RNA (1998), *4*:331–339. Cambridge University Press. Printed in the USA.

Defining the orientation of the human U1A RBD1 on its UTR by tethered-EDTA(Fe) cleavage

DAVID L. BECK,[1] W. THOMAS STUMP,[2] and KATHLEEN B. HALL[2]

[1]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110, USA
[2]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, USA

ABSTRACT

The N-terminal RNA binding domain of the human U1A protein (RBD1) specifically binds an RNA hairpin of U1 snRNA as well as two internal loops in the 3′ UTR of its own mRNA. Here, a single cysteine has been introduced into Loop 1 of RBD1, which is subsequently used to attach (EDTA-2-aminoethyl) 2-pyridyl disulfide-Fe^{3+} (EPD-Fe). This EDTA-Fe derivative is used to generate hydroxyl radicals to cleave the proximal RNA sugar–phosphate backbone in the RNA–RBD complexes. RBD1(K20C)–EPD-Fe cleaves the 5′ strand of the RNA hairpin stem, centered four base pairs away from the base of the loop, and cleaves the UTR in two places, again centered on the 5′ side of the fourth base pair from each internal loop. These data, extrapolated to the position of Lys 20 in RBD1, orient the two proteins bound to the UTR, and provide direct biochemical evidence for the proposed model of the RBD1:UTR complex.

Keywords: hydroxyl radical cleavage; RNA–protein complexes; U1A RBD1–EDTA(Fe); UTR

INTRODUCTION

RNA binding domains (RBDs) [or RNA recognition motifs (RRMs) or ribonucleoproteins (RNPs)] constitute a large family of RNA binding proteins (Birney et al., 1993; Burd & Dreyfuss, 1994). The domains are predicted to adopt $\beta_1\alpha_1\beta_2$-$\beta_3\alpha_2\beta_4$ secondary structures and an α/β sandwich global fold (Ghetti et al., 1989), in which the four-stranded antiparallel β-sheet packs against the two α-helices. Both β1 and β3 strands are conserved in RBDs, as are the amino acid residues that form the hydrophobic core of the domain, but the five loop sequences show no sequence conservation, and are variable in length (Birney et al., 1993).

The human U1A protein contains two RBDs, but only the N-terminal RBD1 is required to bind to its RNA targets (Scherly et al., 1989; Jessen et al., 1991; Hall & Stump, 1992). The structure of RBD1 has been solved by X-ray crystallography (Nagai et al., 1990) and NMR (Avis et al., 1996), and the cocrystal of RBD1 with its U1 snRNA hairpin target was the first illustration of how these domains bind RNA (Oubridge et al., 1994). The RNA sits on the β-sheet surface; residues in Loop 3 (between β2 and β3) make sequence-specific contacts with seven nucleotides (5′ AUUGCAC) of the RNA loop. Loop 1 of the protein contains four residues ($Asn_{18}Glu_{19}Lys_{20}Ile_{21}$); Lys 20 together with Lys 22 at the start of α1 have been suggested to electrostatically interact with the phosphate backbone of the RNA stem, localized to the 5′ strand of the duplex (Oubridge et al., 1994).

U1A RBD1 also binds to two asymmetric internal loops in the 3′ UTR of its own mRNA, which contain the sequence 5′ AUUGC/UAC (Box 1 and Box 2, see Fig. 1A); binding is required for autoregulation of polyadenylation (Boelens et al., 1993; van Gelder et al., 1993). Footprinting of RBD1 bound to the UTR sites showed that there is one protein bound to each adjacent internal loop (van Gelder et al., 1993; Teunissen et al., 1997). An NMR structure of RBD1 bound to a single UTR site equivalent to Box 2 shows the RNA loop nucleotides positioned on the β-sheet, with amino acids of Loop 3 again making many specific contacts (Allain et al., 1996, 1997; Gubser & Varani, 1996). The first base pair of the UTR stem 2 is in proximity to the side chain of Lys 23, which is in the first turn of α1, but no contacts are reported between the RBD and stem 3.

A model of two RBD1 molecules bound to the UTR was proposed by Jovine et al. (1996) based on the cocrystal structure and mapping data. In addition to interactions between the RBD1 Loop 3 and the surface of the β-sheet with the UTR RNA, this model positions

Reprint requests to: Kathleen B. Hall, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, USA; e-mail: hall@bionmr3.wustl.edu.

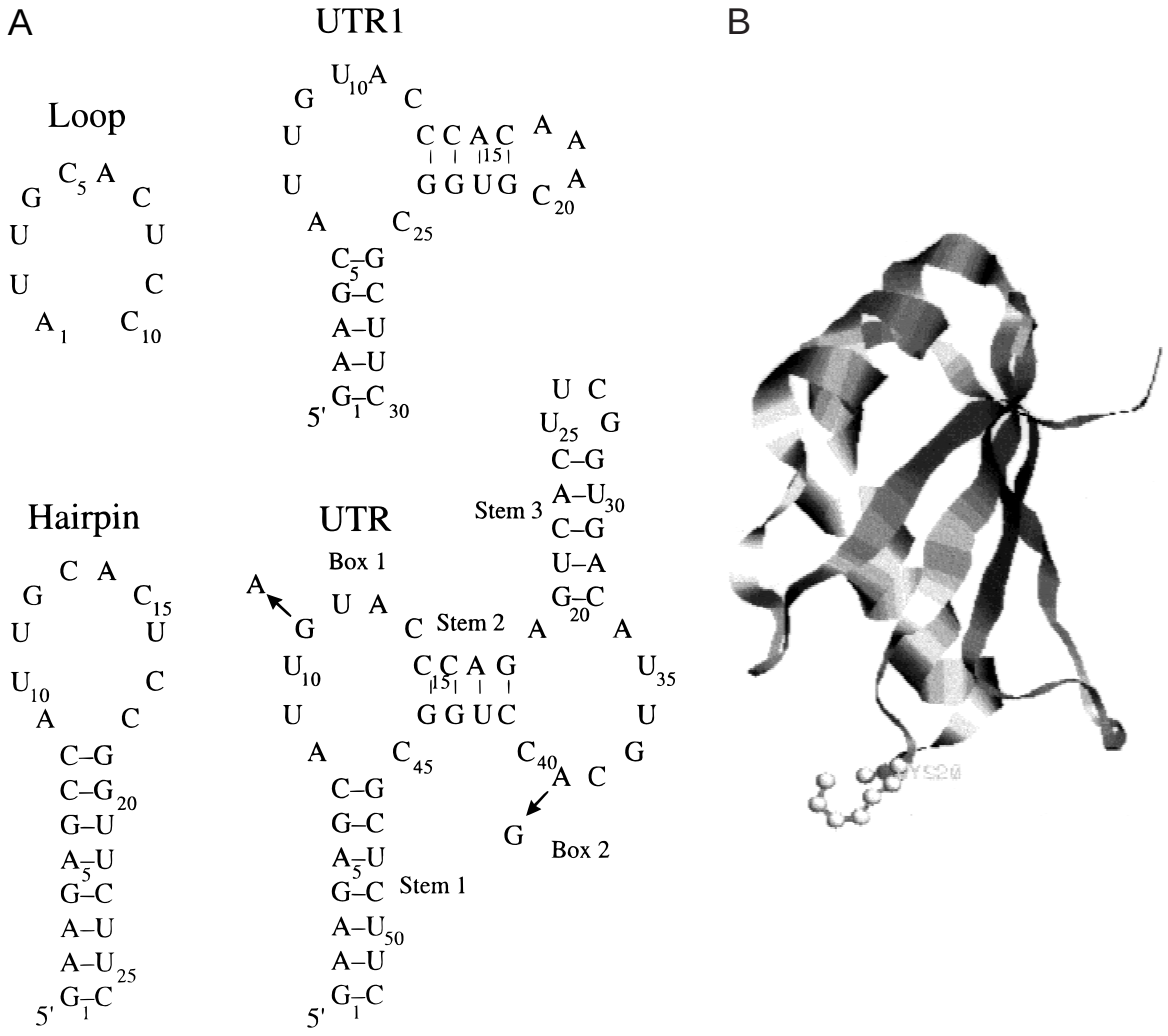

FIGURE 1. Sequence of the RNAs used in experiments. **A:** Numbering scheme of the hairpin loop, and the sequence of the hairpin used for binding experiments and cleavage experiments. UTR1 contains one binding site from the U1A mRNA 3′ UTR. 3′ UTR of U1A mRNA with the positions of mutations marked. **B:** Tertiary fold of RBD1 showing Lys 20 in Loop 1 (based on van Gelder et al., 1994). The RNA sits on the surface of the β-sheet.

RBD1 Loop 1 very close to UTR stems 1 (adjacent to Box 1) and 3 (adjacent to Box 2), based on Loop 1 contacts in the RBD1 cocrystal with the RNA hairpin. Details of the interactions of the two RBDs bound to the UTR relate to possible regulatory features, for there are data that suggest that protein binding to these two sites is cooperative (van Gelder et al., 1993) and also that occupancy of both sites is required for interactions with poly(A) polymerase and subsequent control of polyadenylation (Gunderson et al., 1994). Thus, to understand these processes, it is critical to have an accurate description of the interactions between RNA and protein, as well as between adjacent protein domains.

In these experiments, the human U1A RBD1 has been modified by site-directed mutagenesis to introduce a single cysteine in Loop 1, replacing Lys 20. It is important to note that RBD1 Loop 1 does not make sequence-specific contact with the RNA, so mutations should not change specificity. Introduction of K20C provides a single reactive site in RBD1 for (EDTA-2-aminoethyl) 2-pyridyldisulfide complexed to iron (EPD-Fe) (Ebright et al., 1992; Ermácora et al., 1992), which was attached covalently to the cysteine through its sulfhydryl group. Hydroxyl radicals are generated by the bound iron upon addition of ascorbate by the Fenton reaction (Imlay et al., 1988), which directs cleavage of RNA in its vicinity. EPD-Fe has a reported cleavage radius of 15 Å (Ermácora et al., 1994), which is roughly twice the length of a lysine side chain. This substitution should therefore approximate the position of Lys 20 in the RBD1:RNA complex. With the RBD1(K20C)-EPD-Fe protein bound to RNA, we have mapped the cleavage around the U1 snRNA hairpin and U1A mRNA UTR binding sites. These biochemical data orient the protein on the RNA in good agreement with the model of two RBD1 domains bound to its UTR (Jovine et al., 1996).

RESULTS

Based on the cocrystal structure of U1A RBD1 with the U1 snRNA stem/loop II (Oubridge et al., 1994), positions Glu 19 and Lys 20 (see Fig. 1B) were chosen for the introduction of a unique cysteine. The Glu 19 Cys protein was not soluble; the Lys 20 Cys was stable and soluble, so it was used for subsequent experiments. Nitrocellulose filter binding experiments (see Fig. 2) showed that the RBD1(K20C) protein bound the U1 snRNA stem loop II (hairpin) with a slightly weaker affinity [$K_D = 6.6\ (\pm\ 0.6) \times 10^{-10}$ M] than did wild-type RBD1 [$K_D = 9.5\ (\pm\ 0.7) \times 10^{-11}$ M]. The weaker affinity might reflect the loss of an electrostatic interaction between Lys 20 ϵ-NH_2 and the phosphate backbone of the RNA.

Attachment of EPD-Fe to RBD1(K20C)

After removing the DTT from the buffer, Ellman's reagent (DTNB) was added to the protein to determine the reactivity of the cysteine. Ellman's reagant has approximately the same reactivity toward cysteines as EPD (R. Fox, pers. comm.) and thus it serves as a reasonable measure of the expected degree of modification by EPD. Between 50 and 60% of the RBD1(K20C) was consistently modified by Ellman's reagent. This value is similar to those obtained for other proteins, where for each disulfide the equivalent thiol concentration detected by Ellman's reagent was 1.34, or 67% of the cysteines (Anderson & Wetlaufer, 1975). Because EPD-Fe absorbs strongly at 290 nm, it was not possible to determine the precise protein concentration once the protein had been modified with this reagent.

Binding of the RBD1(K20C)-EPD-Fe

To compare association of RBD1, RBD1(K20C), and RBD1(K20C)-EPD-Fe to RNA, complexes were observed using mobility shifts in nondenaturing gels. Experiments were done with each of the RNAs and were performed for each batch of RBD1(K20C)–EPD-Fe produced. As illustrated for RBD1 and RBD1(K20C) (Fig. 3), there are two bands in the samples of protein bound to RNA hairpin and to the UTR1 RNA, which contains one binding site equivalent to Box 1 of the UTR; the lower band corresponds to free RNA and the upper band to the bound complex. In the UTR:protein lanes, three bands are observed, corresponding to free RNA and RNA with one or both sites occupied (single-

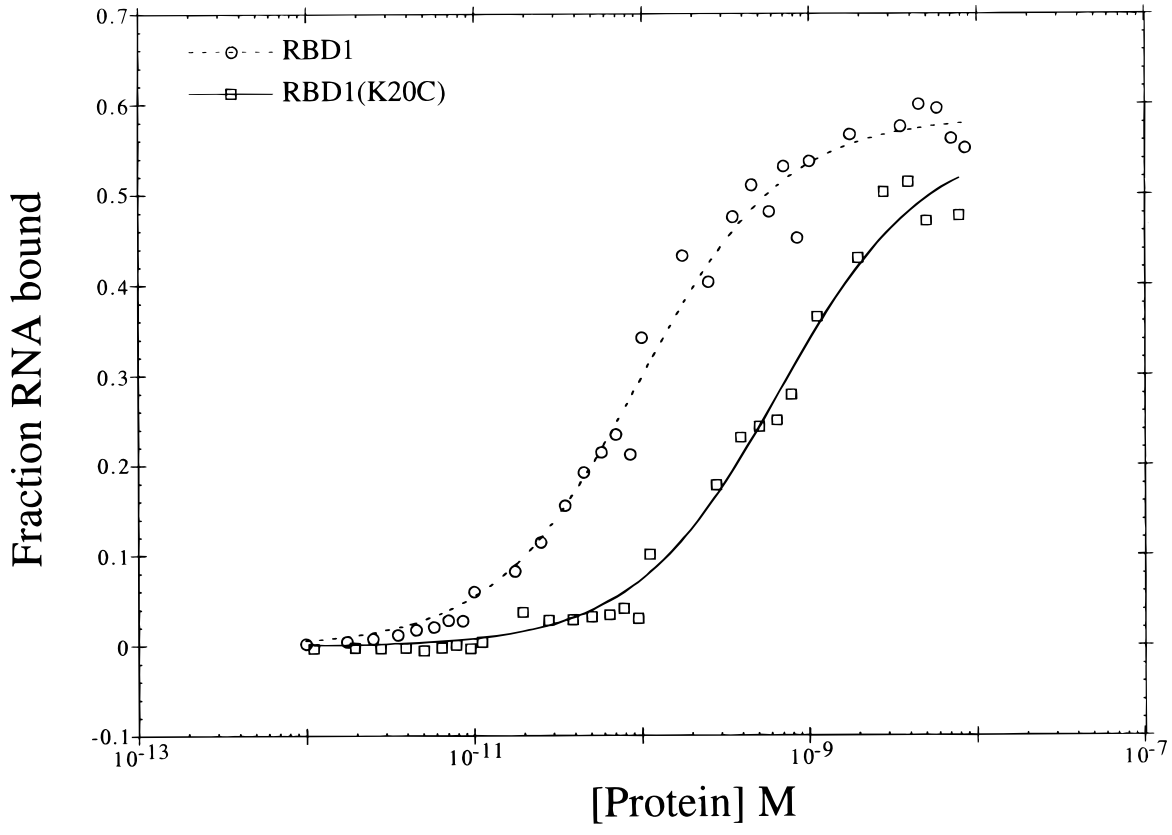

FIGURE 2. Binding isotherm of RBD1 and RBD1(K20C) proteins binding to the RNA hairpin. Data points are shown; smooth line is the fit to a Langmuir isotherm, assuming a two-state model and 1:1 stoichiometry. Binding assays include 150 mM NaCl, 1 mM $MgCl_2$, 10 mM sodium phosphate, pH 7.0, 1 mM DTT, 20 μg/mL BSA, and 50 μg/mL tRNA at 22 °C.

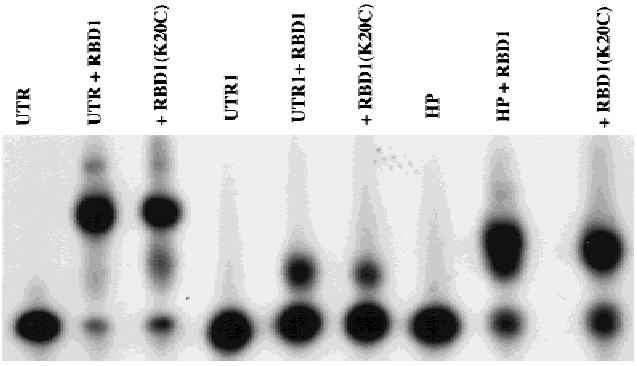

FIGURE 3. Gel shifts of UTR, UTR1, and hairpin (HP) by RBD1 and RBD1(K20C). Reactions contained 100 cps of end-labeled RNA per lane and either 3.6 μM RBD1 or 2.2 μM RBD1(K20C).

site occupancy is more apparent in the RBD1(K20C) sample). The band-shifts of RNA with RBD1(K20C)–EPD-Fe binding were virtually indistinguishable from those with RBD1(K20C) (data not shown). These band-shift experiments were also used to estimate how much RBD1(K20C)–EPD-Fe was required to occupy both sites on the UTR but to bind only a single site on the UTR mutants (UTR G11A and UTR A39G). Based on these results, cleavage reactions were performed to assign cleavage events to individual RBD1 binding sites.

Cleavage of U1 snRNA hairpin and UTR1

When RBD1(K20C)–EPD-Fe is used to cleave the wild-type hairpin (Fig. 1A), the cleavage centers at A5 in the stem (Fig. 4, lane 2). Significant cleavage is also seen at G4 and G6 and minor cleavage at positions A3 and C7. No cleavage is detectable on the complementary 3′ side of the stem between G20 and U24, which suggests that the 3′ side of the RNA stem is beyond the 10–15-Å reactive range of the C20-EPD-Fe tether. These data support the cocrystal structure of the U1A(95) and a slightly shorter hairpin molecule (Oubridge et al., 1994) in which K20 and K22 appear to interact with the phosphodiester backbone on the 5′ strand of the stem. Notably, when an RNA hairpin with a five-base pair stem was incubated with RBD1(K20C)–EPD-Fe, no cleavage was observed.

The UTR1 (30mer) (Fig. 1A) has only one binding site for RBD1, and is equivalent to Box 1. With this RNA, cleavage would be expected at the A2 position of the RNA stem, but no cleavage was detected with RBD1(K20C)–EPD-Fe (data not shown), even at very high concentrations of protein. As with the shorter RNA hairpin stem, it seems likely that the terminal base pair of the stem may be fraying and therefore unable to form a sufficiently stable interaction with RBD1 Loop 1 for efficient cleavage.

Cleavage of UTR RNAs

The UTR (52mer) (Fig. 1A) has two RBD1 binding sites, Box 1, flanked by stems 1 and 2, and Box 2, flanked by

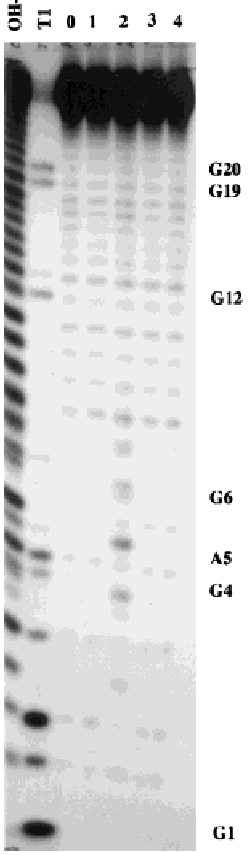

FIGURE 4. Cleavage of hairpin RNA by RBD1(K20C)–EPD-Fe. Lane 0, RNA with ascorbate; lane 1, with RBD1(K20C)–EPD-Fe; lane 2, with RBD1(K20C)–EPD-Fe and ascorbate; lane 3, with RBD1(K20C); lane 4, with RBD1(K20C) and ascorbate.

stems 2 and 3. Both sites are bound by RBD1(K20C)–EPD-Fe, as observed in gel-shift experiments. Two distinct cleavage sites were observed, centered around position G4 (cleavage at A2, A3, G4, A5) in stem 1 and around position U30 (cleavage at G28, G29, U30, G31, A32) in stem 3 (see Figs. 5A, 6). This is in agreement with the cleavage pattern observed for the hairpin RNA, for both stems are cleaved at an analogous site.

To confirm that cleavage in stem 1 was due to binding at the Box 1 binding site (see Fig. 1A), and that cleavage in stem 3 was due to binding at Box 2, the binding sites were individually mutated. It has been observed in RBD1 binding to hairpin RNAs that a G4 to A4, or an A6 to G6 mutation in the loop nucleotides (Fig. 1A) results in a decrease in the binding affinity by four orders of magnitude (Hall, 1994). In the UTR, these positions correspond to G11A and A13G in Box 1 and G37A and A39G in Box 2. Two individual UTR mutants were therefore constructed, UTR G11A and UTR A39G, which should allow single occupancy at the remaining wild-type binding site.

The cleavage pattern of UTR A39G, with Box 2 binding site mutated, was consistent with binding at Box 1 only, for cleavage was centered around position G4

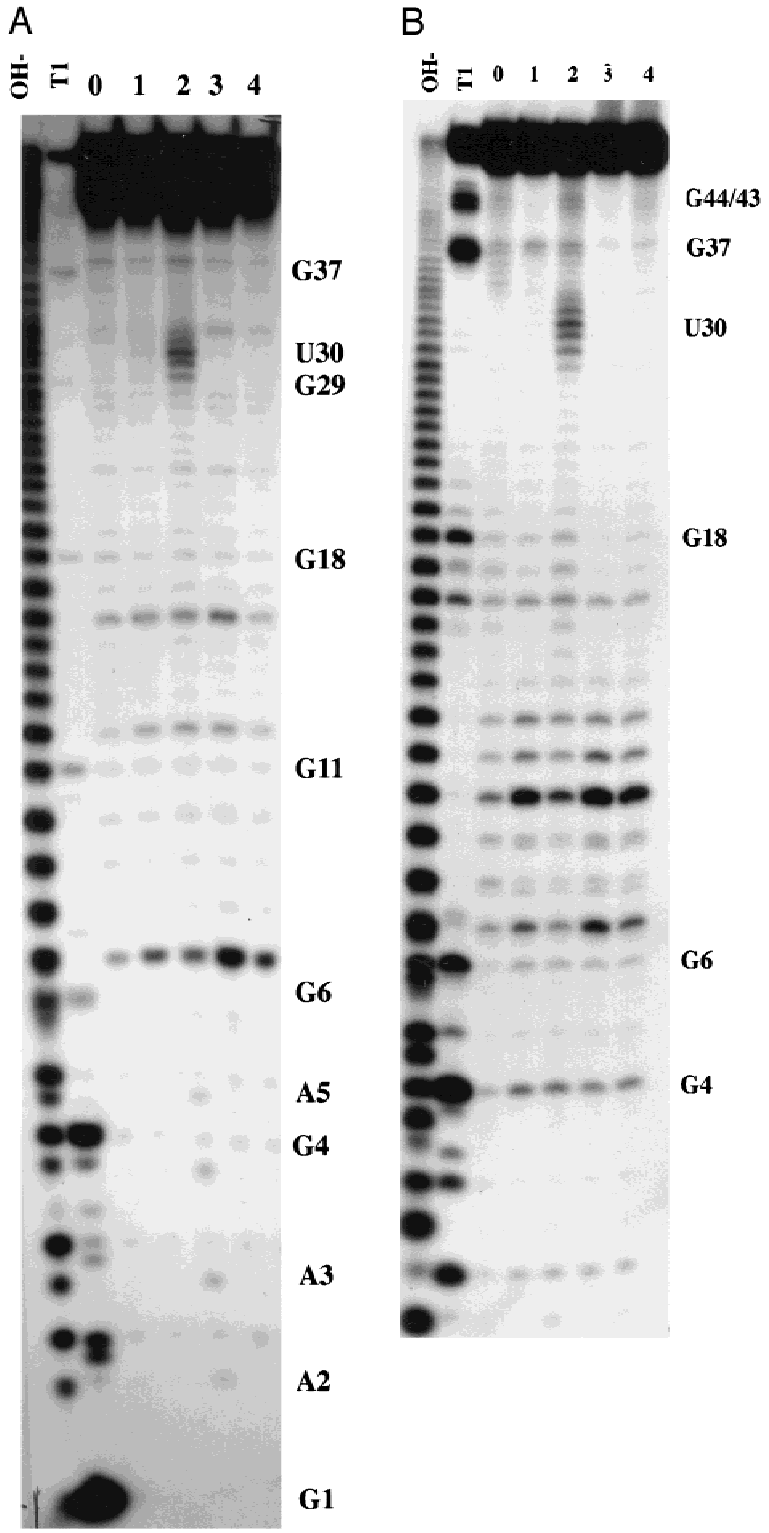

FIGURE 5. Cleavage of UTRs by RBD1(K20C)–EPD-Fe. **A:** WT UTR. Lane 0, with ascorbate; lane 1, with RBD1(K20C)–EPD-Fe; lane 2, with RBD1(K20C)–EPD-Fe and ascorbate; lane 3, with RBD1(K20C); lane 4, with RBD1(K20C) and ascorbate. **B:** UTR G11A. Lane 0, with ascorbate; lane 1, with RBD1(K20C)–EPD-Fe; lane 2, with RBD1(K20C)–EPD-Fe and ascorbate; lane 3, with RBD1(K20C); lane 4, with RBD1(K20C) and ascorbate.

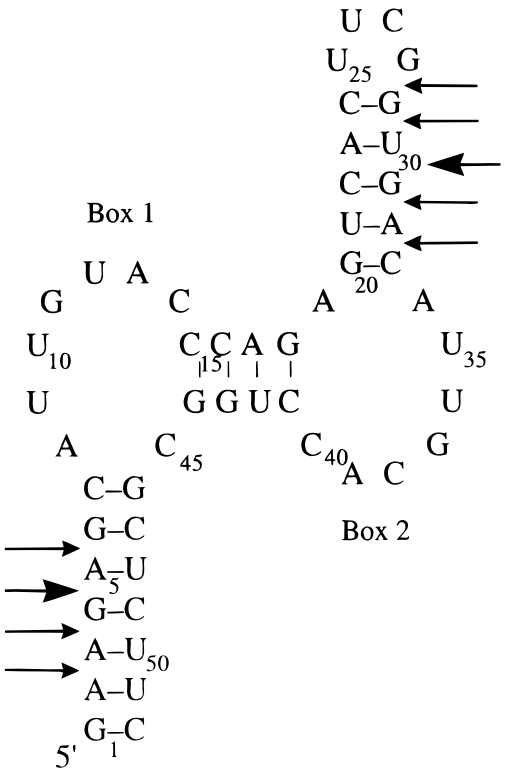

FIGURE 6. Cleavage sites (→) on the UTR around Box 1 and Box 2. Larger arrow is the most prominant cleavage. The cleavage pattern is independent of the number of binding sites occupied.

(cleavage at A2, A3, G4, A5, and slightly at G6) in stem 1 (data not shown) and no cleavage was detected around position U30 in stem 3. The cleavage pattern of UTR G11A (Fig. 5B) was consistent with binding at Box 2 only, for cleavage was centered around position U30 (cleavage at A32, G31, U30, G29, G28) in stem 3. Note that there is no cleavage observed around C22, A23, and C24 on the opposite strand of stem 3, which further constrains the relative geometry of the stem in the complex.

DISCUSSION

EDTA–metal complexes incorporated at selected sites within a protein have been used previously to study protein–protein interactions (Ermácora et al., 1992, 1994) and protein–DNA interactions (Ebright et al., 1992; Mazzarelli et al., 1993). More recently, directed hydroxyl radical cleavage has been used to probe the three-dimensional structure of RNA–protein complexes in the context of the 30S ribosome, where ribosomal proteins S4 and S13 containing tethered BABE-Fe were used to localize their binding on 16S rRNA (Heilek et al., 1995; Heilek & Noller, 1996). These experiments with the U1A RBD1 are used to demonstrate the applicability of the method to proteins in the RBD family, as well as to provide experimental evidence for a proposed structure of RBD1 with its 3′ UTR.

RBD1 and RNA binding

One result of the initial experiments with RBD1(K20C) binding to the RNA hairpin is the observation that it has lost about 1 kcal/mol of binding free energy, which thus implicates Lys 20 as one source of electrostatic contributions in the RNA–RBD1 complex. The EPD-Fe cleavage patterns indicate that this interaction undoubtedly involves Lys 20 and the phosphate backbone of the RNA stem, centered three or four base pairs distant from the base of the RNA loop. The cleavage patterns provide the first direct biochemical evidence that RBD1 Loop 1 residues are juxtaposed to specific positions on the RNA duplex adjacent to the binding site. The cocrystal structure of U1A RBD1 and U1 snRNA hairpin suggested that Lys 20 could interact with the phosphodiester backbone on the 5′ strand of the hairpin stem (Oubridge et al., 1994), although the RNA residues involved in the interaction were not identified. In biochemical studies of the interaction of U1 snRNA and RBD1 (Jessen et al., 1991), protection from ethylation was seen in residues corresponding to positions G1–G6 of the hairpin, with strongest protection at positions A3 and G4, suggesting that this region of the phosphodiester backbone is interacting with the protein. RBD1(K20C)–EPD-Fe cleavage is centered at A5, which is consistent with these data.

These EPD-Fe cleavage results are in agreement with the chemical and nuclease footprinting data of Teunissen et al. (1997), which mapped the binding of U1A on the Ag RNA (the complete 3′-UTR including the polyadenylation signal). In those experiments, the accessibility of the Ag RNA to nucleases was compared in the presence or absence of excess U1A. The U1A:Ag RNA complex reduced cleavage of nucleotides corresponding to positions 28–32 on the 5′ side of stem 3 adjacent to Box 2, whereas some protection was noted at nucleotides corresponding to the 3′ side of stem 1 adjacent to Box 1; no protection was observed at positions 2–6 on the 5′ side of stem 1. The RBD1(K20C)–EPD-Fe cleavage of the UTR G11A most clearly illustrates that cleavage occurs only along the 5′ side of the duplex adjacent to the internal loop, thus positioning this strand of the duplex next to the body of the RBD. The presence of the the downstream polyadenylation element in the Ag RNA may alter the stem 1 conformation in the complex, which could explain the apparent differences observed.

Jovine et al. (1996) proposed a model of the interaction of RBD1 with the 3′-UTR of the U1A mRNA based on the cocrystal of RBD1 and U1 snRNA hairpin. In this model, RBD1 binding to Box 1 positions the side chains of Lys 20 and Lys 22 for interaction with the phosphate groups of residues G4–C7 in stem 1, and

for Box 2 binding, residues G29–A32 in stem 3. Their model places Lys 20 near A5 and G31 (using our UTR numbering scheme), leading to the prediction that RBD1(K20C)–EPD-Fe would center its cleavage there. The observed cleavage instead centers at G4 and U30, where the model predicts Lys 22 to be located, but, because the reactive radius of the EPD-Fe is longer than a single lysine side chain, the observed cleavage may extend beyond the reach of Lys 20 in the complex. The model was based on the RBD1–hairpin cocrystal structure; NMR studies of the interaction of RBD1 with Box 2 of the UTR did not report any contacts between Lys 20 and the RNA. However, those NMR data indicate that the side chain of Lys 23, which is in α1, is positioned near stem 2 (Allain et al., 1996, 1997; Gubser & Varani, 1996). Our biochemical results, when combined with the NMR data, show that the two RNA stems flanking the internal loop are both used by the protein as sources of electrostatic interactions, and possibly function also to position the RNA on the β-sheet surface.

Despite some uncertainty in the position of Lys 20 on the RNA stem, these RBD1–EPD data do allow further biochemical and structural constraints to be put on the complexes: specifically, whether one site or both sites of the UTR are occupied, the cleavage patterns at each site are identical (Fig. 6), indicating that the presence of two adjacent RBDs does not alter the interactions between RNA and protein (at least at this level of resolution). Also, cleavage is centered at the fourth base pair from the base of the loop, whether in the UTR or hairpin binding site. Thus, there does not appear to be any (detectable) difference in the interaction of RBD1 Loop 1 with the RNA hairpin stem or the comparable UTR stem.

Applications to other RBDs

In U1A RBD1, amino acid side chains from Loop 1 interact with the bound RNA, where they contribute nonspecific electrostatic binding free energy. Given the juxtaposition of this loop to bound RNA, it makes a convenient site for introduction of a cleavage agent that will not interfere with specific recognition. Assuming that all RBDs will bind RNAs in a similar geometry, their Loop 1 sequences could be modified for introduction of a cleavage reagent. However, Loop 1 sequences are not conserved in RBDs, and it is likely that in some proteins this loop will be essential for protein folding and stability (Glu 19 in RBD1 could not be replaced by Cys, for example), making the actual position subject to experiment. Of course, other sites on the RBD, such as C-terminal sequences, Loop 3 sequences, and even a residue on the surface of the β-sheet, might be candidates for introduction of a reactive site. This example of site-directed RNA cleavage suggests a method for mapping RNA binding sites of proteins that contain multiple RBDs [e.g., U2AF, poly(A) binding protein, sex-lethal] to understand how the RNA is distributed among the multiple β-sheet surfaces. Because EDTA-Fe will also cleave proteins (Ermácora et al., 1994, 1996), it could be attached to RBDs to map the relative orientations of proteins in RNA–protein complexes.

MATERIALS AND METHODS

Construction of the cysteine mutant

A cysteine was introduced into RBD1 of the U1A protein (Sillekens et al., 1987) by site-directed mutagenesis using recombinant PCR. A K20C mutant was constructed using the following primers: N-terminus, TAATACGACTCACTATAG; C-terminus, GGCCAAGCTTCTTCTACTAGAAGGTGCCTTTCATC; K20C (top), AACCTCAATGAGTGTATCAAGAAGGATG; and K20C (bottom), CATCCTTCTTGATACACTCATTGAGGTT. The DNA was cloned into a plasmid under the control of the tac promoter (Hall & Stump, 1992) and sequenced.

Protein purification

The protein was isolated as described (Hall & Stump, 1992) with differences noted below. The protein was overexpressed in *Escherichia coli* and the cells were pelleted. The pellet was resuspended in 50 mM PIPES, pH 6.5, 50 mM NaCl, 100 mM DTT, 20 μg/mL PMSF, 1 μg/mL pepstatin, 2 μg/mL leupeptin, and DNAse II. The cells were French pressed, followed by ammonium sulfate precipitation, and resuspended in 50 mM PIPES, pH 6.5, 50 mM NaCl, 10 mM DTT, 20 μg/mL PMSF. The protein was further purified by cation exchange chromatography, followed by ammonium sulfate precipitation and extensive dialysis against a solution of 50 mM PIPES, pH 6.5, 50 mM NaCl, and 10% glycerol that had been treated with Chelex 200 (Bio-Rad) to remove all divalent cations. The buffer was supplemented with 10 mM DTT and 20 μg/mL PMSF, and the dialyzed protein was stored at 4 °C. The purity of the protein as determined by silver stain was greater than 98%. Protein concentration was determined spectroscopically, using $\epsilon_{280} = 5{,}120\ M^{-1}\ cm^{-1}$, based on its four tyrosines [ϵ_{280} (tyr) $= 1{,}280\ M^{-1}\ cm^{-1}$] (Edelhoch, 1967).

RNA synthesis

RNA molecules were synthesized using SP6 RNA polymerase from DNA oligonucleotides as described (Stump & Hall, 1993). The sequences of the four RNAs used here are shown in Figure 1. The UTR RNAs differ from the wild-type U1A 3′-UTR in that they do not contain the polyadenylation sequence of the full-length 3′-UTR. In addition, the normal 3′-UTR stem 1 is shorter by two base pairs and position G4–C48 is C-G; this sequence was changed here to transcribe with SP6 RNA polymerase. Positions 24–30 (UGUCCCCA) were changed to ACUUCGGU to introduce a UUCG tetraloop to stabilize the structure. Before constructing the mutations in the UTR sequence, the RNAs were folded using *m-fold* (Zuker, 1989) to compare their predicted secondary structures with that of the wild-type UTR sequence. The predicted lowest-energy structures for the mutants and wild-type were identical.

RNAs were transcribed with either α-^{32}P-CTP or α-^{32}P-UTP for binding experiments, or 5′-labeled with γ-^{32}P-ATP by polynucleotide kinase after treatment with calf intestine alkaline phosphatase for cleavage reactions. All final RNA products were purified from polyacrylamide urea gels.

Filter binding assays

Nitrocellulose filter binding assays were performed as described (Hall & Kranz, 1995) in 150 mM NaCl, 1 mM $MgCl_2$, 10 mM sodium phosphate, pH 7.0, 1 mM DTT, 20 μg/mL BSA, and 50 μg/mL tRNA at 22 °C, using Schleicher and Schuell 0.2-μm BA-S83 supported nitrocellulose membrane filters. A limiting concentration of ^{32}P-RNA was titrated with RBD1(K20C) and incubated 20 min before filtering with a modified dot blot apparatus (Wong & Lohman, 1993). Bound radiolabeled RNA was quantified using a Storm 840 phosphorimager (Molecular Dynamics). The fraction of RNA bound was calculated as $(R_B - R_0)/R_T$, where R_B is bound RNA, R_0 is the RNA retained in the absence of protein, and R_T is the total RNA added in each reaction. All binding experiments were repeated at least twice.

Protein–EDTA-Fe conjugation

An aliquot of RBD1(K20C) was purified away from DTT by spin-column chromatography (Penfesky, 1979) using Bio-Gel P2 (100–200 mesh, Bio-Rad) that had been washed extensively with Chelex-treated 50 mM PIPES, pH 6.5, 100 mM NaCl. In order to monitor the protein flowthrough and separation from DTT (MW 154), a duplicate column was loaded with blue dextran (MW 8×10^6) and orange G (MW 452). The protein concentration was determined as described above. The reactivity of the cysteine in the RBD1(K20C) protein was determined using Ellman's reagant and the assay was performed as described (Ellman, 1959). Twenty microliters of 4 mg/mL Ellman's reagent in 0.1 M Na_2HPO_4, pH 8.0, was added to 20 μL of protein in 960 μL 50 mM PIPES and 50 mM NaCl. The absorbance at 412 nm was read after 15 min, and the concentration of reactive sulfhydryl groups was determined spectroscopically, using $\epsilon_{412} = 1.36 \times 10^4$ M^{-1} cm^{-1}. The ratio of reacted sulfhydryl groups to protein concentration (mol reacted)/(mol total) was consistently between 50 and 60%. RBD1(K20C) was reacted with a twofold molar excess of [3 mM] EPD-Fe (Ermácora et al., 1992, 1994, 1996; Platis et al., 1993) in the presence of chelexed 50 mM PIPES, pH 7.0, 100 mM NaCl for 1 h at 22 °C. Modified protein was purified from free EPD-Fe using spin column chromatography as described above. To the RBD1(K20C)–EPD-Fe solution was added an equal volume of 100% glycerol that had been filtered through a 0.2-μm nitrocellulose centrifugal microfilter (Schleicher & Schuell Centrex MF). The final concentration of RBD1(K20C)–EPD-Fe could not be determined accurately because EPD-Fe absorbs strongly at 290 nm (Ermácora et al., 1992). Modified protein was stored at −20 °C and used within 2–3 weeks. Slight protein degradation has been observed if modified proteins are stored for an extended period of time (R. Fox, pers. comm.).

Gel shift assays

Binding of the RNAs by RBD1, RBD1(K20C), and RBD1(K20C)–EPD-Fe were compared by examining the shift in mobility of radiolabeled RNA using native gel electrophoresis (8% acrylamide, 40:1 bis, 50 mM Tris-glycine, pH 8.0). Gels were prerun for at least 30 min at 7 V/cm at 4 °C. Radiolabeled RNA and protein were incubated in 50 mM PIPES, pH 7.0, 100 mM NaCl with 100 μg/mL tRNA; glycerol was then added to make 20% glycerol for a final sample volume of 10 μL. Gels were run for 5 h at 4 °C and the bands were visualized by autoradiography. Various concentrations of protein were used to titrate the two binding sites on the UTR constructs. Gel shifts were repeated for each batch of RBD1(K20C)–EPD-Fe.

Cleavage reactions

All cleavage reactions (final volume 10 μL) were performed in chelexed 50 mM PIPES, pH 6.5, 100 mM NaCl, using 5′-labeled ^{32}P-RNA. An appropriate amount of modified protein, as determined from the gel shift assays, was added to the RNA (approximately 100 cps). Two microliters of freshly prepared 300 mM sodium ascorbate in chelexed 50 mM PIPES, 100 mM NaCl, pH 7.0, that had been filtered through a 0.2-μm nitrocellulose centrifugal microfilter were added and the reactions incubated for 50 min at 37 °C.

For control reactions, RNA was incubated without protein in the presence of ascorbate, RBD1(K20C)–EPD-Fe in the absence of ascorbate, or with RBD1(K20C) in the presence and absence of ascorbate. A hydroxyl ladder was generated by incubating RNA with 0.1 M $NaHCO_3$, pH 9.8, for 10 min at 95 °C in the presence of 11 μg tRNA. A T1 digest was generated by incubating RNA with 0.1 units T1 endonuclease (Boehringer) for 10 min at 65 °C in the presence of 20 μg tRNA.

All reactions were terminated by phenol extraction followed by ethanol precipitation in the presence of 10 μg glycogen as carrier. The samples were resuspended in 16 μL loading buffer, and 8 μL were loaded on a 20% (20:1) polyacrylamide urea sequencing gel. Gels were visualized via autoradiography. All RNAs were cleaved with at least two different preparations of RBD1(K20C)–EPD-Fe.

ACKNOWLEDGMENTS

EPD was a generous gift from Prof. Robert Fox (UTMB, Galveston, Texas). We thank Jim Kranz for making the RNA hairpin. D.L.B. was supported by NIH training grant GM07067-22 and is supported as a Lucille P. Markey Pathway predoctoral fellow. This work was supported by the NIH (GM46318) to K.B.H.

Received November 24, 1997; returned for revision December 11, 1997; revised manuscript received December 18, 1997

REFERENCES

Allain FHT, Gubser C, Howe PWA, Nagai K, Neuhaus D, Varani G. 1996. Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. *Nature 380*:646–650.

Allain FHT, Howe PWA, Neuhaus D, Varani G. 1997. Structural basis of the RNA-binding specificity of the human U1A protein. *EMBO J 16*:5764–5774.

Anderson WL, Wetlaufer DB. 1975. A new method for disulfide analysis of peptides. *Anal Biochem 67*:493–502.
Avis JM, Allain FHT, Howe PWA, Varani G, Nagai K, Neuhaus D. 1996. Solution structure of the N-terminal RNP domain of U1A protein: The role of C-terminal residues in structure stability and RNA binding. *J Mol Biol 257*:398–411.
Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res 21*:5803–5816.
Boelens WC, Jansen EJ, van Venrooij WJ, Stripecke R, Mattaj IW, Gunderson SI. 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell 72*:881–892.
Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science 265*:615–621.
Ebright YW, Chen Y, Pendergast PS, Ebright RH. 1992. Incorporation of an EDTA–metal complex at a rationally selected site within a protein: Application to EDTA–iron DNA affinity cleaving with catabolite gene activator protein (CAP) and Cro. *Biochemistry 31*:10664–10670.
Edelhoch H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*:1948–1954.
Ellman GL. 1959. Tissue sulfhydryl groups. *Arch Biochem Biophys 82*:70–77.
Ermácora MR, Delfino JM, Cuenoud B, Schepartz A, Fox RO. 1992. Conformation-dependent cleavage of staphylococcal nuclease with a disulfide-linked iron chelate. *Proc Natl Acad Sci USA 89*:6383–6387.
Ermácora MR, Ledman DW, Fox RO. 1996. Mapping the structure of a non-native state of staphylococcal nuclease. *Nature Struct Biol 3*:59–66.
Ermácora MR, Ledman DW, Hellinga HW, Hsu GW, Fox RO. 1994. Mapping staphylococcal nuclease conformation using EDTA-Fe derivative attached to genetically engineered cysteine residues. *Biochemistry 33*:13625–13641.
Ghetti A, Padovani C, Di Cesare G, Morandi C. 1989. Secondary structure prediction for RNA binding domain in RNP proteins identifies $\beta\alpha\beta$ as the main structural motif. *FEBS Lett 257*:373–376.
Gubser CC, Varani G. 1996. Structure of the polyadenylation regulatory element of the human U1A pre-mRNA 3′-untranslated region and interaction with the U1A protein. *Biochemistry 35*:2253–2267.
Gunderson SI, Beyer K, Martin G, Keller W, Boelens WC, Mattaj LW. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell 76*:531–541.
Hall KB. 1994. Interactions of RNA hairpins with the human U1A N-terminal RNA binding domain. *Biochemistry 33*:10076–10088.
Hall KB, Kranz JK. 1995. Thermodynamics and mutations in RNA–protein interactions. *Methods Enzymol 259*:261–281.
Hall KB, Stump WT. 1992. Interaction of N-terminal domain of U1A protein with an RNA stem/loop. *Nucleic Acid Res 20*:4283–4290.
Heilek GM, Marusak R, Meares CF, Noller HF. 1995. Directed hydroxyl radical probing of 16S rRNA using Fe(II) tethered to ribosomal protein S4. *Proc Natl Acad Sci USA 92*:1113–1116.
Heilek GM, Noller HF. 1996. Directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S13 using tethered Fe(II). *RNA 2*:597–602.
Imlay JA, Chin SM, Linn S. 1988. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. *Science 240*:640–642.
Jessen TH, Oubridge C, Teo CH, Pritchard C, Nagai K. 1991. Identification of molecular contacts between the U1A small nuclear ribonucleoprotein and U1 RNA. *EMBO J 10*:3447–3456.
Jovine L, Oubridge C, Avis JM, Nagai K. 1996. Two structurally different RNA molecules are bound by the spliceosomal protein U1A using the same recognition strategy. *Structure 4*:621–631.
Mazzarelli JM, Ermácora MR, Fox RO, Grindley NDF. 1993. Mapping interactions between the catalytic domain of resolvase and its DNA substrate using cysteine-coupled EDTA–iron. *Biochemistry 32*:2979–2986.
Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal-structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein-A. *Nature 348*:515–520.
Oubridge C, Ito H, Evans PR, Teo CH, Nagai K. 1994. Crystal-structure at 1.92 angstrom resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature 372*:432–438.
Penfesky HS. 1979. A centrifuged-column procedure for the measurement of ligand binding by beef heart F_1. *Methods Enzymol 56*:527–530.
Platis IE, Ermácora MR, Fox RO. 1993. Oxidative polypeptide cleavage mediated by EDTA-Fe covalently linked to cysteine residues. *Biochemistry 32*:12761–12767.
Scherly D, Boelens W, van Venrooij WJ, Dathan NA, Hamm J, Mattaj I. 1989. Identification of RNA binding segment of human U1A protein and definition of its binding site on U1 snRNA. *EMBO J 8*:4163–4170.
Sillekens PTG, Habets WJ, Beijer RP, van Venrooij WJ. 1987. cDNA cloning of the human U1 snRNA-associated A protein: Extensive homology between U1 and U2 snRNP-specific proteins. *EMBO J 6*:3841–3848.
Stump WT, Hall KB. 1993. SP6 RNA polymerase efficiently synthesizes RNA from short double-stranded DNA templates. *Nucleic Acids Res 21*:5480–5484.
Teunissen SWM, van Gelder CWG, van Venrooij WJ. 1997. Probing the 3′ UTR structure of U1A mRNA and footprinting analysis of its complex with U1A protein. *Biochemistry 36*:1782–1789.
van Gelder CWG, Gunderson SI, Jansen EJR, Boelens WC, Polycarpou-Schwartz M, Mattaj IW, van Venrooij WJ. 1993. A complex secondary structure in U1A pre-mRNA that binds two molecules of U1A is required for regulation and polyadenylation. *EMBO J 12*:5191–5200.
van Gelder CW, Leusen FJ, Leunissen JA, Noordik JH. 1994. A molecular dynamics approach for the generation of complete protein structures from limited coordinate data. *Proteins Struct Funct Genet 18*:174–185.
Wong I, Lohman TL. 1993. A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interactions. *Proc Natl Acad Sci USA 90*:5428–5432.
Zuker M. 1989. On finding all suboptimal foldings of an RNA molecule. *Science 244*:48–52.