Contribution of RNA Conformation to the Stability of a High-Affinity RNA−Protein Complex

Sarah J. Luchansky, Scott J. Nolan, and Anne M. Baranger*

Department of Chemistry, Wesleyan University
Middletown, Connecticut 06459

Received March 8, 2000

RNA folds into complicated secondary and tertiary structures that present RNA-binding proteins with functional groups in diverse orientations.[1] Proteins that bind to RNA sequence-specifically usually bind to single-stranded regions of RNA because they expose a more accessible and sequence-dependent array of functional groups than do A-form RNA helices. Single-stranded regions are often conformationally restrained by adjacent helices. This preorganization may contribute to the stability of many, if not most, RNA−protein complexes, but quantitative analyses of this contribution are rare. We show that RNA secondary structure contributes at least 3.5 kcal/mol to the stability of the U1A−RNA complex, in large part by preorganizing nucleotides of the single-stranded target sequence for optimal interactions with U1A.

U1A is a component of the U1 small nuclear ribonucleoprotein particle (snRNP) that forms part of the spliceosome, which splices most eukaryotic pre-mRNA.[2] An N-terminal RNP domain of U1A is responsible for RNA recognition and binds to stem loop 2 in U1 snRNA (Figure 1A).[3] The RNP domain is one of the most common and best-characterized RNA binding domains. Proteins with this domain are able to bind single-stranded RNA in a variety of structural contexts.[4]

The N-terminal RNP domain of U1A binds RNA with low affinity, though specifically, when its recognition sequence is located in a linear RNA.[5] X-ray crystallography has shown that the N-terminal RNP domain of U1A primarily contacts the first seven nucleotides of the RNA loop (AUUCGAC) and the CG base pair that closes the loop, but makes few contacts with other base pairs in the stem (Figure 1B).[6] Mutagenesis experiments suggest that the most important contacts in the complex are with the RNA loop and the closing CG base pair.[5b,7] Therefore, it is likely that the role of the stem is either to preorganize the loop region or maintain the CG base pair at the base of the loop, or both.

To probe the energetic contribution of stem loop structure to complex stability, we sought a minimal RNA target site in which the loop secondary structure and the CG base pair closing the loop were maintained. Thus, we designed RNA sequences in which the majority of the stem was replaced by a disulfide cross-link (Figure 2). We used a cross-link developed by Glick and co-workers that has been observed to stabilize RNA hairpins and tRNA secondary structures.[8] The disulfide cross-link enabled a direct comparison between the U1A complexes of linear and stem

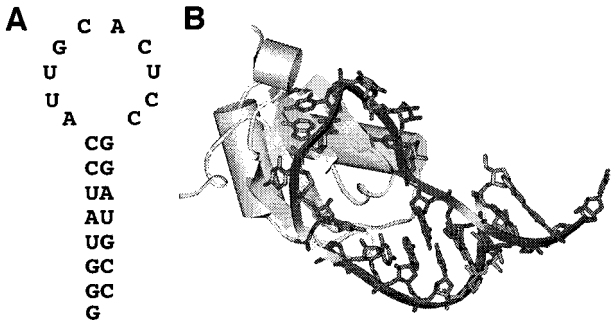

Figure 1. (A) Stem loop 2 of U1 snRNA. (B) Ribbon diagram of the complex formed between the N-terminal RNP domain of U1A and stem loop 2 from the X-ray cocrystal structure.[6]

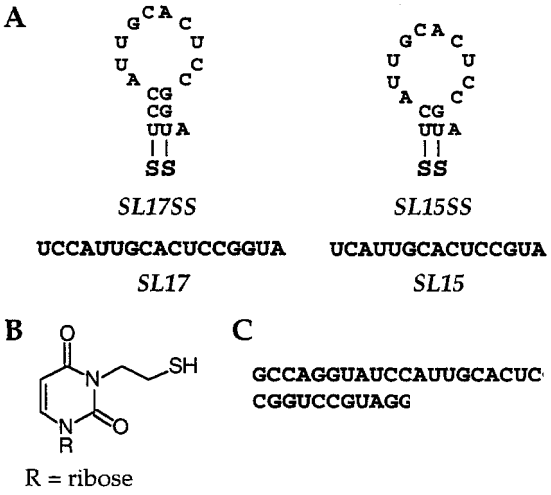

Figure 2. (A) Cross-linked and linear RNA sequences synthesized to probe the contribution of secondary structure to complex stability. (B) Thiol-modified uridine used to form cross-links.[8c] (C) Duplex RNA target site synthesized to probe the contribution of the CG base pair closing the loop to complex stability.

loop RNA oligonucleotides of the same length and sequence.[9] SL17SS maintains two CG base pairs found at the top of the stem, while SL15SS maintains only one (Figure 2A). SL17 and SL15 are identical to the cross-linked sequences except they contain unmodified uridines and therefore lack the stem and the disulfide cross-link. Since SL17 and SL15 are only able to make two and one stem base pairs, respectively, they were linear under the conditions of our binding experiments discussed below.

Thiol-modified RNA sequences protected as the *tert*-butyl disulfide were prepared and subsequently deprotected by reduction with DTT. Upon removal of DTT, cross-link formation occurred spontaneously at room temperature in air, as shown by analytical denaturing PAGE (Figure 3). The compact, oxidized RNA has a higher mobility than the reduced RNA.[9] ^{32}P-labeled oxidized RNA oligonucleotides were gel-purified. Treatment of the oxidized RNA with DTT formed a product with an identical mobility as the reduced RNA, demonstrating reversible oxidation.

Equilibrium binding affinities of the N-terminal RNP domain of U1A (U1A101) for the disulfide cross-linked and linear RNA target sites were measured by polyacrylamide gel mobility shift assays. Both cross-linked stem loops, SL17SS and SL15SS, bound U1A101 with high affinity (Table 1). Of course, these complexes were less stable than the U1A101 complex of the wild-type stem loop. The elimination of favorable interactions with the stem sequences, the creation of unfavorable interactions between the

(1) Varani, G. *Acc. Chem. Res.* **1997**, *30*, 189−195.
(2) Green, M. R. *Annu. Rev. Cell Biol.* **1991**, *7*, 559−599.
(3) Scherly, D.; Boelens, W.; van Venrooij, W. J.; Dathan, N. A.; Hamm, J.; Mattaj, I. W. *EMBO J.* **1989**, *8*, 4163−4170.
(4) Varani, G.; Nagai, K. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 407−445.
(5) (a) Tsai, D. E.; Harper, D. S.; Keene, J. D. *Nucleic Acids Res.* **1991**, *19*, 4931−4936. (b) Hall, K. B. *Biochemistry* **1994**, *33*, 10076−10088.
(6) Oubridge, C.; Ito, N.; Evans, P. R.; Teo, C. H.; Nagai, K. *Nature* **1994**, *372*, 432−438.
(7) (a) Stump, W. M.; Hall, K. B. *RNA* **1995**, *1*, 55−63. (b) Jessen, T.-H.; Oubridge, C.; Teo, C. H.; Pritchard, C.; Nagai, K. *EMBO J.* **1991**, *10*, 3447−3456.
(8) (a) Goodwin, J. T.; Glick, G. D. *Tetrahedron Lett.* **1994**, *35*, 1647−1650. (b) Maglott, E.; Glick, G. *Nucleic Acids Res.* **1998**, *26*, 1301−1308. (c) Goodwin, J. T.; Osborne, S. E.; Scholle, E. J.; Glick, G. D. *J. Am. Chem. Soc.* **1996**, *118*, 5207−5215.
(9) Allerson, C. R.; Verdine, G. L. *Chem. Biol.* **1995**, *2*, 667−675.

10.1021/ja000837u CCC: $19.00
Published on Web 07/08/2000

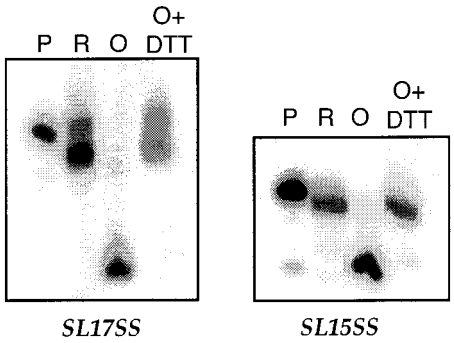

Figure 3. Denaturing polyacrylamide gel electrophoresis of SL17SS and SL15SS. Lane 1: The thiol-modified RNA with thiols protected as *tert*-butyl disulfides (P). Lane 2: The reduced thiol-modified RNA sequences (R). Lane 3: The oxidized, cross-linked RNA (O). Lane 4: Reduction of the oxidized RNA from lane 3 with DTT.

Table 1. Binding Affinities of U1A101 for the Linear and Cross-Linked RNA Target Sites

RNA	Kd ($\times 10^{-7}$ M)[a]	ΔG (kcal/mol)[b]	$\Delta\Delta G$ compared to stem loop 2[c]
stem loop 2	0.005 ± 0.002	-12.7 ± 0.3	0
SL17SS	0.06 ± 0.02	-11.1 ± 0.2	1.6
SL17	4.4 ± 0.5	-8.7 ± 0.1	4.0
SL15SS	0.8 ± 0.1	-9.7 ± 0.1	3.0
SL15	300 ± 40	-6.2 ± 0.1	6.5

[a] Kd values were measured by gel mobility shift assays. Binding reactions were performed at 25 °C for 1 h in 10 mM Tris-HCl (pH 7.4), 250 mM NaCl, 1 mM EDTA, 0.5% Triton X-100, 1 mg/mL tRNA in a total volume of 10 μL. RNA concentration was 25 pM. [b] ΔG is the free energy of association of the complex. [c] $\Delta\Delta G$ is the difference in free energy of U1A101 association with stem loop 2 and with the indicated RNA.

disulfide cross-link and U1A101, or small conformational changes in the RNA stem loop could reduce affinity.

The difference in the binding affinity of U1A101 for the cross-linked stem loop SL17SS and the linear RNA SL17 was 2.4 kcal/mol, while the difference in the binding affinity of U1A101 for the shorter cross-linked stem loop SL15SS and linear RNA SL15 was 3.5 kcal/mol (Table 1). Because SL17SS and SL15SS differ from SL17 and SL15 by only the disulfide cross-links, the differences in stability must result from the secondary structure that is enforced in the cross-linked sequences or from favorable interactions between U1A101 and the disulfide linker itself. Since the ethyl disulfide linkers are unnatural additions to the RNA target site, it is unlikely that interactions between U1A and the disulfide cross-link are responsible for the 73-fold (SL17SS) and 375-fold (SL15SS) decrease in Kd observed for the U1A101 complexes of the cross-linked oligonucleotides compared to the analogous linear sequences. Instead, the high affinity of U1A101 for the disulfide cross-linked oligonucleotides is most likely due to their stable stem loop structures prior to binding.

The difference in binding energy between SL15SS and SL15 (3.5 kcal/mol) was larger than that between SL17SS and SL17 (2.4 kcal/mol). SL17 is able to form two CG base pairs, while SL15 is only able to form one. Although SL15 and SL17 will be linear when free in solution, these sequences may form stem loop structures on the surface of the protein. If so, the difference in stability of the U1A101 complexes of SL17SS and SL17 would be an underestimate of the contribution of secondary structure to complex stability because the two CG base pairs would contribute to stabilizing stem loop structure in SL17. Since SL15 is only able to make one CG base pair, the difference in U1A101 affinity between SL15SS and SL15, 3.5 kcal/mol, is a more accurate estimate of the minimum contribution of stem loop structure to complex stability.

The stabilization of the U1A101 complexes of the cross-linked RNA compared to the linear RNA is not due to favorable interactions between the stem and U1A101, since most of the stem was eliminated in the cross-linked target sites. However, our experiments with the cross-linked RNA do not distinguish between stabilizing contributions from nucleotide preorganization and from interactions with the closing CG base pair. Previous studies have suggested that a precise loop conformation may not be essential for high affinity binding; a nucleotide or a variety of lengths of poly(ethylene glycol) linkers can be inserted into the 3′ side of the loop with no decrease in binding affinity.[10] In contrast, mutation of the CG base pair to GC or UU has been found to result in approximately 2 kcal/mol destabilization of the complex.[7a]

To distinguish between the stabilizing effects of maintaining the CG base pair and of preorganizing loop conformation, two RNA oligonucleotides were synthesized that would anneal to form the stem of the stem loop, but would contain the loop sequence in a linear structure (Figure 2C). The CG base pair is still present in this target site, but the stem loop structure is absent. This duplex RNA contained a longer stem sequence than that found in stem loop 2 so that the duplex would remain annealed during the binding experiments. We confirmed that the longer stem did not alter the affinity of U1A101 for the stem loop target site. The duplex RNA bound U1A101 poorly with a dissociation constant of 1.6×10^{-5} M. We also examined the U1A101 affinity of the longer strand of the duplex, without the complementary shorter strand. This RNA also bound U1A101 poorly with a dissociation constant of 6.8×10^{-5} M. It is unlikely that differences between the duplex target site and stem loop 2 in the kinetics of opening of the CG base pair are responsible for the extremely low affinity of U1A101 for the duplex target site. Although the lifetime of the CG base pair closing the loop in stem loop 2 has not been measured, several studies of base pairs closing loops have found them to be substantially more dynamic than internal helical base pairs and to approach the dynamics of terminal base pairs.[11] Therefore, the low binding affinity of the duplex suggests that the stem loop structure is required to constrain the conformation of the target site.

We have found that RNA secondary structure contributes at least 3.5 kcal/mol to the stability of the U1A101−stem loop 2 complex. The ability of U1A101 to induce stem loop structure in a linear RNA target site that can form two of the stem base pairs is suggested by the higher binding affinity of SL17 than SL15. However, U1A101 is not able to recognize a duplex target site that maintains the stem base pairs, including the CG base pair at the top of the stem, but contains the remainder of the target site in a linear sequence. Therefore, an essential role of RNA secondary structure in the U1A101−stem loop 2 complex is the restriction of target site conformation by adjacent helical regions.

Acknowledgment. We are grateful to Prof. K. Nagai for the expression vector for U1A101. Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this work. Funding was also provided by NIH (GM56857-02). A.M.B. is an Alfred P. Sloan Research Fellow. S.J.L. was supported by awards for undergraduate research from Pfizer Inc. and the Howard Hughes Medical Institute.

Supporting Information Available: Description of peptide and RNA syntheses and peptide−RNA binding reactions, a representative gel shift assay, and plots of the equilibrium dissociation of the RNA−peptide complexes (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA000837U

(10) Williams, D. J.; Hall, K. B. *J. Mol. Biol.* **1996**, *257*, 265−275.
(11) (a) Cain, R. J.; Glick, G. D. *Nucleic Acids Res.* **1997**, *25*, 836−842. (b) Nonin, S.; Jiang, R.; Patel, D. J. *J. Mol. Biol.* **1997**, *268*, 359−374.