4688–4695 *Nucleic Acids Research, 1998, Vol. 26, No. 20*

Mapping of a protein–RNA kissing hairpin interface: Rom and Tar–Tar*

Luis R. Comolli, Jeffrey G. Pelton and Ignacio Tinoco Jr*

Department of Chemistry, University of California at Berkeley and Structural Biology Department, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720-1460, USA

Received June 12, 1998; Revised and Accepted August 18, 1998

ABSTRACT

An RNA 'kissing' complex is formed by the association of two hairpins via base pairing of their complementary loops. This sense–antisense RNA motif is used in the regulation of many cellular processes, including *Escherichia coli* ColE1 plasmid copy number. The RNA one modulator protein (Rom) acts as a co-regulator of ColE1 plasmid copy number by binding to the kissing hairpins and stabilizing their interaction. We have used heteronuclear two-dimensional NMR spectroscopy to map the interface between Rom and a kissing complex formed by the loop of the *trans*-activation response (Tar) element of immunodeficiency virus 1 (HIV-1) and its complement. The protein binding interface was obtained from changes in amide proton signals of uniformly ^{15}N-labeled Rom with increasing concentrations of unlabeled Tar–Tar*. Similarly, the RNA-binding interface was obtained from changes in imino proton signals of uniformly ^{15}N-labeled Tar with increasing concentrations of unlabeled Rom. Our results are in agreement with previous mutagenesis studies and provide additional information on Rom residues involved in RNA binding. The kissing hairpin interface with Rom leads to a model in which the protein contacts the minor groove of the loop–loop helix and, to a lesser extent, the major groove of the stems.

INTRODUCTION

Antisense regulation mediated by RNA loop–loop interactions occurs in many cellular processes in both prokaryotes and eukaryotes (1,2). The antisense regulation of *Escherichia coli* ColE1 plasmid copy number has been studied in great detail (for reviews see 3–6 and references therein). RNA-I and RNA-II are two RNA transcripts encoded next to the ColE1 plasmid replication origin. RNA-II hybridizes to the plasmid DNA and is processed by RNase H to form the primer for DNA synthesis. RNA-I acts as an antisense inhibitor of RNA-II hybridization and primer formation. RNA-I and RNA-II are complementary because they are transcribed in opposite directions from the same DNA sequence. Initial pairing of the RNA transcripts occurs via loop–loop or kissing hairpin interactions between three tandem stem–loop structures in each molecule. The RNA one modulator protein (Rom) acts as a co-regulator by binding to the kissing hairpins and decreasing their dissociation rate (3). Ultimately, the metastable complex folds into a stable RNA-I–RNA-II duplex that is not cleaved by RNase H.

The *trans*-activation response element (Tar) of human immunodeficiency virus 1 (HIV-1) consists of a 59 nt hairpin closed by a 6 nt loop (7–10). The loop binds cellular proteins and is required for HIV replication (11,12). A conserved 14 nt sequence in the HIV *gag* gene can form a 4 bp stem with a 6 nt loop complementary to the Tar loop (13). Tandem repeats of Tar and antisense Tar hairpins have been tested for inhibition of HIV replication in cell extracts (14). Recently, Chang and Tinoco showed that a hairpin, denoted Tar*, with a loop complementary to that of Tar associates specifically with it to form a stable kissing hairpin (13). The structure of the Tar–Tar* complex (Fig. 1) was determined by NMR spectroscopy (15). A kissing complex from the ColE1 antisense control region has also been determined (16). The structures have many features in common. All loop nucleotides form base pairs with their complements, as determined from imino spectra and AH2 cross-strand connectivities in NOESY spectra. The loop bases stack on the 3′-side of each stem forming a quasi-continuous helix. Each loop is closed by bridging of the phosphate backbone across the major groove. Bridging occurs between U5 and U6 and between C5 and C6 for Tar and Tar* respectively. The crossover compresses the major groove and introduces a bend in the helix (15,16). It also brings three pairs of phosphates into close proximity, providing a potential strong binding site for magnesium ions (15).

Although many RNA-binding motifs have been identified (17–20), there are relatively few reports of protein–RNA complexes. Important examples are the NMR (21,22) and X-ray structures (23) of U1A–RNA complexes and X-ray structures of tRNA–tRNA synthetase complexes (24). The structure of Rom has been determined by both NMR in pH 2.3 phosphate buffer (25) and by X-ray crystallography (26); the structures are similar. Rom is a homodimer in which the two monomers are related by a dyad axis of symmetry. The monomer (63 amino acids) consists of two α-helices, H1 (the N-terminal domain) and H2 (the C-terminal domain), connected by a hairpin bend. Each monomer packs with its antiparallel neighbor to form a four helix, coiled

*To whom correspondence should be addressed at: Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720-1460, USA. Tel: +1 510 642 3038; Fax: +1 510 643 6232; Email: intinoco@lbl.gov

coil bundle based on a classic leucine zipper motif. Rom cannot be easily related to other RNA-binding motifs nor has the structure of a protein–kissing hairpin complex been determined.

Rom binding to the wild-type ColE1 kissing hairpin and variants of it has been probed by biochemical methods. Elegant ribonuclease cleavage studies have shown that Rom binds to the loop residues of the kissing complex in a pseudosymmetrical fashion (27). Affinity constants derived from cleavage rates indicate that Rom can bind to 6, 7 or 8 nt loop complexes with similar affinity and recognizes a specific structure rather than a particular sequence (28). Alanine scanning mutagenesis experiments on Rom binding to the wild-type ColE1 complex and two variants have implicated residues in helix H1 and its symmetrical counterpart, H1′. Asn10, Phe14, Gln18 and Lys25 abolish binding when substituted by Ala (29). Lys3 may also play a role. Additional mutational experiments suggest that Phe14 residues in helices H1 and H1′ interact with loop residues in a pseudosymmetrical fashion (29,30). These studies have also shown that helix H2 and its symmetry related helix H2′ do not participate in binding.

Here we report NMR studies of Rom with an RNA kissing complex. Tar–Tar* (Fig. 1) was chosen because of the interest in antisense HIV strategies involving Tar described above (14) and because it is the only kissing complex whose structure has been determined at high resolution (RMSD of ~0.8 Å) (15). The protein–RNA binding interface is inferred from changes in amide and imino proton signals revealed by 2D ^{15}N-^{1}H HSQC spectra. The results are related to previous biochemical data.

MATERIALS AND METHODS

RNA synthesis and purification

All RNA molecules were synthesized *in vitro* using T7 RNA polymerase and a synthetic DNA template (31). Uniformly ^{15}N-labeled Tar was produced from ^{15}N-labeled nucleotide monophosphates (NMPs) isolated from *E. coli* BL21 cells grown on ^{15}N-labeled ammonium chloride (32,33). The NMPs were converted to triphosphates using enzymatic phosphorylation (34). The RNA species were purified using 20% denaturing polyacrylamide gel electrophoresis. RNA extinction coefficients at wavelengths of 260 and 280 nm were calculated using nearest neighbors parameters. The values are 156 and 84/mM/cm at 260 and 280 nm, respectively, for Tar and 154 and 72/mM/cm at 260 and 280 nm, respectively, for Tar*. The Tar–Tar* complex was formed by annealing each hairpin separately prior to combining them together in a 1:1 ratio. The concentration of the Tar–Tar* kissing complex was further confirmed using the sum of extinction coefficients for the individual hairpins as a first approximation to its extinction coefficient. Hypochromicity due to the loop–loop interaction was determined to affect the measurements by <2%. RNA samples were dialyzed into 10 mM sodium phosphate, pH 6.3, and 0.1 mM EDTA.

Protein synthesis and purification

Escherichia coli strain BL21(DE3) transformed with the plasmid p2R that encodes Rom was a kind gift of Professor L.Regan (Yale University). This construct differs from wild-type Rom by a Gly insertion after the N-terminal Met (29). Expression and purification were done as described (29) except for a final reverse phase HPLC step to remove residual RNase activity. The Rom sequence was confirmed and the purity of each preparation was checked by electrospray ionization mass spectrometry. Two species which differed by the presence or absence of an N-terminal Met were found. The purity of each Rom sample was >95%, with the ratios of the two species dependent on sample preparation. Purified Rom was exchanged into a solution containing 10 mM sodium phosphate, 10 mM sodium chloride, 1 mM dithiothreitol (DTT) and 0.1 mM EDTA by ultracentrifugation through a 3 kDa molecular weight cut-off membrane. Samples at different salt concentration were obtained in a similar way, keeping everything else constant. Rom protein concentrations were determined by UV spectroscopy using an extinction coefficient of 0.24 cm^2/mg at 280 nm. Sample concentrations ranged from 1 to 5.5 mM/mol homodimer.

UV melting

Optical absorbance melting studies were performed on a Gilford 250 spectrometer using a heating rate of 0.5°C/min. The RNA was dissolved in 10 mM sodium phosphate buffer, pH 6.3. Studies of the sodium and magnesium dependence of the loop–loop interaction were done using an RNA concentration of 10 μM in the same buffer and varying salt concentrations. Sodium chloride concentrations ranged from 0 to 500 mM and magnesium chloride concentrations ranged from 0 to 5 mM. In the sodium chloride studies the buffer included 0.1 mM EDTA. Optical absorbance melting studies of Rom binding to the kissing complex were studied using RNA and protein homodimer concentrations of 10 μM, combined in a 1:1 ratio in the same buffer, and various salt concentrations.

Protein–RNA titrations

Aliquots of Tar–Tar* were prepared by lyophilizing the appropriate volume of stock solution containing the RNA complex in 10 mM sodium phosphate, pH 6.3. Titration of Rom with Tar–Tar* was done by dissolving each aliquot of dry RNA with the protein solution in the buffer previously described. In this way the volume and sodium chloride or magnesium chloride and DTT concentrations of the NMR sample remained constant throughout the experiment while the concentration of sodium phosphate varied from 10 to 30 mM. Control experiments show that variations in sodium phosphate concentration within this range produce no changes in the spectra. Formation of the kissing hairpin complex was corroborated by comparison of 1D imino spectra with published data (13). Similarly, titration of Tar–Tar* with Rom was performed by dissolving lyophilized aliquots of the protein in sodium phosphate buffer and 10 mM sodium chloride with a solution containing Tar–Tar* dissolved in the final buffer. Typical concentrations of the complexes at a 1:1 Rom homodimer to Tar–Tar* stoichiometry ranged from 1 to 2 mM. RNA imino spectra for the 1:1 Rom homodimer:Tar–Tar* complex and for Tar–Tar* with excess protein (2:1) (not shown) are similar, indicating that one Rom homodimer binds to one Tar–Tar* complex.

NMR spectroscopy

All experiments were recorded on a Bruker AMX 600 NMR spectrometer equipped with a three channel interface. ^{1}H and ^{15}N signals were referenced to DSS as described (35). Data were processed with the NMRPipe suite of programs (36) and Felix95 (Biosym Technologies). Typically the data were apodized in each

dimension with a phase-shifted sinebell followed by zero filling to twice the original size before Fourier transform. Peak volumes were measured with Felix95.

^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled Rom were recorded at 30°C with the fhsqc sequence (37). The ^{15}N carrier was set to 118 p.p.m. with a spectral width of 40 p.p.m. ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled Tar complexed with unlabeled Tar* were recorded at 25°C with the 1:1 echo HMQC sequence (38). The ^{15}N carrier was set to 150 p.p.m. with a spectral width of 30 p.p.m. All experiments were recorded with 128 complex ^{15}N points using the States-TPPI method of quadrature detection.

DQF-COSY spectra (39) of Rom and the Rom–Tar–Tar* complex were recorded at 30°C. A total of 450 points were collected in the indirectly detected dimension of each experiment using the TPPI method for quadrature detection. The spectral width in both dimensions was 6024 Hz.

RESULTS

Ultraviolet melting profiles

Figure 1 shows the derivatives of absorbance melting curves for the Tar–Tar* kissing complex in the presence and absence of Rom. The RNA concentration and Rom homodimer concentration are both 10 μM. Data are shown for sodium chloride concentrations of 50 and 150 mM. The melting temperature of Tar–Tar* increases from ~15°C at 50 mM NaCl to ~25°C at 150 mM NaCl. The addition of Rom to the RNA kissing complex in 50 mM NaCl increases its melting temperature by ~25°C. Increasing the NaCl concentration to 150 mM does not change the melting temperature of this complex, although the transition becomes sharper. $MgCl_2$ at a concentration of 150 μM has the same effect as NaCl at 150 mM, both in the presence and absence of Rom (data not shown).

Control experiments confirm that the protein contribution to the absorbance curves and their derivatives can be neglected (data not shown). The transitions detected by melting absorbance curves at 260 nm and its derivatives correspond to RNA transitions.

Rom signals assignments

Recently we reported (40) ^{1}H, ^{15}N, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ signal assignments for all but the first six residues (Gly1, Thr2, Lys3, Gln4, Asp5 and Lys6) of Rom using standard multidimensional heteronuclear methods (41). Incomplete cleavage of the N-terminal Met (Materials and Methods) caused doubling of NH resonances for Thr7, Ala8, Asn10 and Ala54. Complete aromatic ^{1}H assignments were obtained for one of the two phenylalanine residues (Phe54) and the single tyrosine (Tyr49). Only the H_{δ} assignment was obtained for Phe14 due to chemical shift degeneracy. Side chain ^{1}H-^{15}N assignments were also reported for the three Asn and one of the three Gln residues (Gln34). Many aliphatic side chain resonances were broader than expected for a 14.3 kDa protein. Possibilities for the broadening include the interchange of monomers and side chain dynamics at the helix interfaces. The broadening has not been investigated in detail and has hampered efforts to assign the H_{β} signals as well as the other side chain ^{1}H-^{13}C resonances. The chemical shift assignments have been deposited in the BioMagResBank under BMRB accession no. 4072.

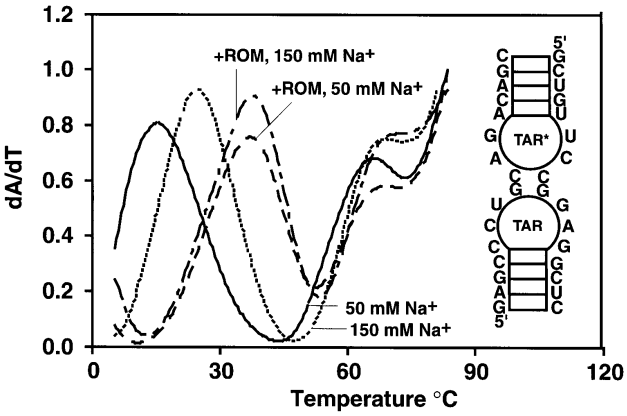

Figure 1. UV melting of 10 μM Tar–Tar* in the presence or absence of the Rom homodimer. The experiments were done in 10 mM $NaPO_4$, pH 6.3, 0.1 mM EDTA and 50 or 150 mM NaCl. The individual hairpins melt at ~65°C for Tar* and 80°C for Tar in 50 mM NaCl. The loop–loop interaction melts at ~15°C in the absence of Rom in 50 mM NaCl and at ~40°C in its presence at a 1:1 ratio. Rom stabilizes the interaction by ~25°C at these concentrations. Increasing the NaCl concentration from 50 to 150 mM stabilizes the Tar–Tar* kissing complex by ~10°C. The melting temperature of the complex in the presence of Rom does not change with increased salt concentration. All curves have been smoothed. The sequences of the Tar and Tar* hairpins are shown at the side.

Mapping of the Rom RNA-binding site

Protein backbone signals. The interaction of Rom with the Tar–Tar* kissing complex was monitored via ^{15}N-^{1}H HSQC spectra of uniformly ^{15}N-labeled Rom recorded with increasing equivalents of Tar–Tar* (Fig. 2). Each signal can be used as a probe for Tar–Tar* binding and dynamics. A chemical shift change reflects a change in environment due to either a structural change in the protein or because of its close proximity to the kissing hairpin. A linewidth change reports on dynamic processes such as the exchange of Tar–Tar* between different binding orientations or different protein molecules. Signals for residues near the RNA binding site are expected to undergo the largest chemical shift changes and therefore decrease in intensity and possibly broaden, in particular at substoichiometric RNA–protein ratios.

The NH signals can be divided into classes based on changes observed in the spectra with successive increases in Tar–Tar*. Most signals do not change significantly in either chemical shift or intensity upon addition of Tar–Tar*. In Figure 2, Phe56 and Leu52 show little change up to a Tar–Tar*:Rom homodimer ratio of 2:1. Residues in this class are far from the RNA binding site. All residues belonging to helix H2 are in this category. Some signals, including Ala8, Asn10, Arg13, Ile15 and Gln18, decrease in intensity by >40% upon addition of Tar–Tar* to a Tar–Tar*:Rom ratio of 0.4:1, indicating that signals for these residues undergo chemical shift changes upon binding of RNA. Asn10 illustrates this behavior in Figure 2. The signals from residues Thr7, Phe14, Ser17, Thr19 and Leu22 disappear when the RNA:protein ratio is increased from 0.4 to 0.8, again indicating chemical shift changes between free and bound forms for these NH signals. In Figure 2, Thr7 and Ser7 show this behavior. All these residues with large chemical shift changes define the RNA-binding interface. New resonances corresponding to protein NH signals

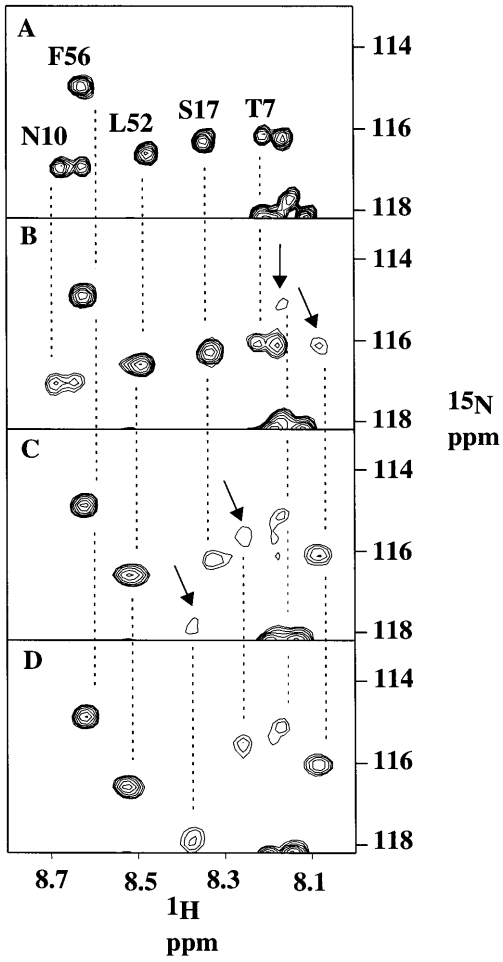

Figure 2. A portion of the ^{15}N-^{1}H HSQC spectra of uniformly ^{15}N-labeled Rom with (**A**) no added Tar–Tar*, (**B**) 0.4, (**C**) 0.8 and (**D**) 1.2 equivalents of Tar–Tar*. Backbone ^{15}NH signals are shown. Assignments for unbound Rom are denoted by the one letter amino acid code and residue number. Arrows denote new amide signals that appear due to interaction with the RNA kissing hairpin. Dashed lines connect signals from the same amino acid in each of the spectra. Doubling of Thr7 and Asn10 was due to incomplete cleavage of the N-terminal methionine (text).

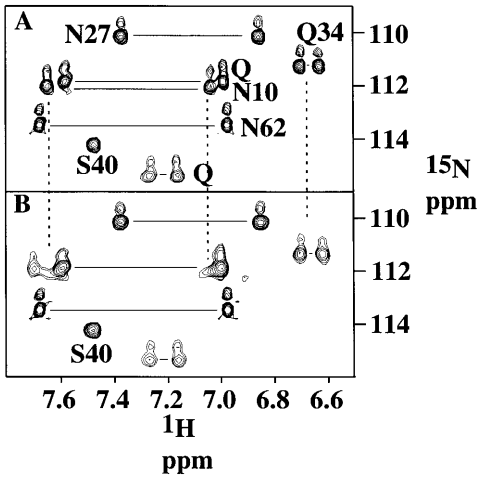

Figure 3. A portion of the ^{15}N-^{1}H HSQC spectra of uniformly ^{15}N-labeled ROM with (**A**) no added Tar–Tar* and (**B**) 0.4 equivalents of Tar–Tar*. Side chain $^{15}NH_2$ signals are shown. Dashed lines connect Asn10 signals in each of the spectra. Asn27 and Asn62 signals do not change as a consequence of the interaction with RNA, while Asn10 broadens (text).

in the bound form appear as the titration proceeds. However, these signals are generally broader than those for the unbound protein and have not been assigned.

The disappearance of some signals for free Rom and the appearance of signals for bound Rom show that the kissing hairpin is in slow exchange between protein molecules when dissolved in 10 mM phosphate buffer (pH 6.3) containing 150 mM sodium chloride at 30°C. Attempts were made to sharpen the resonances by either slowing or increasing the exchange rate between Rom and Tar–Tar*. Changes in sodium chloride concentration ranging from 10 mM to 1.4 M and variations in temperature from 10 to 45°C failed to improve the quality of the spectra. Indeed, the Rom signals sharpened with addition of sodium chloride to ~150 mM. Further increases caused broadening. The effect of magnesium was also investigated. No dramatic changes were observed in HSQC spectra recorded with samples dissolved in 10 mM phosphate, 10 mM sodium chloride and up to 20 equivalents magnesium ion/RNA complex. Magnesium ion had the same effect on linewidths as added sodium chloride but at lower concentrations. Resonances sharpened with increases in magnesium ion up to 5 mM and broadened with subsequent increases. These results agree well with the UV melting studies, which showed the same effects for magnesium and sodium on the protein–RNA complex melting temperature and no change in the melting temperature with increasing salt concentrations.

^{1}H-^{15}N HSQC spectra of free Rom showed the number of signals expected for a symmetrical homodimer. In contrast, the Tar–Tar* complex is asymmetrical, but it may have a pseudosymmetry if only the phosphate backbone is considered. The solution NMR structure reported by Chang and Tinoco (15) is consistent with an angle of 30–90° between the two hairpins. Upon binding, we expected the kissing hairpin to break the symmetry of Rom resulting in a doubling of many NH signals. However, spectra of Rom with excess Tar–Tar* show approximately the same number of signals as the free protein. One explanation is that the asymmetry of the two RNA hairpins causes a doubling and broadening of the bound form signals, rendering them undetectable. An alternative explanation for the absence of more signals is fast exchange of the kissing hairpins between symmetrical binding sites (alternative sampling by Tar and Tar* of the same Rom residues) coupled with slow exchange between the RNA and protein. In this case, there would be one averaged bound signal broadened by the slow exchange between the bound and free states of Rom. This would explain the broad resonances in the bound Rom spectra.

As a control experiment, the protein was titrated with the individual Tar and Tar* hairpins and monitored by ^{15}N-^{1}H HSQC spectra of uniformly ^{15}N-labeled Rom. Little or no changes are observed at a RNA to protein homodimer ratio of 0.4:1. At RNA to protein ratios of 2:1 we see some slight line broadening and shifting of resonances, mostly of the same protein signals as for the RNA kissing complex at a 40% RNA to protein ratio.

Protein side chain signals. Side chain NH_2 signals for the three Asn and three Gln residues were monitored in ^{1}H-^{15}N HSQC

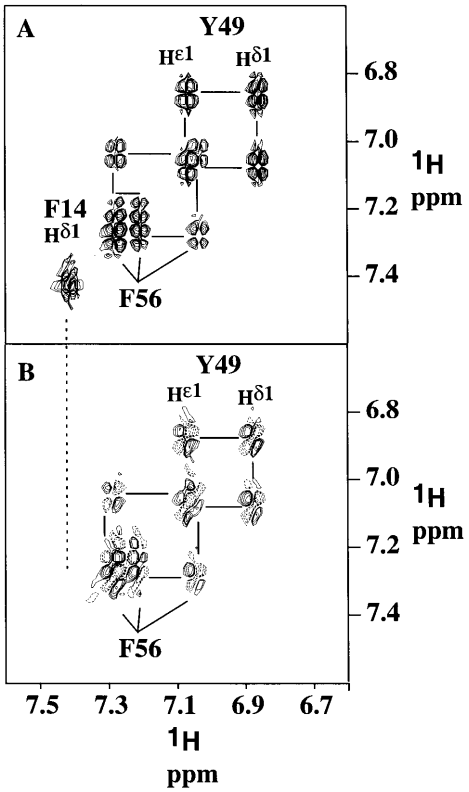

Figure 4. Parts of the DQF-COSY spectra of unlabeled Rom with (**A**) no added Tar–Tar* and (**B**) 0.4 equivalents of Tar–Tar*. The aromatic region is shown. Note that the Phe14 signal (dashed line) of the free protein disappears on addition of Tar–Tar*.

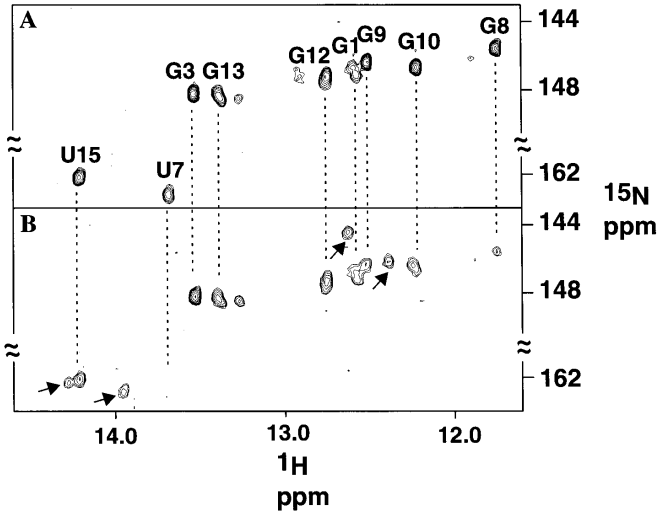

Figure 5. ^{15}N-^{1}H HSQC spectra of ^{15}N-labeled Tar in complex with Tar* and with (**A**) no added ROM and (**B**) 0.5 equivalents of Rom. Assignments for unbound Tar imino ^{15}NH signals are denoted by the one letter code and residue number. Arrows denote new RNA signals that appear due to interaction with Rom. Dashed lines connect signals from the same imino signal in each spectrum.

spectra of uniformly ^{15}N-labeled Rom at increasing RNA:protein ratios (Fig. 3). The NH signals for Asn10 broaden at a ratio of 0.4, indicating that the side chain for Asn10 interacts with the RNA or is close to the RNA-binding site. Upon further additions of RNA the signals sharpen again. Signals for the other NH_2 resonances do not change significantly, indicating that they do not play a role in RNA binding.

The aromatic ^{1}H signals for Phe14, Tyr49 and Phe56 were monitored via DQF-COSY spectra of unlabeled Rom with no added RNA (Fig. 4) and at a protein:RNA ratio of 0.4. As can be seen in Figure 4, the signal for Phe14 disappears when the kissing hairpin is added at a ratio of 0.4, indicating that the side chain is close to or interacts with the RNA. Signals for Tyr49 and Phe56 are unchanged, indicating that they do not interact with the RNA.

Tar–Tar* signal assignments

Chang and Tinoco have previously reported imino proton assignments for Tar, Tar* and the kissing complex Tar–Tar* obtained from NOESY spectra recorded in H_2O (13). The ^{1}H assignments were used to assign the ^{1}H-^{15}N signals in ^{1}H-^{15}N HSQC spectra of the complex (Fig. 5).

Interaction of the Tar–Tar* complex with Rom was monitored via ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled Tar in complex with unlabeled Tar* and with increasing concentrations of unlabeled Rom (Fig. 5). Analogous to protein amide signals, each G or U imino ^{15}N-^{1}H signal can be used as a probe for Rom binding and dynamics. The imino proton signals can be divided into classes based on changes observed with 0.5 equivalents of added Rom. Some signals (G3, G12 and G13) show no significant changes in chemical shift or intensity upon protein addition. Other signals (U7 and G8) decrease significantly (>80%) or show moderate decreases (50%) in intensity (G9, G10 and U15). Finally, several new resonances corresponding to Tar imino signals in the protein-bound complex appear at a protein:RNA ratio of 0.5. The appearance of signals corresponding to free and bound forms indicates slow exchange between the two species.

U7 and G8 are on the 5′-side of the loop. Their large changes indicate a strong interaction between the protein and RNA at these positions. Lesser but significant interactions occur for the remaining loop residues (G9 and G10), showing that the loop–loop structure is recognized by Rom. Changes are also found for stem residue U15, indicating a specific protein–RNA interaction in this area.

DISCUSSION AND CONCLUSIONS

Backbone

^{15}N-^{1}H signals of uniformly ^{15}N-labeled Rom were used as probes of RNA binding. Similar experiments have been used previously to map protein–DNA interfaces (42,43). Strongly perturbed signals, including those of Ala8, Asn10, Arg13, Ile15 and Gln18, show significantly reduced intensities at low RNA:protein ratios (0.4) and define the primary RNA-binding interface (Fig. 6). Other signals, including Thr7, Phe14, Ser17, Thr19 and Leu22, are perturbed at higher RNA:protein ratios (0.8) and define an extended interface. NH groups from both classes are found in the central portion of helices H1 and H1′, which forms one face of the Rom homodimer (Fig. 6). Thus, the NMR results indicate that the central residues of helices H1 and H1′ form the RNA-binding interface, in good agreement with

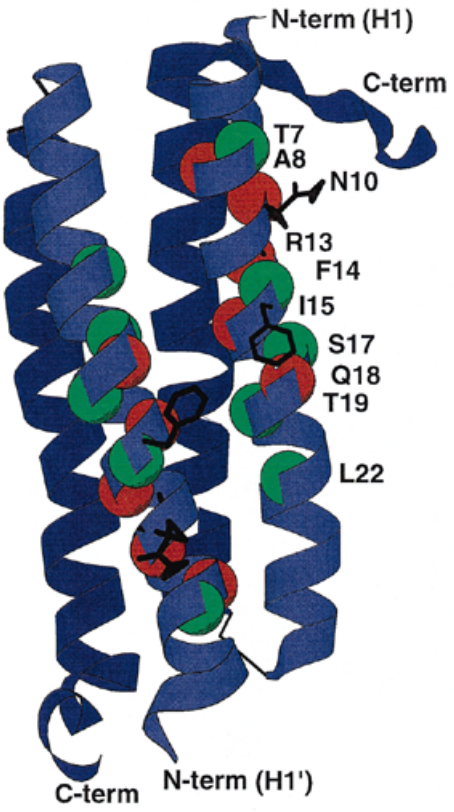

Figure 6. Ribbon model of Rom showing the RNA-binding interface. In each Rom monomer, helices H1 and H1′ are in light blue and helices H2 and H2′ are in dark blue. Helices H1 and H1′ face the viewer. Residues that experience strong and medium perturbations are highlighted in red and green, respectively. The side chains of Asn10 and Phe14, which also broaden upon RNA addition, are shown in black. This figure was created with the program Molscript (58) using coordinates from the NMR structure of Eberle *et al.* (25).

previous results based on mutation studies (29). The results do not require that each of these residues interacts directly with the RNA, but only that the surroundings are significantly perturbed by RNA binding and are therefore in close proximity to the RNA.

The NMR results include as part of the Rom–RNA interaction residues Asn10, Phe14 and Gln18, which were implicated in binding to the ColE1 kissing complex based on alanine scanning mutagenesis combined with gel shift assays (29). Lys25 was also implicated in the previous experiments. In HSQC spectra Lys25 retains significant intensity (>50%) even at a protein:RNA ratio of 1:2. Even though the Lys25 backbone NH signal is not significantly perturbed upon RNA binding, this does not preclude an interaction between its positively charged side chain and the RNA, as implied by the mutation experiments.

The N-terminal segment of Rom contains two lysines, Lys3 and Lys6, that form part of helix H1 in both NMR and X-ray structures of Rom. In the present work NH signals were not observed for the first six residues (Gly1–Lys6), presumably due to rapid exchange of the NH groups at pH 6.3, so no conclusions can be drawn regarding their role in complex formation. Others have suggested that Lys3 (28,29) and Lys6 (28) may interact with the phosphate backbone. The C-terminus is less likely to interact with RNA because it has two aspartic and one glutamic side chains, giving it a highly negative charge.

Protein side chains

The side chains of Asn10 and Phe14 broaden upon RNA addition, indicating that both interact with Tar–Tar* (Figs 3 and 4). Earlier studies suggested that Phe14 and Gln10 (29) also play a role in binding to the ColE1 complex. Because Phe14 is in the center of helices H1 and H1′, it was proposed that the ring intercalates or stacks with loop residues. There are precedents for this type of interaction between aromatic side chains and RNA bases. The RNP2 and RNP1 motifs of RNA-binding domains contain a conserved number of aromatic residues that are essential for the RNA–protein interaction. The role of Phe in the four RNA-binding domains of yeast poly(A)-binding protein has been studied by mutagenesis (44). The crystal structure of the RNA-binding domain of U1A revealed two aromatic side chain–base interactions (45). NMR studies of the N-terminal domain of U1A have confirmed the same interactions (46). Finally, the resolution by X-ray crystallography of a number of aminoacyl RNA synthetases bound to their cognate tRNAs and aminoacyl adenylates or ATP has revealed a variety of stacking interactions between protein side chains and RNA bases. These include stacking between bases and Phe or Tyr side chains as well as the side chains of Pro, Asn and Gln (47–53).

RNA

^{15}N-^{1}H signals of ^{15}N-labeled Tar were used as probes for the binding of Rom to kissing hairpins. Both loop and Tar stem regions were monitored (Fig. 7). The most significant intensity reductions occur for U7 and G8, which are located in the 5′-portion of the Tar loop. Imino signals for G9 and G10 also show reductions in intensity. Together these data show that Rom contacts the loop–loop region of the Tar–Tar* complex. This result is consistent with RNase V1 cleavage studies of ColE1 kissing hairpins by Eguchi and Tomizawa (27), who showed protection of loop–loop helix residues from cleavage in the presence of Rom. Our results are inconsistent with earlier work by Cesarini and Banner (54) and Helmer-Citterich *et al.* (55), who suggested that Rom binds mainly to the ColE1 stems. However, Rom does contact the stems several base pairs from the loop, as evidenced by reduced intensity for U15 upon protein binding. Phosphates 2 and 3 of ColE1 (equivalent to G1 and A2 for Tar–Tar*) were protected from ethylation by ethylnitrosourea, showing that Rom interacts in a similar region with the ColE1 wild-type kissing complex. Thus, our NMR results correlate well with previous findings based on phosphate ethylation and RNase digestion experiments for the wild-type kissing complex as well as sequence variants. Given the lack of sequence similarity between Tar–Tar* and ColE1, our findings support the notion of Eguchi and Tomizawa that Rom recognizes structure as opposed to a particular sequence (28).

Role of magnesium ion

Magnesium and sodium ions are known to stabilize the Tar–Tar* (13) and ColE1 (56,57) loop–loop complexes. At least two magnesium ions are involved for ColE1 (56). In the free RNA complex, short interphosphorus distances are observed between the 5′ phosphate groups of C5 and the 5′ phosphate groups of U7 and G8 of Tar, as well as between U*5 and C*7 and C*8 of Tar*. These are potential binding sites for magnesium. Nonetheless, no dramatic changes were observed in ^{15}N-^{1}H HSQC spectra of a

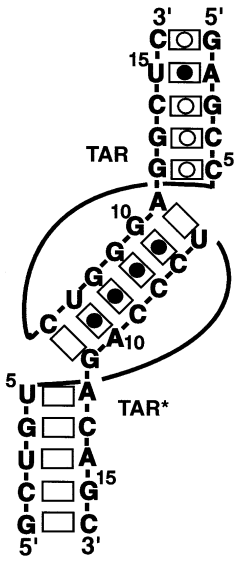

Figure 7. Schematic diagram of the Tar–Tar* complex showing the primary protein-binding interface. Rectangles connect bases that form hydrogen bonded pairs in the kissing complex. Imino signals that decrease in intensity on addition of Rom by 50% or more are indicated by black circles; signals that do not change are indicated by open circles. Only imino signals from the ^{15}N-labeled Tar hairpin can be detected.

1:1 Rom–Tar–Tar* complex with up to 20 equivalents of magnesium chloride. Added magnesium ions had the same effect on linewidths as sodium chloride, but at lower concentrations. This result agrees with the absorbance melting curves studies, that show equivalent roles for sodium and magnesium. We conclude that magnesium ions do not play a significant structural role in protein binding over that of sodium chloride. Additional experiments are needed to determine if these results apply to the interaction of Rom with other RNAs.

Complex formation

Using previously determined structures (15,25,26), we docked Rom to Tar–Tar* consistent with the present NMR results as well as previous mutagenesis (29) and ribonuclease cleavage data (27,28) for the ColE1 kissing complex (Fig. 8).

The free kissing hairpin complex in solution bends towards the major groove of the loop–loop helix with an angle of from 30 to 90° (15). However, the major groove is blocked by the phosphate backbones that cross it. The convex face of the bent helices exposes the minor groove of the loop–loop helix. Four of the 6 bp of this helix interact with Rom, as shown by the NMR experiments. The protein is also bent, with its concave face formed by helices H1 and H1′. We propose that the convex minor groove side of the loop–loop helix interacts with this concave protein surface.

We have made a model of the Rom–Tar–Tar* complex using Insight (Biosym Technologies) that fits our NMR data and the previously published biochemical data (28,29). In our model Lys3 and possibly Lys6 are located in the major groove near the Tar U15·Tar* A2 base pair. Although the amide signals for Lys3 and Lys6 were not assigned, alanine scanning mutagenesis (29) and genetic studies (30) suggest that Lys3 is important for binding. Docking Rom to Tar–Tar* in this manner accounts for

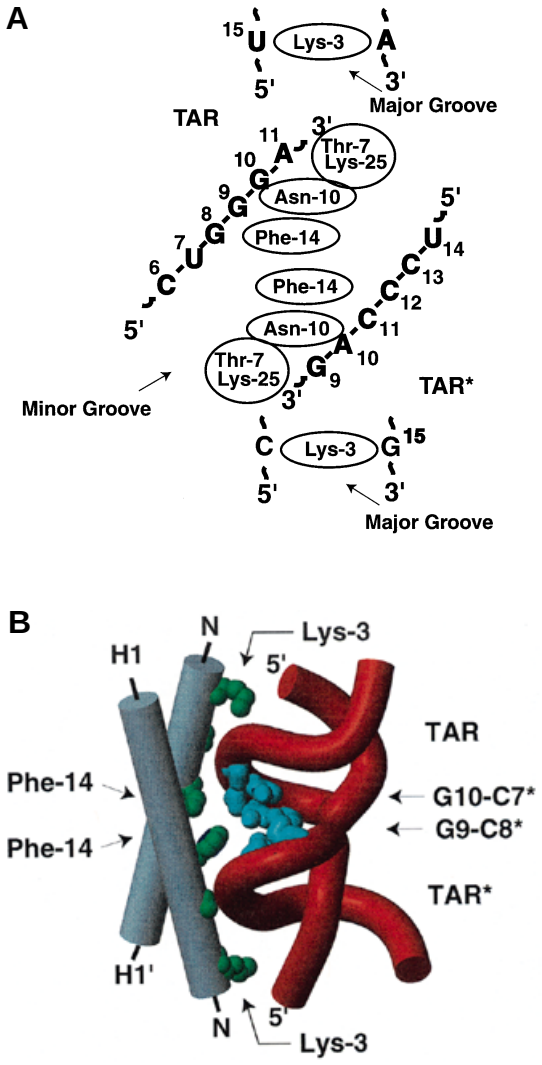

Figure 8. (**A**) A scale model of the binding of Rom to the Tar–Tar* complex based on changes observed in amide and imino NH signals during titrations, as well as results from mutagenesis (29) and ribonuclease cleavage studies (27,28). (**B**) A ribbon diagram of the complex showing the relative orientation of Rom and Tar–Tar*. Lys3, Asn10 and Phe14 (in green) of the equivalent α-helices, H1 and H1′, of Rom interact with the major grooves of the hairpin stems and the minor groove of the loop–loop helix of Tar–Tar*. Base pairs G9·C*12 and G10·C*13 of Tar–Tar* are shown in blue. The coordinates for Rom are from Eberle *et al.* (25); the Tar–Tar* coordinates are from Chang and Tinoco (15). The program MOLMOL (59) was used to make the figure.

the reduced intensity of the U15 imino signal (Fig. 7) and protection at this site from ribonuclease cleavage (27) upon protein binding. With Lys3 anchored in the major groove near Tar U15, the side chains of Asn10 and Phe14 contact the minor groove of the loop–loop helix near G8, G9 and G10 (Fig. 8). This orientation accounts for the broadening of Asn10 and Phe14 side chain signals and abolition of binding when these side chains are mutated to alanine (29). Of the other amide signals perturbed on binding, Ala8 is near G10 and A11, Arg13 and Ile15 flank Phe14 near U7, G8 and G9 and Gln18 is near U7 and G8. Evidence suggests that the complex has a pseudo 2-fold symmetry (27–29).

In our model symmetry-related residues in helix H1′ interact with Tar* in an analogous fashion. Thus, residues whose amide proton signals show the most pronounced intensity reductions can be placed in the minor groove of the loop–loop helix adjacent to the RNA residues whose imino proton signals show large intensity reductions (Fig. 8). There is a precedent of a coiled coil interacting with an RNA minor groove in the crystal structure of *Tetrahymena thermophilus* seryl tRNA aminoacyl synthetase complexed with $tRNA^{Ser}$ (47).

In addition to Asn10 and Phe14, Gln18 and Lys25 were shown to abolish binding when mutated to Ala (29). In our model the side chain of Gln18 is near U7 and G8 while Lys25 is close to the phosphate backbone near A11. A more detailed structure was not obtained because of the broad resonances seen in the complex.

A less likely, alternative model has Rom docked on the opposite face of Tar–Tar*. In this orientation Lys3 is located in the minor groove of Tar near U15 and both Asn10 and Phe14 are located in the major groove of the loop–loop helix. In order to close the hairpin loops, the phosphate backbone bridges the major groove between C5 and C6 for Tar and between U5 and U6 for Tar* (13). To span this distance the major groove is compressed and the kissing hairpin is bent (15), which hinders protein binding. Moreover, the backbone crossover blocks entry of Rom into the major groove at the junctions. For these reasons we favor the parallel binding orientation (Fig. 8), in which the central residues of α-helices H1 and H1′ interact with the RNA loop–loop helix from the minor groove face.

The results presented here provide the foundation for the study of other Rom–kissing hairpin complexes. Such studies should provide insight into the extent that Rom recognizes structure rather than sequence and may aid in the design of antisense strategies to control biological processes.

ACKNOWLEDGEMENTS

We thank Professor L.Regan of Yale University for the Rom plasmid, Dr David King, Mr Paul Schnier and Professor Evan Williams for mass spectral analysis, Ms Barbara Dengler for general assistance and Mr David Koh for synthesizing DNA templates. Dr Kung-Yao Chang started studies on the Rom–kissing hairpin complex; we are very grateful for his help. This work was supported in part by National Institutes of Health grant GM 10840, by Department of Energy grant DE-FG03-86ER60406 and through instrumentation grants from the Department of Energy (DE-FG05-86ER75281) and from the National Science Foundation (DMB 86-09305).

REFERENCES

1 Knee,R. and Murphy,P.R. (1997) *Neurochem. Int*., **31**, 379–392.
2 Wagner,E.G.H. and Simons,R.W. (1994) *Annu. Rev. Microbiol*., **48**, 713–742.
3 Eguchi,Y., Itoh,T. and Tomizawa,J. (1991) *Annu. Rev. Biochem*., **60**, 631–652.
4 Cesareni,G., Helmer-Citterich,M. and Castagnoli,L. (1991) *Trends Genet*., **7**, 230–235.
5 Polisky,B. (1988) *Cell*, **55**, 929–932.
6 Simons,R.W. and Kleckner,N. (1988) *Annu. Rev. Genet*., **22**, 567–600.
7 Weeks,K.M., Ampe,C., Schultz,S.C., Steitz,T.A. and Crothers,D.M. (1990) *Science*, **249**, 1281–1285.
8 Berkhout,B., Silverman,R.H. and Jeang,K.T. (1989) *Cell*, **59**, 273–282.
9 Feng,S. and Holland,E.C. (1986) *Nature (Lond.)*, **334**, 165–167.
10 Muesing,M.A., Smith,D.H. and Capon,D.J. (1987) *Cell*, **48**, 691–701.
11 Sheline,C.T., Milocco,L.H. and Jones,K.A. (1991) *Genes Dev*., **5**, 2508–2520.
12 Wu,F., Garcia,J., Sigman,D. and Gaymor,R. (1991) *Genes Dev*., **5**, 2128–2140.
13 Chang,K.-Y. and Tinoco,I.,Jr (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8705–8709.
14 Graham,G.J. and Maio,J.J. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 5817–5821.
15 Chang,K.-Y. and Tinoco,I.,Jr (1997) *J. Mol. Biol*., **269**, 52–66.
16 Marino,J.P., Razmic,S.G., Scanlovszki,G.J. and Crothers,D.M. (1995) *Science*, **268**, 1448–1454.
17 Draper,D. (1995) *Annu. Rev. Biochem*., **64**, 593–620.
18 Nagai,K. (1996) *Curr. Opin. Struct. Biol*., **6**, 53–61.
19 Chang,K.-Y. and Varani,G. (1997) *Nature Struct. Biol*., **4**, 854–857.
20 Ramos,A., Gubser,C.C. and Varani,G. (1997) *Curr. Opin. Struct. Biol*., **7**, 317–323.
21 Allain,F.H., Gubser,C.C., Howe,P.W., Nagai,K., Neuhaus,D. and Varani,G. (1996) *Nature (Lond.)*, **380**, 646–650.
22 Allain,F.H., Howe,P.W., Neuhaus,D. and Varani,G. (1997) *EMBO J*., **16**, 5764–5772.
23 Oubridge,C., Ito,N., Evans,R.P., Teo,C.-H. and Nagai,K. (1994) *Nature (Lond.)*, **372**, 432–438.
24 Cusack,S. (1997) *Curr. Opin. Struct. Biol*., **7**, 881–889.
25 Eberle,W., Pastore,A., Sander,C. and Rösch,P. (1991) *J. Biomol. NMR*, **1**, 71–82.
26 Banner,D.W., Kokkinidis,M. and Tsernoglou,D. (1987) *J. Mol. Biol*., **196**, 657–675.
27 Eguchi,Y. and Tomizawa,J. (1990) *Cell*, **60**, 199–209.
28 Eguchi,Y. and Tomizawa,J. (1991) *J. Mol. Biol*., **220**, 831–842.
29 Predki,P.F., Nayak,L.M., Gottlieb,M.B.C. and Regan,L. (1995) *Cell*, **80**, 41–50.
30 Castagnoli,L., Scarpa,M., Kokkinidis,M., Banner,D.W., Tsernoglou,D. and Cesareni,G. (1989) *EMBO J*., **8**, 621–629.
31 Milligan,J.F. and Uhlenbeck,O.C. (1989) *Methods Enzymol*., **180**, 51–62.
32 Batey,R.T., Inada,M., Kujawinski,E., Puglisi,J.D. and Williamson,J.R. (1992) *Nucleic Acids Res*., **20**, 4515–4523.
33 Nikonowicz,E.P., Sirr,A., Legault,P., Fiona,M.J., Baer,L.M. and Pardi,A. (1992) *Nucleic Acids Res*., **20**, 4507–4513.
34 Haynie,S.L. and Whitesides,G.M. (1990) *Appl. Biochem. Biotechnol*., **23**, 205–220.
35 Wishart,D.S., Bigam,C.G., Yao,J., Abildgaard,F., Dyson,H.J., Oldfield,E., Markley,J.L. and Sykes,B.O. (1995) *J. Biomol. NMR*, **6**, 135–140.
36 Delaglio,F., Grzesiek,S., Vuister,G., Zhu,G., Pfeifer,J. and Bax,A. (1995) *J. Biomol. NMR*, **6**, 277–293.
37 Mori,J., Abeygunawardan,C., O'Neil-Johnson,M. and van Zijl,P.C.M. (1995) *J. Magn. Resonance, B*, **108**, 94–98.
38 Bax,A., Ikura,M., Kay,L.E., Torchia,D.A. and Tschudin,R. (1990) *J. Magn. Resonance*, **86**, 304–318.
39 Rance,M., Sorensen,O.W., Bodenhausen,G., Wagner,G., Ernst,R.R. and Wüthrich,K. (1983) *Biochem. Biophys. Res. Commun*., **117**, 479–485.
40 Pelton,J.G. and Comolli,L.R. (1998) *J. Biomol. NMR*, **11**, 1–2.
41 Bax,A. and Grzesiek,S. (1993) *Acc. Chem. Res*., **26**, 131–138.
42 Dekker,N., Cox,M., Boelens,R.V.,C.P., van der Vliet,P.C. and Kaptein,R. (1993) *Nature (Lond.)*, **362**, 825–855.
43 Vuister,G.W., Keim,S.J., Orosz,A., Marguardt,J., Wu,C. and Bax,A. (1994) *Nature Struct. Biol*., **1**, 605–614.
44 Sacks,B.A. and Deardoff,J.A. (1997) *J. Mol. Biol*., **272**, 82–94.
45 Nagai,K., Teo,C.H., Evans,R.P., Nobutoshi,I. and Oubridge,C. (1994) *Nature (Lond.)*, **372**, 432–438.
46 Varani,G., Neuhaus,D., Nagai,K. and Howe,W.A. (1994) *EMBO J*., **13**, 3873–3881.
47 Cusack,S., Biou,V., Yarem,C.A., Tukalo,M. (1994) *Science*, **263**, 1404–1410.
48 Cusack,S., Yaremchuk,A., Tukalo,M. (1996) *EMBO J*., **15**, 6321–6334.
49 Goldgur,Y., Mosyak,L., Reshetnikova,L., Ankilova,V., Lavrik,O., Khodyreva,S. and Safro,M. (1997) *Structure*, **5**, 59–68.
50 Moras,D., Cararelli,J., Rees,B., Ruff,M., Thierry,J.C. (1993) *Nature (Lond.)*, **362**, 181–184.
51 Moras,D., Arney,J.G., Hans,D.C., Mitschler,A., Rees,B. and Franckly,C.S. (1995) *EMBO J*., **14**, 4143–4155.
52 Steitz,T.A., Rould,M.A., Peroma,J.J. and Soll,D. (1989) *Science*, **246**, 1135–1142.
53 Steitz,T.A., Perona,J.J. and Rould,M.A. (1991) *Nature (Lond.)*, **352**, 213–218.
54 Cesareni,G. and Banner,D.W. (1985) *Trends Biochem. Sci*., **10**, 303–306.
55 Helmer-Citterich,M., Anceschi,M.M., Banner,D.W. and Cesareni,G. (1988) *EMBO J*., **7**, 557–566.
56 Gregorian,R.S.,Jr and Crothers,D.M. (1995) *J. Mol. Biol*., **248**, 968–984.
57 Tomizawa,J. (1990) *J. Mol. Biol*., **212**, 695–708.
58 Kraulis,P.J. (1991) *J. Appl. Crystallogr*., **24**, 946–950.
59 Koradi,P., Billeter,M. and Wüthrich,K. (1996) *J. Mol. Graphics*, **14**, 51–55.