The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA

MICHAEL J. LAW, ANDREW J. RICE,1 PATTI LIN, and ITE A. LAIRD-OFFRINGA
Departments of Biochemistry and Molecular Biology and of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California 90089-9176, USA

ABSTRACT
The N-terminal RNA Recognition Motif (RRM1) of the spliceosomal protein U1A interacting with its target U1 hairpin II (U1hpII) has been used as a paradigm for RRM-containing proteins interacting with their RNA targets. U1A binds to U1hpII via direct interactions with a 7-nucleotide (nt) consensus binding sequence at the 5’ end of a 10-nt loop, and via hydrogen bonds with the closing C–G base pair at the top of the RNA stem. Using surface plasmon resonance (Biacore), we have examined the role of structural features of U1hpII in binding to U1A RRM1. Mutational analysis of the closing base pair suggests it plays a minor role in binding and mainly prevents “breathing” of the loop. Lengthening the stem and nontarget part of the loop suggests that the increased negative charge of the RNA might slightly aid association. However, this is offset by an increase in dissociation, which may be caused by attraction of the RRM to nontarget parts of the RNA. Studies of a single stranded target and RNAs with untethered loops indicate that structure is not very relevant for association but is important for complex stability. In particular, breaking the link between the stem and the 5’ side of the loop greatly increases complex dissociation, presumably by hindering simultaneous contacts between the RRM and stem and loop nucleotides. While binding of U1A to a single stranded target is much weaker than to U1hpII, it occurs with nanomolar affinity, supporting recent evidence that binding of unstructured RNA by U1A has physiological significance.

Keywords: Biacore; hairpin; kinetics; RNA-binding; RRM; U1A

INTRODUCTION
RNA-binding proteins (RNA-BPs) typically recognize both RNA structure and sequence in order to specifically interact with their correct RNA targets, allowing the resulting protein/RNA complexes to perform their proper cellular functions (Draper 1999; Antson 2000; Hall 2002; Maris et al. 2005). Structural studies have revealed that many RNA-BPs bind single stranded RNA areas restrained by adjacent helical regions, which limit the mobility of the recognition sequence (Antson 2000; Maris et al. 2005). The role of RNA sequence versus structure in RNA–protein interactions remains a very interesting area of investigation.

Here we study U1A, the A protein of the spliceosomal U1 small nuclear ribonucleoprotein (U1snRNP), and its interaction with hairpin II of the U1snRNA (U1hpII; Fig. 1A), to gain more insight into this question. U1A has been widely used as a model system to study the interactions between RNA and proteins containing an RNA recognition motif, or RRM. RRMs are the most commonly found RNA-binding domains in eukaryotes and are present in one or more copies in hundreds of proteins (Varani and Nagai 1998; Rubin et al. 2000). These versatile globular RNA-binding platforms mediate binding to single stranded and/or structured RNAs in a multitude of proteins.

1Present address: University of Vermont College of Medicine, Burlington, VT 05405-0068, USA.

Abbreviations: bp, base pair(s); RRM, RNA recognition motif; U1snRNP, U1 small nuclear ribonucleoprotein; U1A, A protein of the U1 snRNP; U1hpII, U1 hairpin II.

Reprint requests to: Ite A. Laird-Offringa, USC/Norris Cancer Center, 1441 Eastlake Avenue, Room NOR6420, Los Angeles, CA 90089-9176, USA; e-mail: ilaird@usc.edu; fax: (323) 865-0158.

Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.75206.
Role of RNA structure in U1A/U1hpII binding

FIGURE 1. Schematic representation of RNA mutants tested, closing base pair mutations, and interactions of interest in the U1A/U1hpII complex. (A) Representation of the RNA mutants investigated in this study. Nucleotides U-5 to G15 are identical to those in the wild-type U1hpII RNA target. The consensus binding sequence is indicated in the U1hpII diagram with an asterisk (*); permutations of the U1hpII U1hpII RNA target. The consensus binding sequence is indicated in
structure are highlighted with boxes or underlines. The numbering scheme is based on numbering of the loop residues from 1 to 10, with
the U1hpII diagram with an asterisk (*); permutations of the U1hpII
U1hpII RNA target. The consensus binding sequence is indicated in
structure are highlighted with boxes or underlines. The numbering

vitro selection (SELEX) experiments demonstrated that U1A displays high specificity for a 7-nucleotide (nt) recognition sequence, AUUGCAC (Tsai et al. 1991). Optimal binding to this sequence occurs when it is presented in the context of a stem–loop type structure, as in U1hpII or the polyadenylation inhibition element (PIE), an autoregulatory element in the U1A mRNA that binds two molecules of U1A (Scherly et al. 1990; Tsai et al. 1991; Boelens et al. 1993; van Gelder et al. 1993; Hall 1994). Both U1hpII and the PIE display the AUUGCAC sequence as a single stranded region, and contain a closing C–G base pair at the top of an adjacent RNA stem, but they do so in different structural contexts. In U1hpII the target bases are part of a loop; in the PIE they are present in a bulge between stems. Surprisingly, despite the demonstrated importance of a stem for the interaction (Scherly et al. 1989; Hall 1994; Luchansky et al. 2000), U1A also appears to bind RNA in vivo in the absence of a stem; it interacts with consensus-type motifs that exhibit no clear secondary structure in the immunoglobulin M secretory mRNA, thereby regulating B-cell–specific polyadenylation (Phillips et al. 2001; Ma et al. 2006). The ability of U1A to bind RNA in different structural contexts raises the question: How relevant is RNA structure to U1A binding, and by what mechanism does structure contribute to the high affinity of the interaction? In order to answer this question, we undertook a kinetic study of the interaction between the N-terminal U1A RRM (herein referred to as U1A) and a series of different target RNAs. The design of the targets was based upon previous structural and biochemical studies of the U1A/U1hpII interaction, as outlined below.

The U1A/U1hpII cocrystal structure reveals multiple hydrogen bonds between U1A and the RNA. Most of these are to the 7-nt recognition sequence in the U1hpII loop (Fig. 1A). However, there are also hydrogen bonds to the closing C–G base pair of the RNA stem (Fig. 1C; Oubridge et al. 1994): N7 and O6 of G11 (Fig. 1B) interact with both amide groups (Nζ1 and 2) of Arg52 while the phosphate group of G11 interacts with Leu49. Arg52 also forms a hydrogen bond with A1 in the RNA loop (via Nζ1), a step thought to occur early in complex formation (Oubridge et al. 1994; Tang and Nilsson 1999; Law et al. 2005). These interactions may promote a series of further contacts that lock U1A and U1hpII together. Thus, the stem and the closing base pair may play a role in initiating close-range interactions (Law et al. 2005). In addition, the stem also interacts electrostatically with U1A. The phosphate backbone of the stem provides a structured negatively charged region that could attract positively charged residues in U1A, thereby luring the protein to U1hpII (Katsamba et al. 2001). Kinetic studies from the Laird-Offringa lab have shown that electrostatic interactions between the stem and Lys20 and Lys22 (Fig. 1C) are important for both association and complex stability (Katsamba et al. 2001; Law et al. 2006). Lastly, the stem
functions to constrain the mobility of the RNA loop nucleotides. The first 7 nt of the 10-nt U1hpII loop constitute the canonical AUUGGAC binding site (Fig. 1A). The last three loop nucleotides are nonconserved (Tsai et al. 1991), but we have shown that one or more of them may interact electrostatically with Lys23 (Fig. 1C; Law et al. 2006). The functional replacement of the nonconserved loop residues with a synthetic polyethylene glycol linker suggests that the most important role of these nucleotides is to tether the loop to the stem, limiting loop flexibility (Williams and Hall 1996). This idea is supported by the apparently mobile nature of loop nucleotides 8-10 in structural analyses of U1A/U1hpII complexes (Oubridge et al. 1994; Tang and Nilsson 1999; Reyes and Kollman 2000; Blakaj et al. 2001; Pitici et al. 2002). Reducing the size of the RNA loop by removing one or more of the “spacer nucleotides” is known to be deleterious to RNA binding (Williams and Hall 1996; Katsamba et al. 2001), presumably because the shortened loop cannot accommodate the protein properly. Increasing the size of the RNA loop using non-nucleotide spacers has little impact on binding affinity (Williams and Hall 1996), suggesting that some loss of constraint is well tolerated. In summary, the function of the RNA stem may be threefold: to provide an interaction of constraint is well tolerated. In summary, the function of the RNA stem may be threefold: to provide an interaction and maintain complex stability, and to constrain the mobility of the loop to facilitate the interaction.

Here we used a surface plasmon resonance biosensor (Biacore) to further explore the various roles of the stem. Biacore analysis provides high quality kinetic measurements that allow one to determine whether certain features are important for complex formation (through measurements of the dissociation rate or $k_d$) versus complex stability (through measurements of the association rate or $k_a$) (Katsamba et al. 2002b), thereby providing unique mechanistic insights that can be used to determine the underlying cause for differences in affinity ($K_{D}$, which equals $k_d/k_a$). We first studied the function of the closing base pair by mutating it to all three possible alternative identities. We next examined the electrostatic role of the stem by extending it with 5 or 10 bp. Lastly, we assessed its tethering function by either lengthening the loop or releasing the loop from the stem at either end. We also dispensed with the base-pairing of the stem altogether by mutating it, maintaining the same length of RNA target but in a single stranded form.

RESULTS AND DISCUSSION

Mutation of the closing base pair has little impact on binding kinetic

In order to assess how the closing base pair at the top of the stem contributes to the high affinity interaction with U1A, we analyzed the kinetics of U1A RRM1 binding to hairpins with three permutations of these 2 nt: C–G was mutated to G–C, A–U, and U–A (Fig. 1A, B). Alteration of the closing base pair to any one of the alternate base pairs had small effects on the association rate of the complex; only the U–A mutant showed a statistically significant difference from wild-type RNA ($k_a$ reduced to 0.8-fold the wild-type value) Figs. (2, 3; Table 1). Complex stability was moderately impacted in all three cases. Mutation to G–C caused a small, statistically significant loss of complex stability (1.9-fold increase in $k_d$), while mutation to A–U or U–A showed a fivefold and 4.2-fold loss in stability, respectively.

These losses in stability could be due to different factors. One possibility is that mutating the closing base pair would result in the loss of hydrogen bonds that normally exist between U1A and U1hpII. In the mutations to G–C and A–U, the N7 interaction between the closing base pair nucleotide and Arg52 N$\text{z}$1 would be lost. In contrast, mutations to G–C and U–A would result in a loss of the O6 interaction with Arg52 N$\text{z}$2 (Fig. 1B,C). This could affect the positioning of Arg52 and alter its ability to make subsequent intraprotein contacts, affecting the ability of U1A to interact with U1hpII. If the loss of hydrogen bonds between U1A and U1hpII resulted in reduced complex stability, one would expect the G–C mutant RNA to form the least stable complex. However, this idea does not mesh with the kinetic data. The complex of U1A with the A–U mutant, which retains O6, is the least stable, while that of the G–C RNA, which loses both N7 and O6, is least affected.

An alternative explanation is that the loss in complex stability results from a reduction in the number of hydrogen bonds between the closing RNA bases. In the G–C and C–G closing base pair RNAs, three hydrogen bonds hold the top of the stem together, while only two bridge the A–U and U–A closing base pairs. The single hydrogen bond loss causes a destabilization of the RNA by 2 kcal/mol (as determined using mfold [http://bioweb.pasteur.fr/sequal/interface/mfold.html]), resulting in a potential increase in the flexibility of the RNA loop and “breathing” of the closing base pair. Opening up the closing base pair would lengthen the loop and alter the position of the target site with respect to the RNA stem. Insertion of bases 5’ to the canonical U1A recognition sequence has been shown to perturb RNA binding (Williams and Hall 1996), supporting the notion that spacing of the stem and the loop target sequence is important. The fact that the A–U and U–A mutant RNA targets show similar losses in complex stability (5.0- and 4.2-fold, respectively), while the G–C mutant shows a (statistically significant) 1.9-fold loss in RNA stability, lends further credence to this idea. Thus, we presume that the 1.9-fold loss in stability seen in the G–C mutant is due to the loss in interactions between U1A and U1hpII, while the additional instability of the A–U and U–A complexes arises from destabilization of the RNA stem.
FIGURE 2. Sensorgrams showing kinetic analyses of wild-type U1A with U1hpII and mutant RNAs. Protein concentrations injected are as indicated. Black lines represent triplicate injections which were performed in random order over an indicated RNA surface. A 1-min association was followed by a 5-min dissociation phase. Red lines represent the global fit of data sets using CLAMP (Myszka and Morton 1998). Kinetic parameters obtained from the experiments are given in Table 1.
The modest but statistically significant difference in stability between the complexes with the A–U versus U–A closing base pair suggests that of the two Arg52-interacting groups (O6 and N7), the latter is most important (Fig. 1B,C). In conclusion, it appears as if the closing base pair contributes marginally to association and, at most, fivefold to the stability of the U1A/U1hpII complex. Our data suggest that its role in positioning the loop is more important than the contacts it makes to Arg52. We and others (Tang and Nilsson 1999; Law et al. 2005) had previously proposed a key role for Arg52 in initiating close range interactions between protein and RNA. The modest effect of closing base-pair mutations suggests that the electrostatic role of Arg52, its interaction with A1, and its positioning of the neighboring amino acids Arg47, Ser48, and Gln54 (Law et al. 2005) are more important than its interaction with G11 (Fig. 1C). It would be of interest to examine the role of Arg52 more closely in future studies.

TABLE 1. Kinetic parameters for the interaction of U1A with U1hpII and RNA mutants

<table>
<thead>
<tr>
<th>RNA target</th>
<th>( k_d ) (M(^{-1}) s(^{-1}))</th>
<th>Fold (mut/wt)</th>
<th>( k_a ) (s(^{-1}))</th>
<th>Fold (wt/mut)</th>
<th>( K_D ) (M)</th>
<th>Fold (mut/wt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1hpII</td>
<td>((1.3 \pm 0.07) \times 10^7)</td>
<td>1.1</td>
<td>((4.3 \pm 0.05) \times 10^{-4})</td>
<td>1.9</td>
<td>((3.4 \pm 0.2) \times 10^{-11})</td>
<td>1.8</td>
</tr>
<tr>
<td>CG to GC</td>
<td>((1.4 \pm 0.05) \times 10^7)</td>
<td>1.1</td>
<td>((8.4 \pm 0.3) \times 10^{-4})</td>
<td>5.0</td>
<td>((6.0 \pm 0.2) \times 10^{-11})</td>
<td>4.5</td>
</tr>
<tr>
<td>CG to AU</td>
<td>((1.5 \pm 0.2) \times 10^7)</td>
<td>0.8</td>
<td>((2.2 \pm 0.1) \times 10^{-3})</td>
<td>4.2</td>
<td>((1.8 \pm 0.1) \times 10^{-10})</td>
<td>5.3</td>
</tr>
<tr>
<td>CG to UA</td>
<td>((1.0 \pm 0.03) \times 10^7)</td>
<td>1.6</td>
<td>((1.2 \pm 0.03) \times 10^{-3})</td>
<td>2.7</td>
<td>((5.8 \pm 0.6) \times 10^{-11})</td>
<td>1.7</td>
</tr>
<tr>
<td>Stem + 5 bp</td>
<td>((2.0 \pm 0.3) \times 10^7)</td>
<td>2.0</td>
<td>((1.3 \pm 0.05) \times 10^{-3})</td>
<td>2.9</td>
<td>((5.6 \pm 1.3) \times 10^{-11})</td>
<td>1.7</td>
</tr>
<tr>
<td>U1hpII U8C+5 bp</td>
<td>((1.4 \pm 0.2) \times 10^7)</td>
<td>1.1</td>
<td>((5.9 \pm 0.3) \times 10^{-4})</td>
<td>1.4</td>
<td>((4.3 \pm 0.5) \times 10^{-11})</td>
<td>1.3</td>
</tr>
<tr>
<td>U8C + 2C</td>
<td>((2.8 \pm 0.3) \times 10^7)</td>
<td>2.0</td>
<td>((9.2 \pm 0.7) \times 10^{-4})</td>
<td>2.1</td>
<td>((3.5 \pm 0.4) \times 10^{-11})</td>
<td>1.0</td>
</tr>
<tr>
<td>U8C + 4C</td>
<td>((2.6 \pm 0.2) \times 10^7)</td>
<td>1.1</td>
<td>((2.8 \pm 0.5) \times 10^{-3})</td>
<td>6.5</td>
<td>((1.1 \pm 0.1) \times 10^{-10})</td>
<td>3.2</td>
</tr>
<tr>
<td>3’ loose loop</td>
<td>((9.8 \pm 1.0) \times 10^6)</td>
<td>0.8</td>
<td>((1.7 \pm 1.0) \times 10^{-2})</td>
<td>40</td>
<td>((1.8 \pm 0.2) \times 10^{-9})</td>
<td>54</td>
</tr>
<tr>
<td>5’-loop strand</td>
<td>((1.2 \pm 0.3) \times 10^7)</td>
<td>1.3</td>
<td>((1.3 \pm 0.2) \times 10^0)</td>
<td>2900</td>
<td>((7.3 \pm 0.4) \times 10^{-8})</td>
<td>2200</td>
</tr>
<tr>
<td>Single stranded</td>
<td>((4.3 \pm 1.5) \times 10^6)</td>
<td>0.3</td>
<td>((2.5 \pm 0.7) \times 10^{-1})</td>
<td>590</td>
<td>((6.3 \pm 0.4) \times 10^{-8})</td>
<td>1900</td>
</tr>
</tbody>
</table>

Average and standard error of the mean are given. Values in bold italics represent a statistically significant difference from wild-type U1hpII. To calculate the standard deviation, the error bars were then used to compute the average and standard error of the mean. Ratios for the fold change were calculated to most frequently result in a number >1 (mut/wt for \( k_d \) and \( K_D \), wt/mut for \( k_a \)). Fold changes are approximations represented to two significant digits.
introduction of a nucleotide at the 5' side of the loop. These small gains in association were offset by a loss in stability of the protein/RNA complex; both the +5 bp and +10 bp RNAs dissociated more quickly than the wild-type RNA construct, showing statistically significant 2.7- and 2.9-fold losses in complex stability, respectively (Fig. 3; Table 1).

The observed changes in association and stability might be due to stabilization of the stem by the added base pairs, but we believe this is unlikely for two reasons: First, the added base pairs are at the bottom of the stem, away from the loop and are unlikely to diminish "breathing" of the loop; secondly, if stem stability were important for binding, we would expect complexes with RNAs carrying longer stems to be more rather than less stable, because the stem-loop structure would be better maintained. Thus, it seems more likely that the observed kinetic effects are due to increased electrostatic attraction between U1A and the RNA. We have previously shown, using kinetic analysis and salt-dependence experiments, that electrostatic attraction plays an important role in the association of U1A with U1hpII (Katsamba et al. 2001; Law et al. 2006). Increasing the length of the RNA stem would result in an increase in net negative charge in this region, associated with the added phosphate groups. This might result in a more efficient recruitment of the protein. However, in the bound complex, it might also stimulate dissociation by drawing stem-interacting lysines, such as Lys20 and Lys22, down the stem, destabilizing the complex. In molecular dynamics simulations of the U1A/U1hpII complex, we have observed a "sliding" of Lys20 and 22 down the stem, away from the RNA loop (Fig. 1C; Law et al. 2006). The kinetic differences between our standard U1hpII target and the ones with lengthened stems are of interest, as the natural U1 snRNA is much longer and would carry many additional and more widely distributed negative charges. In the future, it would be important to examine these interactions in the context of the full-length U1A protein and the full-length U1snRNA, preferably in the presence of the remaining U1snRNP proteins. In conclusion, our results indicate that increasing the size of an RNA target may facilitate binding of RNA-BPs via favorable electrostatic interactions. However, this may lead to electrostatic attraction of the RNA-BP to an inappropriate region of the RNA, which might ultimately result in a destabilization of the protein/RNA complex.

Increasing loop length results in faster association while decreasing complex stability

We next wanted to examine the role of loop flexibility in the interaction by elongating the loop. Because the introduction of a nucleotide at the 5' side of the loop is deleterious to binding, while insertion of one at the 3' side (following the nonconserved nucleotides) is not (Williams and Hall 1996), we chose to elongate the loop at its 3' side. Prior to testing the kinetic impact of increasing loop size, we first tested the impact of mutating U8 to C (Fig. 1A). This mutation was necessary to minimize any alternate RNA base-pairing within the loop upon loop extension. U1A bound similarly to U1hpII and the U8C RNA, showing only a small, but statistically significant loss in complex stability (1.4-fold; Figs. 2, 3; Table 1). This suggests that mutation of U8 has minor kinetic effects, confirming the notion that the last 3 nt of the RNA loop need not be conserved. We proceeded to insert two or four cytosine residues into the 3' side of the RNA loop (U8C + 2C and U8C + 4C), increasing both the flexibility and negative charge associated with the loop (Fig. 1A). RNA folding analysis using mfold indicated that these RNA targets would provide a larger loop while maintaining a base-paired stem identical to our original U1hpII RNA (Fig. 1A).

Analysis of the kinetic impact of increasing the loop size showed a statistically significant approximately twofold increase in association rate for both RNAs (Figs. 2, 3; Table 1). In the case of U8C + 2C, this increase was negated by a 2.1-fold loss in complex stability, resulting in no change in the affinity of U1A for this RNA target. With the U8C + 4C mutant, the increase in association rate was coupled with a statistically significant 6.5-fold loss in complex stability, resulting in a 3.2-fold net loss in affinity. These results agree well with those made by Williams and Hall (1996), who showed a very modest approximately twofold loss in affinity when the three spacer nucleotides were replaced by polyethylene glycol linkers two to three times longer than the original spacer sequence. Our data show that similar affinities can sometimes hide kinetic changes; the mechanistically different faster-on faster-off interaction of U1A with the longer loop targets can only be observed when interaction rates are measured. The increase in association could be due to the increased flexibility of loop nucleotides, which might facilitate complex formation. However, structural analyses of U1hpII indicate that there is already substantial mobility of the RNA bases with a 10-nt loop (Tang and Nilsson 1999; Reyes and Kollman 2000; Blakaj et al. 2001; Pitici et al. 2002), and too much flexibility might in fact be deleterious. Alternatively, the increased association might be related to the augmented negative charge of the loop, which could facilitate the association phase through enhanced attraction of U1A to the recognition sequence. Molecular dynamics simulations of the U1A/U1hpII complex have suggested an electrostatic interaction between the spacer nucleotides and Lys23 (Fig. 1C; Law et al. 2006). This interaction might be facilitated by lengthening the spacer.

Analogous to the stem extension, increasing the loop size appears to lead to destabilization of the complexes. This loss in stability may be due to the added negative charge, which could electrostatically draw U1A away from the target nucleotides, in a similar way a lengthened stem might draw on Lys20 and Lys22. Alternatively, the destabilization may result from the increased flexibility of the RNA loop.

Role of RNA structure in U1A/U1hpII binding
and the energetic consequences related to increased entropy of the unbound longer-loop RNAs. In order to further study the consequences of increased loop flexibility, we analyzed binding of U1A to RNA targets in which the loop was released from the stem (see below).

Releasing the RNA loop from the 5’ end has a greater kinetic impact than releasing it from the 3’ end

The importance of tethering the loop to the stem was tested by kinetic analysis of U1A interacting with targets in which the loop was released from the stem at its 3’ or 5’ side (3’ loose loop and 5’ loose loop; Fig. 1A). Using an annealed RNA target, it had previously been shown that uncoupling the RNA loop at its 3’ side resulted in a loss in affinity for U1A of 4 orders of magnitude (Luchansky et al. 2000). Whether this was caused by an association or a dissociation defect was not examined, nor was a 5’ untethered loop studied. Our kinetic measurements show that release of the loop on either end results in a small, nonsignificant loss in the association rate (Figs. 2, 3; Table 1), suggesting that all the components for association are present. The inability of U1A to bind to the tetraloop RNA, combined with our observation that targets with longer stems or loops show an increased association rate, suggests that the total charge and the availability of the target sequence are key factors for proper association. The loose loop RNAs present a slightly elevated negative charge (due to the presence of linker nucleotides at the bottom and the top of the stem) and carry the full canonical target. Hence, proper association is perhaps not all that unexpected. The slightly enhanced negative charge might compensate for potentially deleterious effects of increased loop flexibility, but the insignificant overall effect on association indicates that these are minor factors. In contrast, the effects on complex stability are very pronounced; the 3’ loose loop construct formed a 40-fold less stable complex than the wild-type U1hpII, while the 5’ loose loop construct generated a 2900-fold less stable complex than the wild-type U1hpII (Luchansky et al. 2000). These observations point to the need for a stem, but the requirement were not entirely clear. We designed our single stranded U1hpII to maintain the sequence of the loop and closing base pair, but removed the secondary structure provided by the stem (single stranded target; Fig. 1A). To ensure specificity of the interaction, we also tested the ability of U1A to bind to a single stranded AU-rich RNA of similar length and found no binding at the concentrations examined (Fig. 2). When measuring the kinetics of the interaction of single stranded U1hpII RNA with U1A, a statistically significant threefold loss in association rate was observed (Figs. 2, 3; Table 1). This represents the largest statistically significant loss in association rate of the RNA targets tested in this study (excluding the nonbinding tetraloop and single stranded AU-rich RNA, for which no association rate could be measured). The threefold reduction in $k_a$ is smaller than losses in association rate observed when two spacer nucleotides are deleted from the loop (~15-fold) (Katsamba et al. 2001) or when the protein is altered in key residues, for example by substituting basic by acidic residues (up to 45-fold loss in association) (Law et al. 2006) or mutating aromatic residues such as Tyr13 (14–20-fold) (Law et al. 2005). The statistically significant but modest loss in association indicates that most components necessary for complex formation are present: the negative charge of the nucleotides, and the consensus target sequence. Presumably, this allows the formation of sufficient close-range interactions to allow the two molecules to begin association. Based on our previous work, this involves, among others, electrostatic interactions between the RNA and Lys20, 22, and 50, and the stacking of Tyr13 on C5 (Katsamba et al. 2001, 2002a; Law et al. 2005, 2006). Thus, surprisingly, the unstructured nature of the target sequence is not a major
Implications of single stranded RNA binding by U1A

Although we observe a loss in affinity of 3 orders of magnitude when the RNA target is presented in a linear context, the equilibrium dissociation constant of U1A for the unstructured RNA is still in the nanomolar range (~63 nM). This would be considered a high-affinity interaction, certainly when it is compared to the RNA-binding ability of single RRs from multi-RRM-containing proteins such as HuD, hnRNPA1, sex-lethal, and poly-A binding protein; individually, these RRs bind to RNA with a dramatically lower affinity (μM range) and reduced specificity than the multiple linked RRM domains (Burd et al. 1991; Shamoo et al. 1994; Kanaar et al. 1995; Samuels et al. 1998; Deo et al. 1999; Park et al. 2000; Park-Lee et al. 2003). Structural analyses show that these multi-RRM proteins utilize more than one RRM to generate an RNA-binding platform large enough to provide interactions with the RNA, which is draped over the binding surface in a semilinear fashion (Burd et al. 1991; Shamoo et al. 1994; Kanaar et al. 1995; Samuels et al. 1998; Deo et al. 1999; Wang and Tanaka Hall 2001; Hall 2002). In contrast, the N-terminal U1A RRM has evolved to interact with the highly structured U1hpII RNA, in which the spatially constrained loop nucleotides splay out to make many specific contacts on a single RRM (Oubridge et al. 1994). This explains the high affinity of U1A RRM1 for U1hpII and also provides an explanation for the relatively high affinity of U1A for a single stranded RNA; the bound single stranded target RNA probably follows a similar curved path over the RRM surface, allowing it to make specific contacts to a large surface area. To our knowledge, only one other case of nM affinity binding of an unstructured RNA by an isolated RRM has been described: that of the splice-regulatory UGCAUGU RNA element interacting with the alternative splicing regulator Fox-1 (Auweter et al. 2006). The recently solved NMR structure of the Fox-1/RNA complex shows the three 3′ target bases (UGU) interacting with the β-sheet surface of the RRM, as expected. However, the 5′ UGCA nucleotides behave abnormally: They surround a phenylalanine residue in one of the loops flanking the RRM. The 4 nt form noncanonical intra-RNA interactions that allow the bases to curve around the phenylalanine side chain, generating a hydrophobic pocket and positioning three of the bases to make specific contacts to the protein. Like the interaction of U1A RRM1 with an unstructured target, the Fox-1 interaction derives its high affinity from a high density of contacts achieved by a non-linear RNA arrangement. As studies of RNA–protein interactions expand to include structural analyses of ever more complexes, it is likely that further examples of such high density RNA contacts to single RRs will be uncovered.

Recent work has implicated free (non-U1snRNP) U1A in cleavage and polyadenylation of mRNAs (O’Connor et al. 1997; Liang and Lutz 2006). Non-U1snRNP U1A functions in a multiprotein complex to regulate polyadenylation, further suggesting a link between splicing and polyadenylation. U1A has also been shown to be important for the switch from membrane-bound to secreted Immunoglobulin M (IgM) that occurs during B-cell differentiation (Phillips et al. 2001, 2004; Phillips and Gunderson 2003; Ma et al. 2006). In immature B-cells, free U1A inhibits the polyadenylation of secretory IgM by binding to elements that resemble degenerate consensus U1A binding sequences.
(A(U/G)GC(N1–3)C) in the IgM pre-mRNA (Phillips et al. 2001, 2004; Phillips and Gunderson 2003; Ma et al. 2006). In mature B-cells, U1A levels are decreased, allowing polyadenylation of the transcript and production of secreted IgM. Given their seemingly unstructured nature and their degenerate target sequence, the affinity of U1A for these sites would be expected to be very low. However, we have shown that U1A can bind quite well to an unstructured target. The presence of multiple adjacent sites may substantially strengthen the interaction through cooperative binding, increasing the overall affinity of U1A for the IgM pre-mRNA. This idea is based on the observed dimerization of U1A on the PIE, which contains one fully conserved and one degenerate U1A target sequence, presented in adjacent bulges of the PIE in opposing orientations (van Gelder et al. 1993; Klein Gunnewiek et al. 2000; Varani et al. 2000). Cooperativity is mediated by the interaction of the two bound full-length U1A molecules via the hinge region connecting RRMs 1 and 2. Given the demonstrated ability of U1A RRM1 to bind with nM affinity to linear RNA, we could imagine similar cooperative binding to the secretory IgM 3′ untranslated region. Indeed, gel shift analysis shows binding of multiple full-length U1A molecules to the IgM pre-mRNA (Phillips et al. 2004). The demonstration that U1A can bind tightly to different RNA structures pertaining to splicing, autoregulation, and secretory IgM regulation suggest additional U1A functions in the cell remain to be discovered. Our analysis of the interaction of U1A RRM1 with different U1hpII-derived target RNAs not only sheds light on the mechanism of this particular interaction but also shows the versatility of this RNA-binding protein. Kinetic analyses of the full-length U1A interacting with a variety of target RNAs may provide further clues to its biological functions in cells.

MATERIALS AND METHODS

Construction of U1A and protein purification

Throughout these studies, an N-terminal fragment of the human U1A protein (herein referred to as “U1A”) containing the first RRM (amino acids 1–101) was used (Katsamba et al. 2001). This fragment has been demonstrated to be necessary and sufficient for specific and high-affinity binding to U1hpII (Scherly et al. 1989; Lutz-Freyermuth et al. 1990). The recombinant protein contains a C-terminal His, tag that was used for purification with Nickel agarose beads and a c-myc tag. Bound protein was washed three times to remove nonspecifically bound proteins and eluted using buffer containing increasing concentrations of imidazole (50–200 mM imidazole). The concentration of each protein was measured by Coomassie blue staining of an extensive protein dilution series next to a standard on SDS-PAGE gels.

RNA preparation

RNA targets were designed in order to probe the role of RNA structure in mediating the kinetic interaction of U1A. Following verification that the RNAs would adopt the desired structure (using the mfold program [http://bioweb.pasteur.fr/seqanal/interfaces/mfold.html]), RNAs were chemically synthesized to carry biotin tags at locations designated in Figure 1 (Dharmacon Research), allowing RNAs to be coated on streptavidin-coated sensor chips (SA chips, Biacore Inc.). RNA coating densities were varied according to the affinity of U1A for each respective RNA target (higher densities for weaker interactions).

Biosensor analysis

Binding experiments were performed on a Biacore 2000 instrument (Biacore Inc.). RNA was diluted to a final concentration of 1 μM in HBS buffer (10 mM HEPES at pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% surfactant P20 [Biacore Inc.]) followed by heating at 80°C for 10 min and cooling to room temperature to allow annealing of the stem. The sample was then diluted 500-fold in running buffer (10 mM Tris/HCl at pH 8.0, 150 mM NaCl, 5% glycerol, 125 μM rRNA [Roche], 62.5 μg/mL acetylated bovine serum albumin [New England Biolabs], 1 mM dithiothreitol, 0.05% surfactant P20) and injected over the sensor chip surface at 10 μL min⁻¹ at 20°C. To provide an optimal comparison of the results obtained from all different RNA mutants, we prepared surface densities on the sensor chip that would allow signal to be in range for accurate kinetic measurements (at least 10 RU at the highest injected protein concentration). The variation in response seen in Figure 2 is therefore due to variations in the amount of RNA coated on the sensor chip. Protein was serially diluted in running buffer to the concentrations indicated in Figure 2 and injected at 20°C at a flow rate of 50 μL min⁻¹ for 1 min. Disruption of any complex that remained bound after a 5-min dissociation was achieved using a 1-min injection of 2 M NaCl at 20 μL min⁻¹. Samples with different concentrations of protein were injected in random order, and every injection was performed in triplicate within each experiment. All experiments were done 4–10 times. In order to subtract any background noise from each data set, all samples were also run over an unmodified sensor chip surface and random injections of running buffer were performed throughout every experiment (“double referencing”; Myszka 1999). Data were processed using Scrubber and analyzed using CLAMP XP (Myszka et al. 1998) (developed by the Biomolecular Interaction Facility at the University of Utah [http://www.cores.utah.edu/interaction]) and a simple 1:1 Langmuir interaction model with a correction for mass transport (Myszka and Morton 1998). The results for all mutants were compared to (the wild-type RNA and to each other) using the Student’s t-test to determine whether or not they were statistically significant. Equal and unequal variance for the samples was determined using the F-test. Equilibrium binding constants were calculated for each individual experiment, and these values were used to compute the averages and standard error of the means reported in Table 1.

ACKNOWLEDGMENTS

We thank Dr. Ian Haworth, Dr. Lennart Nilsson, and members of the Laird-Offringa lab for helpful discussions. This material is based on work supported by the NSF under grant no. MCB-0131782 to I.A.L.-O.

Received February 27, 2006; accepted April 17, 2006.
REFERENCES


Liang, S. and Lutz, C.S. 2006. p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. RNA 12: 111–121.


www.rnajournal.org 1177


The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA

Michael J. Law, Andrew J. Rice, Patti Lin, et al.

RNA 2006 12: 1168-1178

References
This article cites 50 articles, 16 of which can be accessed free at:
http://rnajournal.cshlp.org/content/12/7/1168.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.