3×10^6 M^{-1} that were previously determined via a nonequilibrium filter binding assay for the winner and β-globin intron oligos, respectively (Burd & Dreyfuss, 1994). Thus, under equilibrium conditions, A1 hnRNP appears to bind approximately 100-fold more tightly to the winner sequence than to either its randomized control or a β-globin intron sequence. Since

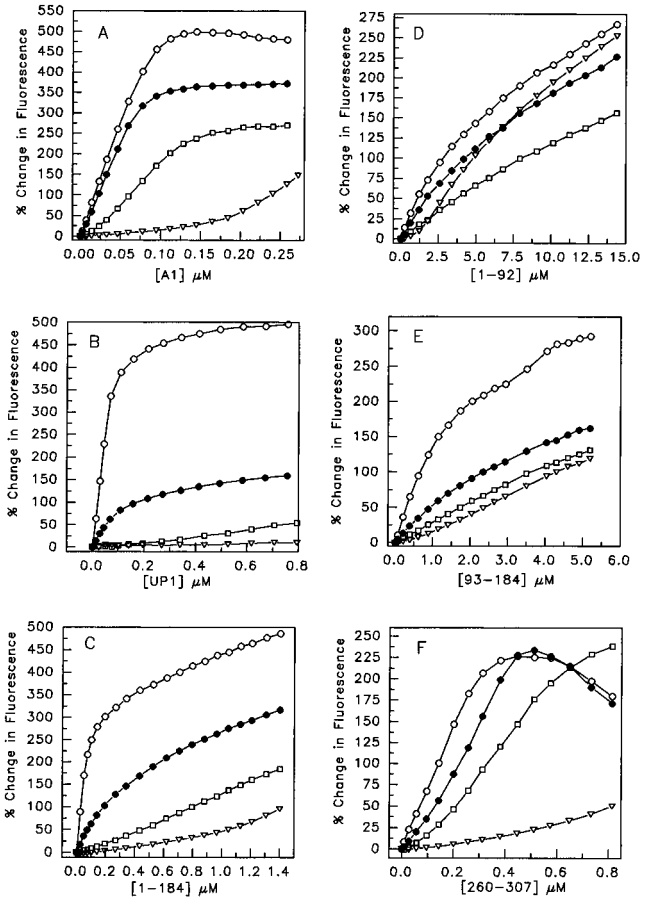

FIGURE 4: Fluorescence titrations of A1 and its fragments with 1 μM poly[r(ϵA)] in the presence of excess 20-mer RNA oligonucleotide competitors. Titrations were carried out in 10 mM Tris/HCl (pH 7.4) at 25 °C with the following oligonucleotide competitors: the high-affinity winner sequence identified by Burd and Dreyfuss (1994) (oligo 7, open inverted triangles), the sequence-randomized version of oligo 7 (oligo 8, open squares), and the 20-mer β-globin intron analogue (oligo 11, closed circles) used by Burd and Dreyfuss (1994). In each instance, a titration of poly[r(ϵA)] that was carried out in the absence of competitor is also shown (open circles). Unless otherwise noted, titrations were carried out in the absence of added NaCl and the competitor concentration was 50 μM (phosphate). The A1 hnRNP fragments used and the residues they span were as follows: (A) intact A1, 1−319, 100 mM NaCl; (B) UP1, 1−195, 25 mM NaCl; (C) 1−184; (D) 1−92; (E) 93−184; (F) 260−307.

the affinities of A1 for all three sequence-randomized controls of the winner oligo (i.e. oligos 8−10) and for the β-globin intron sequence (oligo 11) were within 2-fold of each other, we conclude that under these conditions the approximately 100-fold higher affinity of A1 for the winner sequence results from the unique sequence and/or higher-order structure of the winner oligo as opposed to that of its base composition *per se*.

Both the Sequence and Higher-Order Structure of the Winner Oligo Appear to Contribute to Its Unusually High Affinity for A1. As noted previously, the multiple peaks observed upon anion exchange HPLC of the winner oligo seem to result from a monomer/aggregate equilibrium. That the winner oligo does indeed aggregate is further shown by its behavior on nondenaturing gel electrophoresis and by thermal melting studies. As shown in Figure 5A (lane 1), gel electrophoresis of the winner aggregate gives a major, broad band, which migrates in between the positions expected for a 20-mer and 66-mer DNA oligonucleotide, and a minor band that migrates faster than a 20-mer DNA oligo. On the

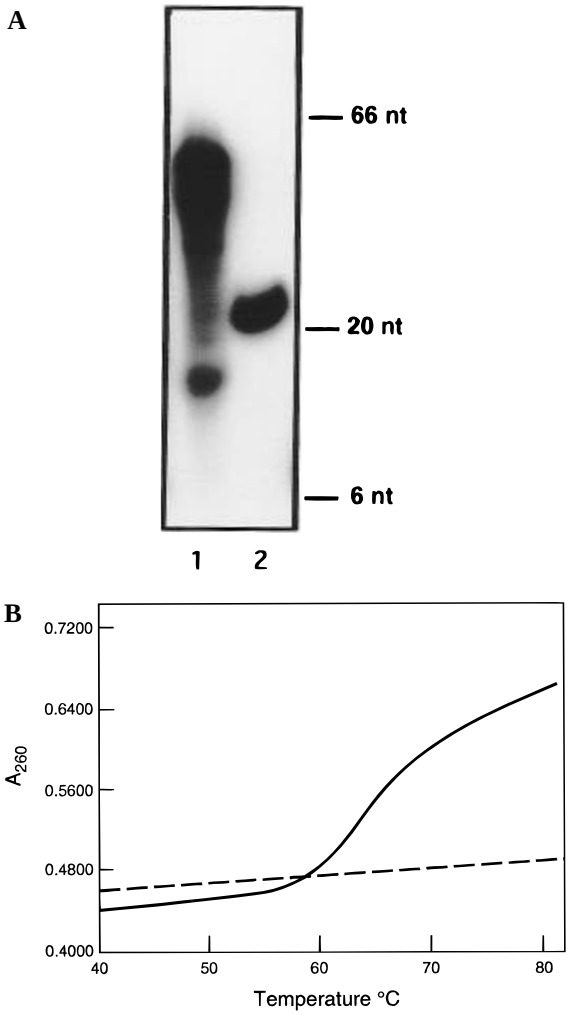

FIGURE 5: Polyacrylamide gel (20%) electrophoresis and thermal denaturation of the 20-mer winner aggregate. (A) Lane 1 contains the ^{32}P-labeled 20-mer RNA winner aggregate following its separation from the winner monomer by anion exchange chromatography. Lane 2 contains the DNA analogue of this sequence, which does not form a stable, higher-molecular weight aggregate. The positions of standards are marked on the right. All samples were prepared in 20 mM Tris/HCl (pH 7.5), 60 mM sodium acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), and 30 mM KCl, and displayed is the autoradiogram obtained after native gel electrophoresis. (B) The absorbance at 260 nm is shown as a function of temperature for winner RNA monomer and aggregate species isolated by anion exchange HPLC (as described in Materials and Methods). In both instances, the RNA concentration was 40 μM (phosphate) and the thermal melting was carried out in 20 mM potassium phosphate buffer (pH 6.0) and 0.1 M potassium chloride. The monomer (dashes) and aggregate (solid line) RNA species appeared as early- and late-eluting peaks, respectively, from the anion exchange HPLC column.

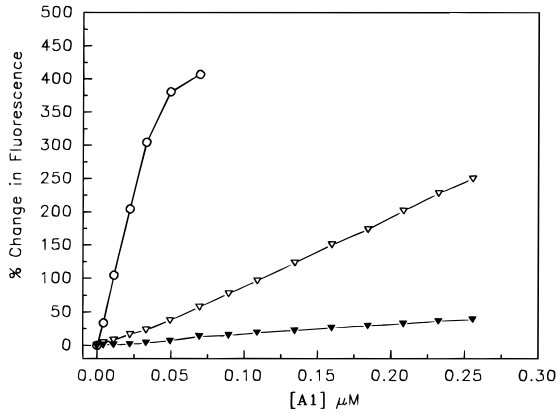

FIGURE 6: Fluorescence titrations of A1 with 1 μM poly[r(ϵA)] in the presence of excess 20-mer winner monomer and aggregate RNA. The top curve (open circles) shows the curve obtained for poly[r(ϵA)] only; the middle curve (open inverted triangles) shows the same titration carried out in the presence of 50 μM winner monomer, and the lower curve (closed inverted triangles) shows the same titration carried out in the presence of 50 μM winner aggregate as isolated by anion exchange HPLC. All titrations were carried out in 10 mM Tris/HCl (pH 7.4) and 150 mM NaCl at 25 °C.

Table 2: Binding of A1 and Its Fragments to Specific and Nonspecific 20-mer RNAs

		K_{app} (M^{-1})[a]		
protein	[NaCl] (mM)	winner RNA (oligo 7)	random RNA (oligo 8)	β-globin RNA (oligo 11)
A1 (1−319)	25	(2.5×10^{10})	(3.0×10^{7})	(1.2×10^{7})
	150	(1.1×10^{8})	(2.9×10^{6})	(3.0×10^{6})
UP1 (1−195)	25	(2.0×10^{8})	(6.0×10^{7})	(2.9×10^{6})
	150	(2.2×10^{6})	(2.3×10^{6})	(1.6×10^{5})
1−184	0	2.9×10^{7}	8.3×10^{6}	7.6×10^{5}
1−92	0	6.1×10^{5}	6.1×10^{5}	1.2×10^{5}
93−184	0	1.6×10^{6}	8.1×10^{5}	2.6×10^{5}
260−307	0	6.6×10^{6}	8.6×10^{5}	4.0×10^{5}

[a] The values in parentheses for A1 and UP1 were estimated by extrapolation of log−log plots.

Table 3: Relative A1 Binding Affinities in 150 mM NaCl

oligo	K_{app} (M^{-1})	fold specificity (compared to that of winner/random)
poly [r(ϵA)]	1.3×10^{8}	
Winner/mix	7.8×10^{7}	26
Winner/monomer	2.6×10^{7}	8.7
Winner/aggregate	1.3×10^{8}	43
Winner/random	3.0×10^{6}	1
β-globin intron sequence (oligo 11)	2.7×10^{6}	−
β-globin 3′-splice site (oligo 1)	1.1×10^{6}	−

basis of these data, the major form of the winner aggregate is most likely a dimer that may exist in more than one isomer. The lower-molecular weight, minor band in lane 1 is most likely a monomer of the winner oligo that contains a hairpin structure and which therefore migrates with a faster mobility than that of its 20-mer DNA analogue (oligo 12) that was used as a standard (see lane 2 in Figure 5A). The thermal denaturation study shown in Figure 5B indicates the winner aggregate has a T_m of about 66 °C in 20 mM potassium phosphate buffer at pH 6.0. Apparently, the stability of winner aggregate is dependent upon the nature of the cation that is present. That is, the T_m decreased to 55 °C when sodium phosphate buffer was substituted for the potassium phosphate buffer used in Figure 5B. Since A1 has a 5-fold higher affinity for the winner aggregate as compared to the winner monomer (Figure 6 and Table 3), it appears that a significant fraction of the specificity of A1 for the winner oligo results from the higher-order structure of the winner aggregate.

The Major Determinants of A1 Specificity for the Winner Oligo Are in the Gly-Rich, COOH Terminal Domain. To gain insight into which A1 domains (see Figure 1) provide most of the apparent A1 specificity for the winner oligo, competition binding studies were carried out on several fragments of A1, including UP1 (residues 1−195), 1−184, both individual RBDs (i.e. 1−92 and 93−184), and a

Table 4: Summary of Binding Data from Salt Sensitivity Studies of A1 and UP1

protein	RNA	oligo	slope from log−log plot	*y*-intercept	predicted maximum number of electrostatic interactions[a]	electrostatic contribution to binding free energy at 150 mM NaCl (%)
A1 (1−319)	winner	7	−3.1	5.5	4.4	32
	random 1	8	−1.3	5.4	1.9	17
	β-globin	11	−0.82	5.8	1.2	10
UP1 (1−195)	winner	7	−2.5	4.3	3.6	32
	random 1	8	−1.8	4.9	2.6	23
	β-globin	11	−1.6	3.9	2.3	25

[a] The value of 0.71 from the slope of a log-log plot represents approximately one ionic interaction as estimated by Record *et al.* (1978).

synthetic peptide analogue (corresponding to residues 260−307) of a portion of the glycine-rich, COOH terminal domain of A1. On the basis of the data in Table 4 (see below), going from A1 to its UP1 fragment reduces the apparent binding specificity [as judged by the ratio of the A1 affinity for winner (oligo 7) to random (oligo 8) RNA in 100 mM NaCl] from about 80-fold to unity. This suggests the primary A1 determinants of specificity for the winner oligo reside within the COOH terminal domain of A1 that spans residues 196−319. This conclusion is supported by the relatively slight specificity (i.e. less than 4-fold) that the 1−184 and its constituent 1−92 and 93−184 RBDs have for the winner oligo (Table 2). Consistent with this idea, a synthetic peptide analogue that corresponds to residues 260−307, which thus includes only about one-third of the glycine-rich, COOH terminal domain of A1, binds approximately 8-fold more tightly to the winner RNA oligo than to its randomized control (Table 2). Hence, the specificity of this 48-residue synthetic peptide analogue for binding the winner oligo is more than twice that of the 195-residue UP1.

Approximately One-Third of the Overall Free Energy of Binding for Winner RNA Appears To Derive from Electrostatic Interactions. To further define protein specificity determinants for binding winner RNA, we used salt sensitivity studies to estimate the ionic contribution to the binding of A1 and UP1 to the winner oligo. Especially in the case of A1, binding of the winner oligo was found to be substantially more salt sensitive than that of either its randomized control or the β-globin intron oligo (Table 4). The 4-fold increase in apparent A1 salt sensitivity observed upon going from the β-globin intron to the winner oligo (Table 4) is again consistent with the presumed importance of the highly basic, COOH terminal domain of A1 in recognizing the winner oligo.

A1 hnRNP Retains Moderate Binding Specificity for the DNA Analogue of the Winner Sequence. As in the case of the RNA version of the winner oligo (see above) , the anion exchange HPLC profile of the winner ssDNA analogue (oligo 12) again revealed an early, and a late-eluting peak, with the latter presumably corresponding to the same aggregate that was seen with the RNA winner oligo. However, thermal denaturation and gel electrophoresis (Figure 5A) studies indicate that the DNA winner aggregate is considerably less stable than the corresponding RNA aggregate. As shown in Figure 7, fluorescence titrations indicate that, although the affinities of A1 for the DNA analogues of the winner, random 1, and β-globin oligos are much lower than those observed for the RNA oligos, A1 nonetheless retains significant specificity of binding for the winner sequence. Hence, the estimated affinities (in 180 mM NaCl) of A1 for the ssDNA analogues of the winner, random 1, and β-globin intron sequences (oligos 12−14) were 3.8×10^6, 5.7×10^5, and 1.5×10^5 M^{-1}, respectively. Hence, A1 binds about 7-fold more tightly to the DNA winner sequence than to its randomized control and about 25-fold more tightly to the DNA winner sequence than to the DNA analogue of the β-globin intron sequence.

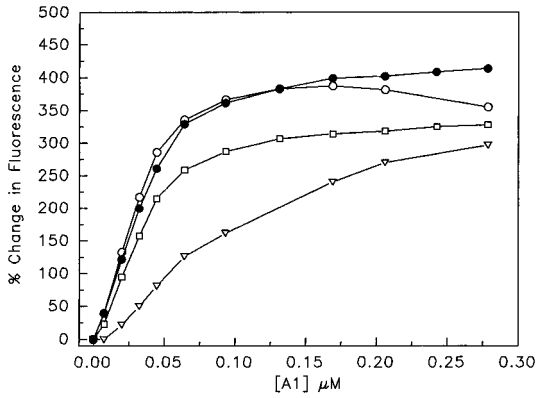

FIGURE 7: Fluorescence titrations of A1 with 1 μM poly[r(ϵA)] in the presence of 20-mer ssDNA analogues of winner, sequence-randomized winner (i.e. random 1), and β-globin intron oligonucleotide competitors. Titrations were carried out in 10 mM Tris (pH 7.4) and 180 mM NaCl at 25 °C. The competitors used were 50 μM winner ssDNA (oligo 12, open inverted triangles), 50 μM sequence-randomized winner ssDNA (oligo 13, open squares), and the 50 μM ssDNA (oligo 14, closed circles) analogue of β-globin intron RNA. The control titration carried out in the absence of competitor is shown by the open circles.

DISCUSSION

One of the more important questions regarding A1 function is its relative ability to recognize and bind with high affinity to specific RNA sequences and/or structures. Although Swanson and Dreyfuss (1988) and Buvoli *et al.* (1990) found that A1 binds preferentially to the polypyrimidine tract at the 3′-end of IVS1 of human β-globin pre-mRNA and other introns, they monitored binding via nonequilibrium techniques (i.e. ribonuclease digestion/immunopurification or photo-cross-linking) which provided only a qualitative estimate of the degree of A1 binding specificity. To more quantitatively address this important question under equilibrium conditions, we have used fluorescence assays based either on oligonucleotide-induced quenching of intrinsic A1 fluorescence or on A1-induced enhancement of the fluorescence of poly[r(ϵA)]. Both assays indicated the affinity of A1 hnRNP for the polypyrimidine tract region at the 3′-end of IVS1 of human β-globin pre-mRNA is unaffected by a mutation in the conserved AG splice site dinucleotide or by randomization of its nucleic acid sequence. We conclude that under equilibrium conditions A1 *cannot* specifically recognize the sequence of this region. We found that A1 actually binds 4-fold *more* tightly to an unrelated β-globin

intron sequence (oligo 11, with an affinity of 4.2×10^6 M^{-1} in 0.1 M NaCl being calculated on the basis of data in Table 4) than it does to the β-globin 3′-splice site region contained within oligo 1 that had been reported to represent a high-affinity A1 target (Swanson & Dreyfuss, 1988; Buvoli *et al*., 1990). Since we were able to qualitatively reproduce the photo-cross-linking study of Buvoli *et al*. (1990), which appeared to demonstrate specific binding of A1 to a DNA analogue (oligo 4) of this β-globin 3′-splice site, we conclude that either the photo-cross-linking itself or the irreversible nature of this assay has led to what may be an inappropriate conclusion. Our data are in agreement with the alternative interpretation by LeStourgeon *et al.* (1990) of the earlier Swanson and Dreyfuss (1988) study, which was based on ribonuclease T1 digestion of synthetic pre-mRNAs (added to splicing competent nuclear extracts) followed by immunopurification. That is, LeStourgeon *et al.* (1990) suggest that, since ribonuclease T1 cuts only at G and because the polypyrimidine sequences present in the added pre-mRNAs are devoid of Gs, the only fragments to survive nuclease digestion would be the polypyrimidine sequences themselves. In other words, the only RNA substrate present in their experiments which was competent to bind A1 (i.e. did not contain any G within the 20-mer or so span of sequence needed to cover the binding site of A1) was the very sequence reported to bind with sequence specificity. LeStourgeon *et al.* (1990) further note the same RNA fragment was immunopurified regardless of whether Swanson and Dreyfuss (1988) used antibodies to hnRNP A1, C, or D. Although the data in Figure 2 indicate A1 cannot recognize the β-globin 3′-splice site studied, before it is concluded that A1 cannot generally recognize such polypyrimidine tracts, similar studies need to be carried out on several other 3′-splice site and control oligos.

Since Buvoli *et al.* (1990) found that the UP1 fragment (i.e. residues 1−195) of A1 retained full binding specificity for DNA analogues of 3′-splice sites, we extended our study to this fragment. Once again, however, our data indicate that like A1 UP1 cannot specifically recognize DNA or RNA analogues of this β-globin 3′-splice site. This finding would seem to qualitatively agree with the observation of Ishikawa *et al*. (1993) that binding of UP1 to a β-globin 3′-splice site oligo is too weak to be detected by gel retardation assays.

We have verified that, under equilibrium conditions, A1 does indeed have unusually high affinity for the winner 20-mer RNA sequence uncovered by selection/amplification by Burd and Dreyfuss (1994). Although the latter studies also relied on a nonequilibrium binding assay (i.e. filter binding), our equilibrium binding data are in close agreement with their results. Hence, we conclude that in 0.1 M NaCl A1 binds 100-fold more tightly to the winner 20-mer than it does to either randomized versions of this oligo or a 20-mer oligo that corresponds to an unrelated stretch of β-globin intron sequence (oligo 11). The ability of A1 to target the winner sequence is unusually salt sensitive such that (on the basis of the data in Table 4) increasing the NaCl concentration to 1 M should abrogate the ability of A1 to differentiate the winner from the random sequences. At this salt concentration, A1 should have an approximately equal affinity of about 3×10^5 M^{-1} for both the winner oligo and its three sequence-randomized controls (i.e. oligos 8−10). It is important to note however that (on the basis of the data in Table 2) nearly 80% of the binding energy for the A1/winner sequence complex derives from interactions that are *within* UP1. Hence, while the ability of A1 to discriminate the winner sequence from its randomized control is resident in the COOH terminal domain, most of the binding energy for interaction with the winner actually derives from the UP1 region of A1.

The finding that A1 recognition of the winner sequence apparently results from ionic interactions (which are abrogated by randomization of the winner sequence) provides insight into possible origins of this specificity. In terms of the protein determinants of specificity, we might expect more basic regions of A1 to be at the interface of the winner/A1 complex. Indeed, our studies on A1 fragments suggest the major determinants of specificity reside within its basic, COOH terminal domain. Hence, conversion of the 319-residue A1 to UP1 (residues 1−195) results in a protein that (in 150 mM NaCl) can no longer differentiate between the winner oligo and its sequence-randomized controls (Table 2). Since the COOH terminal domain (residues 196−319) of A1 has a predicted net charge of +6 at pH 7, as compared to the predicted +1 charge of UP1, the primary involvement of the A1 COOH terminal domain in recognition of the winner sequence is consistent with its overall net positive charge. Although Burd and Dreyfuss (1994) suggested both RBDs in A1 are required for recognition of the winner sequence, this conclusion was based on selection/amplification studies on A1 constructs that all contained the COOH terminal domain. Given our results and the protocol Burd and Dreyfuss (1994) followed, it would have been difficult for them to assess the importance of the COOH terminal domain in recognizing the winner sequence. Since the data in Table 2 indicate UP1 retains ∼80% of the overall binding energy (in 0.15 M NaCl) for the winner oligo, most of the *overall* binding energy derives from interactions that are indeed resident within UP1. However, in view of our data, it is not yet apparent why deletion of one or the other of the A1 RBDs resulted in selected RNA sequences that did not contain, or more closely resemble, the winner sequence (Burd & Dreyfuss, 1994).

Since randomization of the winner sequence would not affect its net charge, it was at first difficult to understand how ionic interactions could account for the ability of A1 to differentiate the winner sequence from its sequence-randomized controls. However, the ability of the winner oligo (but not its sequence-randomized controls) to aggregate provides a possible answer to this conundrum. The relatively monotonic thermal transition of the winner oligo (Figure 5b) suggests the winner aggregate may be rather well-defined. Indeed, on the basis of previous work (Sundquist & Klug, 1989; Williamson, 1994; Scaria *et al*., 1992; Kim *et al.*, 1991), the presence of two GGG sequences within the winner oligo would be expected to result in a G-quartet. Two possible structures [formed by dimerization of two foldback (i.e. hairpin) Hoogsteen base-paired structures] are shown in Figure 8. With regard to possible G-tetrad structures for the winner sequence, it should be noted that the three-dimensional structures of G-quartets vary depending on the sequence, the length of oligonucleotides, the ionic environment, the method of sample preparation, the sample concentration, and the history [for a review, see Williamson (1994)]. We believe the winner sequence exists in solution as an equilibrium mixture of a dimeric G-quartet structure and its constituent monomer, with the latter possibly retaining

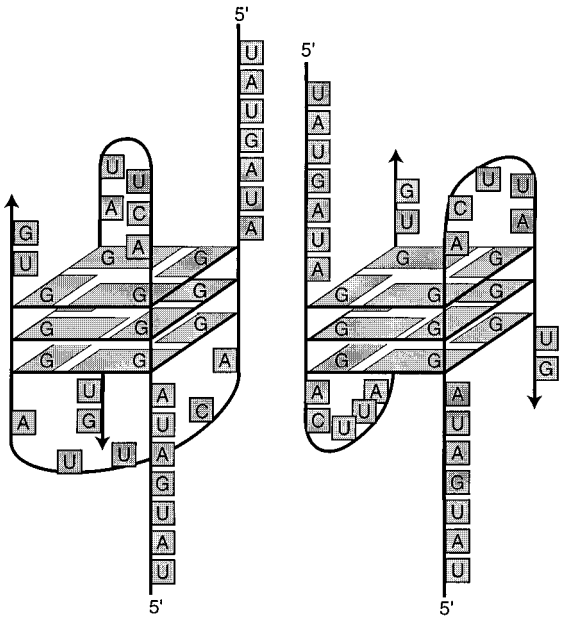

FIGURE 8: Schematic representation of two possible G-quartet structures formed by the winner RNA sequence, UAUGAUAGG-GACUUAGGGUG. The proposed structures are formed by two hairpins that associate so that four guanines are hydrogen-bonded to give a planar array of three G-quartets stacked over each other.

the predicted hairpin. The increased affinity of A1 for the winner sequence aggregate might therefore result either from the increase in negative charge density that accompanies G-quartet formation or simply from the greater number of possible ionic interactions that might occur with the dimeric G-quartet winner sequence as compared to its randomized, monomeric controls. In view of the postulated equilibria in solution between the winner monomer and the G-quartet dimer, it is difficult to assess the absolute degree of "sequence specificity" involved in A1 recognition of the winner oligo. That is, the 9-fold higher affinity of A1 for the monomeric winner oligo as compared to the affinity for its (monomeric) sequence-randomized control (Table 3) may derive from either sequence-specific interactions or G-tetrad formation that may occur during the course of carrying out the binding studies. The higher affinity of A1 for the dimeric G-tetrad, as compared to the affinity for its constituent monomers, leads to the prediction that A1 will shift the monomer $\rightarrow$ G-tetrad equilibrium to the right and promote formation of G-tetrads. Although a variety of unrelated proteins bind preferentially to G-quartet structures *in vitro*, the biological significance of these interactions is not understood (Walsh & Gualberto, 1992; Chung *et al*., 1992; Pearson *et al*., 1993). A common feature of G-quartet binding proteins is that like A1 they all contain one or more highly basic regions. Although data base searches failed to locate the winner RNA sequence in any known pre-mRNAs (Burd & Dreyfuss, 1994), the consensus winner sequence, UAGGGU/A, has some similarity to 5′- (i.e. C/AAGGUA/GAGU) and 3′- (i.e. Py_nNC/UAGGA) vertebrate splice sites. Hence, it was postulated that A1 may interact with both 5′- and 3′-splice site sequences in pre-mRNA and thus help bring them together via A1/A1 protein/protein interactions (Burd & Dreyfuss, 1994). However, the inability of A1 to recognize the polypyrimidine tract region at the 3′-end of IVS1 of human β-globin pre-mRNA and the fact G-tetrads do not occur at splice sites makes this postulate unlikely. Since G-tetrads have been found in the HIV RNA genome (Marquet *et al*., 1993; Sudquist *et al*., 1993), A1 might facilitate dimerization of the HIV RNA genome *in vivo*.

In terms of further assessment of the possible physiological importance of the ability of A1 to recognize G-tetrads, it is interesting to note that, although A1 binds more weakly to DNA than to RNA analogues of the winner sequence, it nonetheless retains specificity for the winner sequence. Hence, A1 has a 7-fold higher affinity for the DNA analogue of the winner sequence (oligo 12) than it does for its sequence-randomized control (oligo 13). Other hnRNP proteins, including A/B, D, and A2/B1 (Ishikawa *et al*., 1993; McKay & Cooke, 1992), also bind oligos containing the TTAGGG sequence found in oligo 12. Interestingly, this sequence matches the vertebrate telomere sequence that is found at chromosome termini (Meyne *et al*., 1989) and that is involved in replication of the ends of eukaryotic chromosomes. The preferential binding of A1 to telomeric-like DNA sequences, which have also been proposed to form G-tetrads (Sundquist *et al*., 1989; Williamson *et al*., 1989), raises the possibility that the role of A1 (and other hnRNP proteins) may extend beyond mRNA biogenesis. In fact, McKay and Cooke (1992) suggest hnRNP A2/B1 may stabilize the double-stranded form of telomeres.

ACKNOWLEDGMENT

We especially thank Dr. Stanley Barnett for carrying out the preliminary fluorescence study that led to this work, Dr. Yousif Shamoo for many useful discussions, and Dr. William Konigsberg for his continued support. We thank the staff of the Oligonucleotide Synthesis Section of the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University for synthesizing the custom oligonucleotides used in this work.

REFERENCES

Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K., & Dreyfuss, G. (1986) *Mol. Cell. Biol. 6,* 2932−2943.
Bandzilius, R. J., Swanson, M. S., & Dreyfuss, G. (1989) *Genes Dev. 3,* 431−437.
Beyer, A. L., Christensen, M. E., Walker, B. W., & LeStourgeon, W. M. (1977) *Cell 11,* 127−138.
Birney, E., Kumar, S., & Krainer, A. R. (1993) *Nucleic Acids. Res. 21,* 5803−5816.
Burd, C. G., & Dreyfuss, G. (1994) *EMBO J. 13,* 1197−1204.
Burd, C. G., Matunis, E. L., & Dreyfuss, G. (1991) *Mol. Cell. Biol. 11,* 3419−3424.
Buvoli, M., Cobianchi, F., Biamonti, G., & Riva, S. (1990) *Nucleic Acids Res. 18,* 6595−6600.
Buvoli, M., Cobianchi, F., & Riva, S. (1992) *Nucleic Acids Res. 20,* 5017−5025.
Caceras, J., Stamm, S., Helfman, D. M., & Krainer, A. R. (1994) *Science 265,* 1706−1709.
Cobianchi, F., Calvo, C., Stoppini, M., Buvoli, M., & Riva, S. (1993) *Nucleic Acids. Res. 21,* 949−955.
Cobianchi, R., Karpel, R. L., Williams, K. R., Notario, V., & Wilson, S. H. (1988) *J. Biol. Chem. 263,* 1063−1071.
Dreyfuss, G., Swanson, M. S., & Pinol-Roma, S. (1988) *Trends Biochem. Sci. 13,* 86−91.
Dreyfuss, G., Matunis, M. J., Pinol Roma, S., & Burd, C. G. (1993) *Annu. Rev. Biochem. 62,* 289−321.
Görlach, M., Wittekind, M., Beckman, R. A., Mueller, L., & Dreyfuss, G. (1992) *EMBO J. 11,* 3289−3295.
Herricks, G., & Alberts, B. (1976) *J. Biol. Chem. 251,* 2133−2141.
Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W., & Keene, J. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88,* 2495−2499.

Ishikawa, F., Matunis, M. J., Dreyfuee, G., & Cech, T. R. (1993) *Mol. Cell. Biol. 13*, 4301−4310.
Kelly, R. C., Jensen, D. E., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7240−7249.
Kenan, D. J., Query, C. C., & Keene, J. D. (1991) *Trends Biochem. Sci. 16*, 214−220.
Kiledjian, M., & Dreyfuss, G. (1992) *EMBO J. 11*, 2655−2664.
Kim, J., Cheong, C., & Moore, P. B. (1991) *Nature 351*, 331−332.
Kowalczykowski, S. C., Paul, L. S., Lonberg, N., Newport, J. W., McSwiggen, J. A., & von Hippel, P. H. (1986) *Biochemistry 25*, 1226−1240.
Kumar, A., & Wilson, S. H. (1990) *Biochemistry 29*, 10717−10722.
Kumar, A., Williams, K. R., & Szer, W. (1986) *J. Biol. Chem. 261*, 11266−11273.
Kumar, A., Casas-Finet, J. R., Luneau, C. J., Karpel, R. L., Merrill, B. M., Wiliams, K. R., & Wilson, S. H. (1990) *J. Biol. Chem. 265*, 17094−17100.
Le Stourgeon, W., Barnett, S. F., & Northington, S. J. (1990) in *The Eukaryotic Nucleus, Molecular Biochemistry and Macromolecular Assemblies* (Strauss, P., & Wilson, S., Eds.) Vol 2., p 477, Telford Press, London.
Marquet, R., Baudin, f., Gabus, C., Darlix, J.-L., Mougel, M., Ehresmann, C., & Ehresmann, B. (1991) *Nucleic Acids. Res. 19*, 2349−2357.
Mayeda, A., & Krainer, A. R. (1992) *Cell 68*, 365−375.
Mayeda, A., Munroe, S. H., Caceras, J. F., & Krainer, A. R. (1994) *EMBO J. 13*, 5483−5495.
McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469−489.
McKay, S. J., & Cooke, H. (1992) *Nucleic Acids Res. 20*, 6461−6464.
Merrill, B. M., & Williams, K. R. (1990) In *The Eukaryotic Nucleus, Molecular Biochemistry and Macromolecular Assemblies* (Strauss, P., & Wilson, S., Eds.) Vol. 2, pp 579−604, Telford Press, London.
Merrill, B. M., LoPresti, M. B., Stone, K. L., & Williams, K. R. (1986) *J. Biol. Chem. 261*, 878−883.
Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., & Williams, K. R. (1988) *J. Biol. Chem. 263*, 3307−3313.
Meyne, J., Ratliff, R. L., & Moyzis, R. K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7049−7053.
Munroe, S. H., & Dong, X. F. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 895−899.
Nadler, S. G., Merrill, B. M., Roberts, W. J., Keating, K. M., Lisbin, M. J., Barnett, S. F., Wilson, S. H., & Williams, K. R. (1991) *Biochemistry 30*, 2968−2976.
Nagai, K., Oubridge, C., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature 348*, 515−520.
Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., & Nagai, K. (1995) *Nature 372*, 432−438.
Pearson, A. M., Rich, A., & Krieger, M. (1993) *J. Biol. Chem. 268*, 3546−3554.
Pinol-Roma, S., & Dreyfuss, G. (1992) *Nature 355*, 730−732.
Pontius, B. W., & Berg, J. M. (1992) *J. Biol. Chem. 267*, 13815−13818.
Query, C. C., Bentley, R. C., & Keene, J. D. (1989) *Cell 57*, 89−101.
Scaria, P. V., Shire, S. J., & Shafer, R. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10336−10340.
Scherly, D., Boelens, W., van Venrooij, W. J., Dathan, N. A., Hamm, J., & Mattaj, I. W. (1989) *EMBO J. 8*, 4163−4170.
Shamoo, A. Y., Abdul-Manan, N., Patten, A. M., Crawford, J. K, Pellegrini, M. C., & Williams, K. R. (1994) *Biochemistry 33*, 8272−8281.
Shamoo, A. Y., Abdul-Manan, N., & Williams, K. R. (1995) *Nucleic Acids. Res. 23*, 725−728.
Sundquist, W. I., & Klug, A. (1989) *Nature 342*, 825−829.
Sundquist, W. I., & Heapy, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3393−3397.
Swanson, M. S., & Dreyfuss, G. (1988) *EMBO J. 11*, 3519−3529.
Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.
Walsh, K., & Gualberto, A. (1992) *J. Biol. Chem. 267*, 13714−13718.
Webster, K. R., Shamoo, A. Y., Konigsberg, W., & Spicer, E. (1991) *Biotech*nology *11*, 658−661.
Weighardt, F., Biamonti, G., & Riva, S. (1995) *J. Cell Sci. 108*, 645−555.
Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 703−730.
Wilson, S. H., Cobianchi, F., & Guy, H. R. (1987) in *DNA, Protein interaction & Gene Regulation* (Thompson, E. B., & Papaconstatinou, J., Eds.) pp 129−146, University of Texas Press, Austin, TX.
Wittekind, M., Görlach, M., Friedrichs, M., Dreyfuss, G., & Mueller, L. (1992) *Biochemistry 31*, 6254−6265.
Yang, X., Bani, M.-R., Lu, S.-J., Rowan, S., Ben-David, Y., & Chabot, B. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 6924−6928.

BI952298P