Complex role of the $\beta_2$-$\beta_3$ Loop in the Interaction of U1A with U1 Hairpin II RNA*

Phinikoula S. Katsamba‡, Melina Bayramyan§, Ian S. Haworth§, David G. Myszka¶, and Ite A. Laird-Offringa‡‡

From the ‡Norris Cancer Center/University of Southern California, Keck School of Medicine, Los Angeles, California 90089-9176, the §Department of Pharmaceutical Sciences, University of Southern California, Los Angeles, California 90089-9121, and the ¶Center for Biomolecular Interaction Analysis, University of Utah, School of Medicine, Salt Lake City, Utah 84112

RNA recognition motifs (RRMs) are characterized by highly conserved regions located centrally on a $\beta$-sheet, which forms the RNA binding surface. Variable flanking regions, such as the loop connecting $\beta$-strands 2 and 3, are thought to be important in determining the RNA-binding specificities of individual RRMs. The N-terminal RRM of the spliceosomal U1A protein mediates binding to an RNA hairpin (U1hpII) in the U1 small nuclear RNA. In this complex, the $\beta_2$-$\beta_3$ loop protrudes through the 10-nucleotide RNA loop. Shortening of the RNA loop strongly perturbs binding, suggesting that an optimal “fit” of the $\beta_2$-$\beta_3$ loop into the RNA loop is an important factor in complexation. To understand this interaction further, we mutated or deleted loop residues Lys$^{50}$ and Met$^{51}$, which protrude centrally into the RNA loop but do not make any direct contacts to the bases. Using BIACORE, we analyzed the ability of these U1A mutants to bind to wild type RNAs, or RNAs with shortened loops. Alanine replacement mutants only modestly affected binding to wild type U1hpII. Interestingly, simultaneous replacement of Lys$^{50}$ and Met$^{51}$ with alanine appeared to alleviate the loss of binding caused by shortening of the RNA loop. Deletion of Lys$^{50}$ or Met$^{51}$ caused a dramatic loss in stability of the U1A-U1hpII complex. However, deletion of both residues simultaneously was much less deleterious. Simulated annealing molecular dynamics analyses suggest this is due to the ability of this mutant to rearrange flanking amino acids to substitute for the two deleted residues. The double deletion mutant also exhibited substantially reduced negative effects of RNA loop shortening, suggesting the rearranged loop is better able to accommodate a short RNA loop. Our results indicate that one of the roles of the $\beta_2$-$\beta_3$ loop is to provide a steric fit into the RNA loop, thereby stabilizing the RNA-protein complex.

*This was supported by National Science Foundation Grant MCB-0131782 (to I. A. L.-O.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†To whom correspondence should be addressed: Dept. of Surgery/Biochemistry and Molecular Biology, University of Southern California, Norris Cancer Center, MC #9176, Rm. NOR 6420, 1441 Eastlake Ave., Los Angeles, CA 90089-9176. Tel.: 323-865-0655; Fax: 323-865-0158; E-mail: ila@usc.edu.

‡This paper is available on line at http://www.jbc.org

© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.

Abstract

The RNA recognition motif (RRM), also known as the ribonucleogprotein (RNP) consensus domain or RNA binding do-

main (RBD), is the most common and best characterized RNA binding domain. It is present in one or more copies in hundreds of RNA binding proteins (1–3). Proteins carrying RRM domains play critical roles in a wide variety of cellular processes, including RNA processing and packaging, mRNA export, translation, and RNA degradation. RRM domains are about 90 amino acids long and fold into a globular structure consisting of a four-stranded antiparallel $\beta$-sheet (the RNA binding surface) backed by two $\alpha$-helices (see Fig. 1, A and B). RRM domains are characterized by the presence of two highly conserved stretches of 8 and 6 amino acids, known as the RNP1 and RNP2 consensus sequences (see Fig. 1A). These consensus sequences lie strategically in the center of the $\beta$-sheet surface and contain conserved aromatic residues critical for RNA binding (2, 4–14). Because the highly conserved nature of the RNP sequences precludes a major role in controlling the specificity of the interaction (1, 2, 11, 15, 16), the less conserved loop regions surrounding the $\beta$-sheet surface are thought to be the predominant determinants of the target specificity of individual RRM proteins. The mechanism by which these flanking regions confer specificity is of considerable interest. Here we describe studies of one such region, the $\beta_2$-$\beta_3$ loop, which plays an important role in RNA binding by the U1A protein. U1A, the A protein of the U1 small nuclear ribonucleoprotein, is the most extensively studied RRM protein, and is often used as a paradigm for RRM domain/RNA interaction. Its N-terminal RRM is necessary and sufficient for high affinity binding to U1 hairpin II (U1hpII), a stem-loop structure in the U1 small nuclear RNA (see Fig. 1D) (17, 18). The $\beta_2$-$\beta_3$ loop of U1A (amino acids 46–52, see Fig. 1, A–C) inserts into the RNA loop, causing the RNA loop residues to splay out (5). This allows the presentation of the bases to highly conserved amino acids within the RNP sequences and the establishment of the critical interactions that mediate high affinity binding (5). The interaction between the RNA and protein loops also appears to be important for providing stability to the RNA-protein complex, as suggested by the increased dissociation rate of complexes containing a Leu$^{49}$ → Met U1A mutant or RNAs with shortened loops (9, 19). Here we describe experiments aimed at testing whether proper complex stability might require a snug “fit” of the RNA and protein loops. We hypothesized that a reduction in the bulk of the $\beta_2$-$\beta_3$ loop might allow it to interact better with RNAs with shorter loops. Using a surface plasmon resonance-based biosensor (BIACORE) (20–22) we investigated the effects of $\beta_2$-$\beta_3$ loop mutants on the interaction of U1A with U1hpII RNAs carrying loops of different sizes. We show that removal of two amino acids from the $\beta_2$-$\beta_3$ loop favors the interaction with smaller RNA hairpins, supporting a steric role of this part of U1A.
**EXPERIMENTAL PROCEDURES**

**Construction of U1A Mutants and Protein Purification**—Throughout these studies, a N-terminal fragment of the human U1A (amino acids 1–101, herein referred to as U1A) containing the first RRMs was used (9). This fragment has been demonstrated to be necessary and sufficient for specific and high affinity binding to U1hpII (17, 18). The U1A fragment was inserted into a modified pET3d vector such that a mscc and a His tag were appended to the C terminus of the RRM, as described previously (9). For cloning purposes, engineered restriction sites were introduced within the U1A coding region. All clones were confirmed using both restriction digests and sequencing. A His6 tag was appended to the C terminus of the RRM, as then “sealing” the backbone by minimizing (1000 cycles) the region around the deletion (from Leu44 to Glu49) in wild type numbering, encompassing the β2-β3 loop while keeping the remainder of the protein and the RNA fixed. Using a similar procedure, the double deletion protein (ΔLYs48,ΔMet51) was constructed by deleting Met51 from the minimized ΔLYs48 protein (Δβ2-β3 loop). The single and double deletion protein:RNA complexes were then minimized for 4000 cycles, using a 12-Å non-bonded cutoff and applying a distance-dependent dielectric constant (ε = 4r) to the electrostatic interactions. The β2-β3 loop region of each of these models (from Leu44 to Glu49, using the numbering of the wild type protein) was then subjected to a brief molecular dynamics simulation (10 ps, time step 0.002 ps, 400 K, other conditions as above) followed by a minimization (using the conditions above) of the structure formed after the 10-ps simulation. This procedure was repeated 50 times for each complex, each time starting from the previous simulation. As a control, a similar series of 50 dynamics and minimization cycles was also performed for the wild type protein:RNA complex.

**RESULTS AND DISCUSSION**

**Side-chain Replacements of β2-β3 Loop Residues Lys50 and Met51 Affect Binding to U1hpII in Kinetically Distinct Ways**—The β2-β3 region of the N-terminal U1A RRMs domain includes amino acids 46–52 (Fig. 1, A and B). Of these, Ser, Leu, Lys, and Met appear to protrude into the RNA loop, with Leu-Met inserted deeply into the middle of the RNA loop (Fig. 1C). Of the three deeply inserted residues, Leu contacts the RNA through a water-mediated hydrogen bond from the main-chain carboxyl to N3 of G4. Lys and Met do not appear to form short-range interactions with the RNA at all. Therefore, these latter two residues are the best suited to examine a steric role of the β2-β3 loop in complex formation and stability. To study the effects of decreasing the bulk of the β2-β3 loop, we therefore either reduced the size of side chains of Lys or Met or removed these residues altogether.
Lys$^{50}$ and Met$^{51}$ were replaced individually or simultaneously with alanine, in a U1A fragment containing the N-terminal RRM (9). The ability of the alanine replacement mutants to bind to wild type U1hpII was first tested, using BIACORE. Highly accurate kinetic data, yielding affinity constants that agree well with those obtained by other methods, can be obtained by BIACORE analysis of the U1A-U1hpII interaction (9). We had previously shown that the high affinity of the U1A-U1hpII complex (∼35 pM) results from a very high association rate and very low dissociation rate (9) (Fig. 2A and Table I). Replacement of Lys$^{50}$ or Met$^{51}$ with an alanine reduced the affinity by 10- and 7-fold, respectively (Fig. 2, D and G, and Table I). Although the main effect of the Lys$^{50}$ → Ala mutation was on the association rate, suggesting that this residue plays a role on the electrostatically mediated association of the two molecules (9), the loss in binding seen with the Met$^{51}$ → Ala mutant was due to a 6-fold increase in the dissociation rate (Table I). This suggests that these two residues stimulate U1hpII binding through different mechanisms. The main role of Lys$^{50}$ appears to be the electrostatic recruitment of the RNA (as confirmed by salt dependence experiments (9)), whereas the main role of Met$^{51}$ appears to be to maximize non-bonded contacts between protein and RNA. The Met$^{51}$ → Ala mutation thus appears to have an effect similar to replacing Leu$^{49}$ with methionine, which also increases complex dissociation (19). When both Lys$^{50}$ and Met$^{51}$ were replaced by alanine, the association rate was slightly faster than the substitution of Lys$^{50}$ alone, suggesting the double replacement mutant might allow compensation for the loss of charge, perhaps by rearrangement of neighboring amino acids (Fig. 2J and Table I). However, the dissociation rate was faster than either of the single mutants; the Lys$^{50}$/Met$^{51}$ → Ala-U1hpII complex exhibited a dissociation rate that was 16-fold faster than that of the wild type U1A-U1hpII complex. A U1A mutant, in which Lys$^{50}$ plus Met$^{51}$ was mutated to Gly plus Ser, has been described previously (32). The Gly plus Ser mutant also exhibited a loss in affinity for U1hpII; however, the effect was only 2.5-fold. The observed difference in the magnitude of the effect of Lys$^{50}$/Met$^{51}$ replacement might arise from the identity of the replacement residues (Ala plus Ala versus Gly plus Ser), from differences in experimental conditions (no absolute $K_D$ values nor kinetic parameters are given by the authors), or from the fact that the Gly plus Ser mutation was made in the context of the full-length protein.

The increased instability of RNA-protein complexes formed by Met$^{51}$ → Ala and the double mutant suggests that these proteins are hampered by a reduced fit between the RNA loop and the protein loop. This reduced fit potentially arises from the reduction in size of the side chains that normally pack against the 5’ half of the U1hpII loop (Fig. 1C). We hypothesized that it might be possible to partially restore the ability of these mutants to bind to RNA by reducing the size of the RNA loop. Accordingly, we analyzed the interaction between the alanine substitution mutants and RNAs lacking one or two loop nucleotides.

**The Double Alanine Substitution Appears to Alleviate the Effect of RNA Loop Shortening**—Of the 10 nucleotides in the U1hpII loop (Fig. 1D), the sequence of only the first seven bases is crucial, as determined by in vitro selection and biochemical experiments (9, 11, 24–27). In contrast, the identity of loop nucleotides 8–10 is irrelevant, because they can be replaced by a polyethylene glycol linker (27). We therefore chose to delete one or two of the spacer nucleotides to examine steric interactions between the β$_2$-β$_3$ loop and the RNA loop.

We compared the interaction of U1A or the alanine mutants with U1hpII, to that of these same proteins with U1hpII lacking C9 or U8-C9 (U1hpIIΔC9 and U1hpIIΔUC, Fig. 1D). In agreement to what was shown previously (9, 27), the interaction of U1A with U1hpII weakens progressively with the removal of each nucleotide (Fig. 2, A–C, and Table I). This is due mainly to an increase in the dissociation rate (Table I and Ref. 9). When the alanine replacement mutants were used, reducing the RNA loop length also caused a marked increase in dissociation rates compared with the same proteins bound to U1hpII (Fig. 2, D–L, and Table I). Interestingly, the double alanine replacement mutant showed a 10-fold smaller negative effect of RNA loop shortening as visualized by comparison of the dissociation constant ($K_D$) of the RNA lacking two loop nucleotides relative to the wild type RNA (Table I, penultimate column). Although the affinity of wild type U1A for the mutant RNAs decreased ~3000-fold with the removal of two nucleotides, the affinity of the double alanine replacement mutant appeared to be less affected (~300-fold decrease). Thus, removing spacer residues from the RNA appears to be less deleterious for the double alanine mutant than for wild type U1A or the single alanine replacements, suggesting that a reduction in size of residues in the β$_2$-β$_3$ loop might compensate to a limited extent for RNA loop shortening. However, the double alanine replacement (which reduces binding affinity) and the shortening of the RNA loop (which has a dramatic effect on binding) were not sufficiently complementary to offset the loss in binding ability, so that a net increase in $K_D$ was still observed for the mutant protein-RNA combination (Table I, final column).

To verify that the binding we observed to the various RNAs in these experiments was not due to nonspecific binding that might occur at higher protein concentrations, all proteins were tested for binding to a U1hpII target (“reverseU1hpII”), in which the order of the loop nucleotides had been inverted (from 5’-AUUGGACGUCC to 5’-CCUCACGUUA), using protein concentrations identical to those used for the wild type and deletion target RNAs. No binding to this RNA was observed for any of the proteins at these concentrations. Thus, the observed ability of U1A and alanine mutants to bind to the wild type and deletion RNAs is specific and depends on the identity of the loop nucleotides.

Together, our observations suggest that, although the double alanine replacement is somewhat beneficial in interactions with RNAs carrying smaller loops, a larger adjustment than that provided by two side-chain substitutions may be required to restore binding. We therefore explored the effect of completely removing Lys$^{50}$ and/or Met$^{51}$ from the β$_2$-β$_3$ loop.

**Deletion of Residues from the β$_2$-β$_3$ Loop Dramatically Reduces Binding to U1hpII**—Three deletion mutants were generated: ΔLys$^{50}$, ΔMet$^{51}$, and ΔLys$^{50}$/Met$^{51}$. Deletion of Lys$^{50}$ resulted in a ~9,000-fold loss in binding affinity for the wild type RNA target, which can be attributed largely to an increase in the dissociation rate (~1,000-fold, Fig. 2M and Table II). The removal of Met$^{51}$ was even more deleterious, resulting in a ~50,000-fold loss in binding affinity, again attributable almost exclusively to a large decrease in complex stability (~8,000-fold increase in dissociation rate, Fig. 2P and Table II). Surprisingly, deletion of both residues was less calamitous, showing a ~900-fold loss in affinity for U1hpII compared with the wild type U1A (Fig. 2S and Table II). This loss was again based largely on an increase (~200-fold) in the dissociation rate. Thus, in all cases, the loss in affinity was almost entirely due to a dramatic loss in complex stability. Importantly, the association rate changed less than 10-fold for each of the three mutants (Table II), indicating that the initial step in complex formation is largely unaffected (9). Because the weak affinity of the deletion mutants necessitated the use of very high protein concentrations, we tested the ability of the proteins to bind to
**Fig. 2.** Binding studies of U1A or \( \beta_2\beta_3 \) loop mutants with U1hpII or U1hpII loop deletion RNAs. BIAcore analysis of the interaction between U1A variants and RNAs. Panels from left to right show data obtained with a given protein, as indicated at left: U1A (A–C), Lys\(^{50} \rightarrow\) Ala (D–F), Met\(^{51} \rightarrow\) Ala (G–I), Lys\(^{50}\)Met\(^{51} \rightarrow\) Ala (J–L), Lys\(^{50}\) (M–O), Met\(^{51}\) (P–R), and Lys\(^{50}\)Met\(^{51}\) (S–U). Columns show the interaction of various proteins with a given RNA, as indicated at the top: U1hpII (A, D, G, J, M, P, and S), U1hpIIAC9 (B, E, H, K, N, Q, and T), and U1hpIIUC (C, F, I, L, O, R, and U). To allow optimal comparison of binding behaviors, the same RNA surfaces carrying the three different synthetic 5'-biotinylated RNAs, each coated on a different flow cell, were used for all data sets shown (A–U). Increasing protein concentrations (in triplicate and in randomized order) were injected over the RNA surfaces at the concentrations shown in the leftmost column and were identical for all RNAs except in the bottom row, where the concentrations used in panel S (indicated) were lower than those used in panels T and U (indicated in T). The
reverse U1hpII and an unrelated single-stranded 19-nucleotide RNA. None of the proteins bound detectably to the single-stranded RNA. However, wild type U1A and the deletion mutants were able to bind the reverse U1hpII RNA at high concentrations, with an apparent affinity in the low micromolar range (≈4 μM for wild type U1A). For ΔLys50 and ΔLys50-Met51, this background binding was at least 10-fold weaker than the binding to U1hpII, supporting the specificity of the interactions of these proteins with U1hpII. However, ΔMet51 bound reverse U1hpII with an affinity of ≈6 μM, which approaches its affinity for wild type U1hpII (≈2 μM, Table II). This suggests that the ability of ΔMet51 to bind to U1hpII is largely due to sequence-independent binding. This is in agreement with the observation that the binding of ΔMet51 to U1hpII appears to be primarily based on association and that the formed complex is highly unstable. As we have suggested previously, association of U1A with U1hpII appears to be mediated by electrostatic interactions between positively charged residues on the protein and negative charges on the RNA (9). Because such negative charges would be derived from the phosphodiester backbone, the affinity of U1A for U1hpII is increased when these charges are removed by the naturally occurring protein U1A.

| Protein | RNA       | $k_a$ | $k_d$ | $K_D$ | Compared to the same protein on wild type U1hpII | Compared to wild type U1A on the same RNA |
|---------|-----------|-------|-------|-------|-----------------------------------------------|------------------------------------------|
| U1A     | U1hpII    | 9.5 ± 0.1 × 10^6 | 3.2 ± 0.4 × 10^-4 | 3.4 ± 0.4 × 10^{-11} | 1 | 2 |
| ΔLys50  | U1hpII    | 1.5 ± 0.1 × 10^6 | 5.4 ± 0.6 × 10^-4 | 3.6 ± 0.5 × 10^{-10} | 1 | 1 |
| ΔMet51  | U1hpII    | 3.8 ± 0.3 × 10^6 | 1.9 ± 0.1 × 10^-3 | 2.3 ± 0.2 × 10^{-10} | 1 | 1 |
| ΔLys50-Met51 | U1hpII | 2.7 ± 0.1 × 10^6 | 5.1 ± 0.1 × 10^-3 | 1.9 ± 0.1 × 10^-9 | 1 | 50 |
| U1A ΔC9 | U1hpII    | 3.3 ± 0.4 × 10^6 | 2.3 ± 0.1 × 10^-2 | 6.8 ± 0.9 × 10^-9 | 200 | 1 |
| U1A ΔUC  | 1.9 ± 0.9 × 10^6 | 2 ± 1 × 10^-3 | 1.2 ± 0.8 × 10^-7 | 3000 | 1 |
| ΔLys50 | U1hpII    | 1.9 ± 0.3 × 10^6 | 1.3 ± 0.5 × 10^-1 | 7 ± 3 × 10^{-8} | 200 | 10 |
| ΔMet51 | U1hpII    | 1.2 ± 0.3 × 10^7 | 2.5 ± 0.5 × 10^-1 | 2.1 ± 0.7 × 10^-8 | 100 | 5 |
| U1A ΔC9 | U1hpII    | 6 ± 4 × 10^6 | 10 ± 5 × 10^-1 | 2 ± 1 × 10^-7 | 100 | 30 |
| U1A ΔUC  | 1.7 ± 0.8 × 10^6 | 8 ± 3 × 10^-1 | 5 ± 3 × 10^-7 | 300 | 5 |

$^a$ The standard error of the mean, based on three independent experiments (each consisting of randomized injections repeated at least twice), is given.

$^b$ The calculated standard error of the $K_D$ value ($K_D = k_d/k_a$) is given.

$^c$ Fold loss in affinity: $K_D$ value for deletion RNA relative to wild type U1hpII.

$^d$ NA: not applicable. A comparison is difficult to make, because this protein/RNA combination binds with an affinity close to non-sequence-specific binding, as determined by binding to reverse U1hpII.
phate backbone, they are expected to be sequence-independent. Yet association could well be structure- and/or charge-dependent, requiring presentation of the RNA in a defined context for optimal association. This is supported by the observation that no binding of ΔMet51 was observed to the single-stranded RNA target.

The loss in complex stability for the deletion mutants was much greater than that observed with the alanine replacement mutants. Taking into account the lack of any direct contacts between Lys50 or Met51 and the RNA, and the limited effect of the alanine replacement mutations on RNA binding, this suggests that it is not the loss of the Lys50 or Met51 side or main chains (and accompanying electrostatic and steric contributions) per se that dramatically destabilizes the complex but effects of the removal of these residues on the conformation of the β2-β3 loop region. Removal of 1 or 2 residues could lead to rearrangements in the flanking amino acids and the RNA or protein contacts that these residues make. Three of the loop residues directly contact the RNA (5): the main-chain carbonyl of Arg47 makes a hydrogen bond to G11 in the closing base pair of the RNA stem, the main-chain carbonyl group of Leu49 makes a water-mediated hydrogen bond to G4, and Arg42 makes hydrogen bonds to A1 and the closing C-G pair. In addition, certain loop residues are also involved in intramolecular contacts within the protein (5, 33): the main-chain carbonyl of Lys50 contacts Gln54, Arg42 makes a similar contact to Gln54 and contacts the main-chain carbonyl groups of Arg47 and Ser48. Thus, these loop residues appear to form a network of interactions that directly or indirectly affect Gln54. Gln54 is a key residue that appears to be involved in positioning Tyr13, which stacks on C5 and is one of two critically important aromatic residues located in the RNP consensus sequences (the other is Phe56) (4–6, 8, 10–12, 14, 33, 34). Deletion of Lys50 or Met51 could result in a shift in this intricate network of contacts, indirectly interfering with Tyr13/C5 stacking. This is supported by NMR studies that show rearrangements of the β2-β3 loop in Tyr13 → Phe and Gln54 → Gln mutants (33). Loss of the Tyr13 interaction leads to a dramatic loss in binding affinity, reminiscent of the consequences of a Phe56 → Ala mutation in U1A (9, 12, 33). Thus, it would appear that the most likely mechanism for the strong destabilization of the complex seen with the Lys50 and/or Met51 deletion mutants lies in the local rearrangement of the β2-β3 loop region, and the effects thereof, both on RNA contacts by other β2-β3 loop residues as well as on a network of interactions positioning Tyr13. Although their effects are strong, these rearrangements are likely relatively small in the context of the whole RRM domain, because the high association rate of the deletion mutants suggests that the global conformation of the domain and its charge distribution is largely intact.

Simultaneous Deletion of Lys50 and Met51 Partially Compensates for RNA Loop Shortening—Because the β2-β3 loop deletion mutants were still able to bind to U1hpII (albeit with greatly increased dissociation rates) we wished to explore whether the deletions in the β2-β3 loop would affect the ability to interact with the RNAs containing shorter loops. Therefore, we examined the interaction between the β2-β3 loop deletion mutants and U1hpIIAC9 or U1hpIIACUC RNAs. ΔMet51, of which binding to U1hpII appears to be sequence-independent (see above), bound even more weakly (KD ~ 20 μM) to U1hpIIACUC than to reverseU1hpII. This effect is due to a loss in association rate, suggesting that the absence of two phosphates from the RNA loop can weaken even sequence-independent background binding. ΔLys50, which could still bind specifically to U1hpII, when bound to U1hpIIACUC exhibited a similar loss in association rate. When combined with its increased dissociation from U1hpIIACUC, the binding observed (KD ~ 10 μM) dropped to background binding.

Only the double deletion mutant, ΔLys50Met51, bound U1hpIIACUC with an affinity above background. Interestingly, ΔLys50Met51 appears to be affected less by deletion of the two spacer nucleotides than U1A. U1A shows a loss of ~3000-fold in binding affinity when confronted with an RNA loop shortened by two nucleotides (Table I, penultimate column), whereas ΔLys50Met51 becomes only 20-fold weaker (Table II, penultimate column). The partial complementarity of RNA and protein loop deletions can also be visualized by comparing the affinity of ΔLys50Met51 for a given target RNA to the affinity of wild type U1A to that same RNA. Although ΔLys50Met51 exhibits a loss in affinity of ~900-fold in binding to U1hpII, this loss is only ~5-fold on the double deletion RNA (Table II, final column). Again, deleting Lys50 and Met51 is less harmful in the context of an RNA with a smaller loop, suggesting changes in these areas of protein and RNA are somewhat complementary. However, as with the double alanine mutant, the gain in the ability of the double deletion protein to bind to RNA with shortened loops is offset by a much larger loss in general RNA-binding activity (see below), so that the combination of RNA and protein deletion mutants does not result in an absolute gain in affinity.

A Complex Role for the β2-β3 Loop Region—The observed beneficial interaction between RNAs with shorter loops and proteins with reduced β2-β3 loop supports a steric role for the β2-β3 loop region. We envision this role as assisting in the spaying out of the RNA loop residues, as well as in achieving a snug “fit” of this part of the protein into the RNA, stabilizing the complex. However, the dramatic reduction in binding affinity exhibited by the Lys50 and Met51 deletion mutants suggests the β2-β3 loop plays one or more additional roles in complex formation. This region of U1A was originally thought to be disorganized in the free protein, because it was visible in only one of the two subunits of the crystal (4). It has since become clear that the β2-β3 loop contains a single turn of a helix. This turn is clearly seen in NMR studies of the free protein, although some mobility is suggested in this region (35). The same turn can be seen in the complex (5), suggesting it is preformed to some extent and important for complexation. A somewhat rigid structure in the β2-β3 loop would be helpful as the protein interacts with the RNA loop (which is thought to be mobile (24, 35)), because structuring a disorganized β2-β3 loop as the complex forms is likely to be energetically costly (33). The single turn in the β2-β3 loop is composed of residues Leu49-Met51, so that its disruption by the Lys50 and Met51 deletion mutants seems inevitable. Accordingly, we suggest that the β2-β3 loop also has a structural role and that one possible reason for the loss of affinity of ΔLys50 or ΔMet51 is the energy required to accommodate the altered loop into the complex.

Another possible reason for the load in affinity is the likely repositioning of residues flanking Lys50 or Met51, which would affect their direct and indirect interactions with the RNA and amino acids elsewhere in the protein. Of particular concern would be the effect on the network of interactions involving residues in the highly conserved RNP consensus sequences, such as Tyr13, Phe56, and Gln54. NMR studies of U1A have shown that mutations in these residues lead to rearrangements of the β2-β3 loop residues (in particular Leu49 and Lys50) (10, 33). Thus, the β2-β3 loop region can also be seen as an integral part of the RNA-binding platform, connected with and affecting the positioning of critical residues on the β-sheet surface.

In light of these possible roles of the β2-β3 loop, the observation that deletion of Lys50 and Met51 simultaneously is less disruptive than the individual deletions could be explained in
shortened/H9252 rearrangement of the complex over the 50 cycles, implying that a conformational decrease in the energy of the double deletion protein dynamics and energy minimization) is shown. The position of the nature of the conformational change in the double deletion single deletion complexes did not decrease significantly. The focused on the possible conformational changes in the short- the single and double amino acid deletions. These calculations predict the possible structural changes that might result from 50 cycles of molecular dynamics and energy minimization to an attempt to distinguish between these possibilities, we used more strongly affecting the (in)direct interactions supported by ing the deletion more than removal of two nucleotides, thereby might disorganize the trajectory of the RNA is indicated in blue.

Fig. 3. Rearrangement of residues Leu49 and Arg47 in the \(\Delta\text{Lys}^{50}\text{Met}^{51}\)-U1hpII complex. An overlay of the \(\beta_2\beta_3\) loop region of the minimized wild type U1A-U1hpII complex onto a representative structure of the \(\Delta\text{Lys}^{50}\text{Met}^{51}\)-U1hpII complex (generated by molecular dynamics and energy minimization) is shown. The position of the \(\beta_2\beta_3\) loop in the U1A-U1hpII complex is shown in green, whereas that of the shortened \(\beta_2\beta_3\) loop in the \(\Delta\text{Lys}^{50}\text{Met}^{51}\)-U1hpII complex is shown in pink. Note the strongly shortened trajectory of the loop, and the dramatic change in the positions of the side chains of residues Arg47 (bright green in the wild type complex, pink in the deletion mutant complex) and Leu49 (bright green in the wild type complex, red in the deletion mutant complex). Other parts of the protein are shown in gray. The trajectory of the RNA is indicated in blue.

several ways. First, although removal of a single amino acid might disorganize the \(\beta_2\beta_3\) loop and increase its mobility, removal of 2 residues might limit this mobility. Thus, less energy would be required to organize the \(\beta_2\beta_3\) loop during complex formation. Alternatively, removal of a single nucleotide might distort the position of \(\beta_2\beta_3\) loop amino acids flanking the deletion more than removal of two nucleotides, thereby more strongly affecting the (in)direct interactions supported by this region. Lastly, it is possible that removal of 2 amino acids results in major changes in the position of flanking amino acids but that these changes are in fact more favorable to RNA binding than those caused by removal of a single amino acid. In an attempt to distinguish between these possibilities, we used 50 cycles of molecular dynamics and energy minimization to predict the possible structural changes that might result from the single and double amino acid deletions. These calculations focused on the possible conformational changes in the shortened \(\beta_2\beta_3\) loop (from Leu44 to Gln54, wild type protein numbering). Two results of interest emerged. There was a steady decrease in the energy of the double deletion protein-RNA complex over the 50 cycles, implying that a conformational rearrangement of the \(\beta_2\beta_3\) loop might be occurring, whereas the minimized energies of the wild type complex and the two single deletion complexes did not decrease significantly. The nature of the conformational change in the double deletion protein is illustrated in Fig. 3. Arg47 and Leu49 in the \(\Delta\text{Lys}^{50}\text{Met}^{51}\)-U1hpII complex may compensate for the loss of Lys50 and Met51, respectively. In the shortened loop, Leu49 is able to move into the space created by the deletion of Met51, thus playing the role of the hydrophobic space-filling amino acid in the center of the protein loop/RNA loop interaction. Arg47 is able to rotate toward the center of the RNA loop and at least partially compensate for the loss of Lys50. These motions are apparently only possible in the double deletion protein-RNA complex, because neither were observed in the simulations of the \(\Delta\text{Lys}^{50}\text{U1hpII} and \Delta\text{Met}^{51}\text{U1hpII complexes. These observations suggest that repositioning of flanking residues is at least in part responsible for the higher RNA-binding affinity of the double deletion U1A mutants compared with the single deletion mutants. This repositioning of amino acids in response to RNA binding may be illustrative of the conformational flexibility that can occur in RNA-protein interactions. A more complete understanding of such phenomena could aid in the rational design of engineered proteins that will bind specific RNA molecules with high affinity.

Comparison of the Role of \(\beta_2\beta_3\) Loops in Different RRM Proteins—The prominent role of the \(\beta_2\beta_3\) loop in the U1A-U1hpII interaction may be directly related to the fact that U1hpII carries a 10-nucleotide loop, because this sizeable loop offers a docking site for the protein. A smaller RNA loop or an unstructured RNA may not provide as much access to the bases or as good a structural “handle” for the protein. In the autoregulatory interaction of U1A with the Polyadenylation Inhibition Element (PIE) in the U1A mRNA (a complex structure containing two conserved U1A binding sites similar to U1hpII), \(\beta_2\beta_3\) loop residues protrude into the RNA loops (16). The RNA loops are comparable in size to the U1hpII loop but differ because a double-stranded stem linking the two binding sites protrudes from the spacer region. In this complex, \(\beta_2\beta_3\) loop residues make a large number of contacts with the RNA, locking the protein into the hole defined by the RNA loop (16). The U2B+ protein (which is closely related to U1A) forms a ternary complex with the U2A protein and the U2 hairpin IV RNA (a stem with an 11-nucleotide loop). In this complex, the \(\beta_2\beta_3\) loop region is again inserted into the RNA loop and makes extensive contacts with the spacer nucleotides, which form a rigid step ladder (36). Thus, the role of the \(\beta_2\beta_3\) loop is similar in the interactions of U1A and U2B+ with their cognate RNA hairpins, all of which carry loops of comparable sizes. A very different case is presented by nucleolin, which carries four RRM domains, the first two of which are required to bind to the Nucleolin Regulatory Element (NRE), an RNA hairpin with a six-nucleotide loop (37). The smaller RNA loop accommodates only two amino acids, and these are derived from different parts of the protein: Phe56, which lies at the border of a very short \(\beta_2\beta_3\) loop in the first RRM, and Lys94, which lies in the hinge region between RRMs 1 and 2. Phe56 and Lys94 protrude into the loop from different sides and clamp the RNA between the two RRMs (14). Possibly, a loop that is too small limits the extent to which the \(\beta_2\beta_3\) loop region can participate in binding and may thereby necessitate the involvement of additional RRM domains. Obviously, RNAs lacking a loop or other secondary structure altogether would further limit the role of the \(\beta_2\beta_3\) loop. It is therefore of interest that the specific binding of unstructured (single-stranded) RNAs apparently requires two, three, or even four RRMs, as is evident from studies of poly(A) binding protein (38), Sex-lethal (39), hnRNPA1 (40), U2AF (41), HuD (23, 42), and nucleolin (43). In those cases where the structures of complexes between multiple RRMs and unstructured RNA have been solved, the \(\beta_2\beta_3\) loops (which are generally longer in these RRMs than the U1A N-terminal RRM) appear to contribute to RNA binding by extending the RNA binding surface or “reaching out” to the RNA, which is strung across the \(\beta\)-sheet surfaces (14, 38, 39, 42). These observations suggest that the nature of the bound RNA target is a key element defining the possible roles for the \(\beta_2\beta_3\) loop in complex formation. It will be of interest to determine the effects of \(\beta_2\beta_3\) loop mutations on proteins that bind to non-hairpin target RNAs, such as single-stranded RNA.

Acknowledgments—We thank members of the Laird-Offringa laboratory for helpful criticism.

REFERENCES
1. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
2. Varani, G., and Nagar, R. (1998) Annu. Rev. Biophys. Biomol. Struct. 27, 407–445
3. Perez-Canadillas, J. M., and Varani, G. (2001) Curr. Opin. Struct. Biol. 11, 53–58
