Using Phosphorothioate-Substituted RNA To Investigate the Thermodynamic Role of Phosphates in a Sequence Specific RNA−Protein Complex†

Dagmar Dertinger, Linda S. Behlen, and Olke C. Uhlenbeck*

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215

Received July 29, 1999

ABSTRACT: Part of the binding affinity and specificity in RNA−protein complexes is often contributed by contacts between the protein and backbone phosphates that are held in position by the RNA structure. This study focuses on the well-characterized interaction between a dimer of the MS2 coat protein and a small RNA hairpin. Using a short oligoribonucleotide which contains all the necessary sequence elements required for tight protein binding, a single phosphorothioate linkage was introduced at 13 different positions. In each case, the R_P and S_P stereoisomers were separated and their affinities to the MS2 coat protein were determined. Comparison of these biochemical data with the crystal structure of the protein−hairpin complex indicates that introduction of a phosphorothioate only affects binding at sites where a protein−phosphate contact is observed in the crystal structure. This means that phosphorothioate-containing oligoribonucleotides should also be useful for mapping phosphate contacts in RNA−protein complexes for which no crystal structure is available.

Phosphates in the nucleic acid backbone often play an essential role in nucleic acid−protein interactions (*1*−*3*). They are typically well-exposed on the surface of the nucleic acid. In a B-form helix, for example, one of the nonbridging oxygens points outward into solution and the other one toward the major groove. Therefore, these two oxygens are readily accessible to proteins and can contribute to both binding affinity and specificity. The phosphate oxygens can form hydrogen bonds with the protein as well as participate in electrostatic interactions with positively charged amino acid side chains. The contribution of each individual contact to the total binding energy presumably varies with the type of interaction, the orientation of the contacting atoms in space, the hydrophobicity of the local environment, and the solvation of both the protein and the nucleic acid free in solution.

This study utilizes the well-characterized interaction between the MS2 coat protein homodimer and a hairpin in its RNA genome (*4*−*7*) to investigate the role of phosphates in a sequence specific RNA−protein interaction. The 2.7 Å resolution crystal structure of the complex (*8*) indicates that, in addition to several base specific interactions, eight of the phosphates also contact the protein (Figure 1A). The nonbridging phosphate oxygens form ionic contacts with two arginines and three lysines and serve as hydrogen bond acceptors with one tyrosine, one asparagine, and one serine. In addition, one of the phosphates makes an intramolecular contact with a 2′-hydroxyl group, presumably stabilizing the structure of the bound RNA. Several biochemical experiments support the view that the protein−phosphate contacts in MS2 contribute substantially to the overall binding energy. Studies of the binding affinity as a function of the monovalent ionic strength revealed an upper limit of five ionic contacts (*9*), which agrees well with the five ionic contacts seen in the crystal structure of the complex. Under physiological conditions, the electrostatic interactions have been estimated to contribute about 20% of the total free energy of binding (*9*). Previous phosphorothioate substitution experiments with transcribed RNA hairpins containing several phosphorothioate linkages deduced four positions where the modification led to either a reduction or an increase in binding affinity (*10*). Complete analysis of phosphorothioates with transcribed hairpins was hampered by the fact that only R_P isomers are incorporated into the RNA and because it is often impossible to design hairpins that will contain only one substitution.

The work described here examines in greater detail the effects of individual phosphorothioate substitutions on the binding of the RNA to the MS2 coat protein. This is the first time that such a "phosphorothioate walk" with chemically synthesized oligoribonucleotides has been applied to a protein−RNA complex. However, phosphorothioate modifications have been used to study several DNA−protein interactions, particularly complexes between restriction enzymes and their substrates (*11*−*13*). Substitution of oxygen by sulfur is fairly conservative in terms of size and properties, but differences in charge distribution, hydrophobicity, and its ability to hydrogen bond are expected to alter the contribution of the individual protein−phosphate contacts to the overall binding affinity and, therefore, change the stability of the whole complex. The data obtained in these experiments will be used to try to answer the following questions. Is phosphorothioate substitution a good method for identifying protein−phosphate contacts in other RNA−

† This work was financed by NIH Grant GM36944 to O.C.U. and graduate student fellowships from the Gottlieb Daimler- and Karl Benz-Foundation and the German Academic Exchange Service (DAAD) to D.D.

* To whom correspondence should be addressed. E-mail: Olke.Uhlenbeck@Colorado.edu.

10.1021/bi991769v CCC: $19.00
Published on Web 12/09/1999

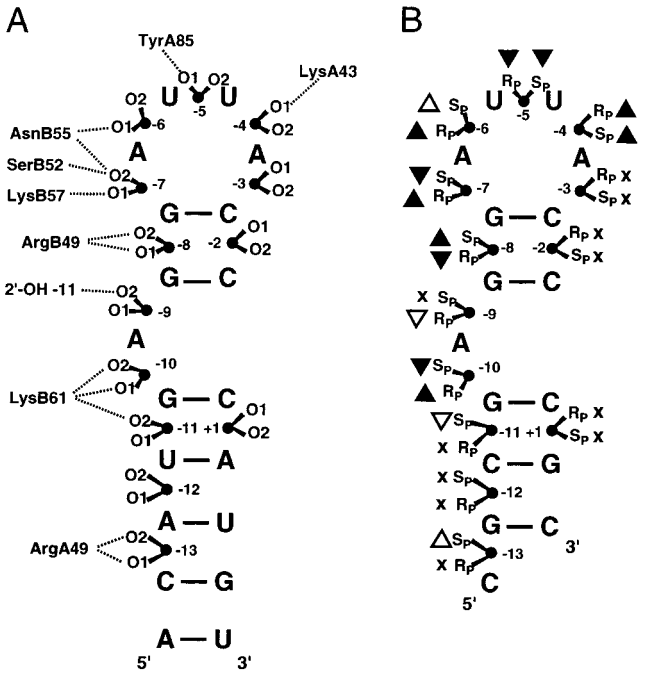

FIGURE 1: (A) Two-dimensional representation of the protein−phosphate contacts seen in the crystal structure of the MS2 coat protein bound to a 21-nucleotide RNA fragment (*8*). The prefixes A and B specify amino acids in the two monomers of the asymmetric MS2 coat protein dimer. Substitution of O1, as defined by the crystal structure, with sulfur leads to the R_P isomer and substitution of O2 to the S_P isomer. (B) Results of the phosphorothioate substitution data superimposed onto hairpin 2. Black triangles (▲) mark the positions of sulfur substitutions which clearly affect binding of the RNA to the protein in all three data sets. The triangles that are pointing up represent binding stronger than that of unmodified RNA, and the triangles that are pointing down represent weaker binding. Effects that are close to the limit of detection are identified by a white triangle (△) and positions where the binding is unaffected by sulfur substitution by an **x**.

protein complexes of unknown structure? Do all of the protein−phosphate contacts observed in the crystal structure contribute thermodynamically to the overall binding energy? Can phosphorothioate substitutions be used to distinguish between ionic contacts and hydrogen bonds? Is it possible to rationalize the results of this modification study on a molecular level?

MATERIALS AND METHODS

RNA Synthesis. All RNAs used in this study were chemically synthesized on an ABI DNA/RNA synthesizer using standard phosphoramidite techniques (*14*). RNA amidites, solid supports, and ancillary reagents were purchased from Glen Research. Phosphorothioate-containing RNAs were obtained by substitution of the iodine/water oxidation step with a sulfurization step using Beaucage reagent (*15*), according to Glen Research. Deprotection of oligoribonucleotides was carried out following one of two protocols described previously (*16*, *17*). The full-length product was isolated by electrophoresis on a denaturing 20% polyacrylamide gel.

The chemical synthesis of RNAs containing a single phosphorothioate linkage yields a mixture of two stereoisomers, which can usually be separated by reversed phase HPLC (*18*−*21*). Most oligonucleotides were purified using a Nucleosil C-18 column (Alltech). The best resolution of the two isomers was achieved at 65 °C and a flow rate of 1.5 mL/min in a 0.1 M ammonium acetate buffer system (pH 7.0). Increasing the acetonitrile content from 2.9% to 4.7% over the course of 25 min allowed a separation of the R_P and S_P isomer peaks by 3−5 min. The earlier-eluting isomer corresponds to the R_P phosphorothioate and the late-eluting isomer to the S_P phosphorothioate (*11*, *18*, *20*). The small amount of unmodified all-oxygen RNA that is introduced during synthesis and workup generally elutes before the peaks of the phosphorothioate-containing RNAs. In some cases, however, the all-oxygen RNA is not resolved from the earlier-eluting isomer. Although this contaminant generally is less than 10% of the total material, a better purification protocol that eliminates this problem was developed.

It has been reported previously that anion exchange columns can be used to purify unmodified RNA from oligonucleotides that contain a phosphorothioate or phosphorodithioate linkage (*22*). A NucleoPac anion exchange column (Dionex) was employed at 65 °C with a flow rate of 2.0 mL/min. A 0.1 M ammonium acetate buffer system (pH 8.0) with 2% acetonitrile and an increasing KCl concentration (from 370 to 390 mM over the course of 30 min) allowed the full separation of all three compounds in one step, independent of the position of the modification within the RNA molecule. This system can be used for those positions that give incomplete separation by reversed phase HPLC and is recommended for future experiments. Purified RNAs were desalted and then stored in water or TE [10 mM Tris-HCl (pH 8.0) and 1 mM EDTA] at −20 °C. Under these storage conditions, the phosphorothioate linkage was stable against hydrolysis and oxidation for at least 3 years as confirmed by reanalyzing the samples by HPLC.

RNAs were 5′-^{32}P-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP or 3′-^{32}P-labeled with T4 RNA ligase and 5′-[^{32}P]pCp, using protocols described previously (*23*−*25*). The labeled product was subsequently purified by electrophoresis on a denaturing 20% polyacrylamide gel. The existence and correct position of the phosphorothioate linkage within an oligonucleotide were confirmed by iodine cleavage and coelectrophoresis with an alkaline hydrolysis ladder on a denaturing 20% polyacrylamide gel (*16*). Partial digestion with T1 ribonuclease was used to confirm the position of guanidine residues within the primary sequence of each oligonucleotide (*26*).

Equilibrium Dissociation and Kinetic Rate Constants. All experiments were performed using the MS2 coat protein dimer carrying a double mutation (Val75Glu/Ala81Gly) that is unable to form capsids, but binds RNA in a manner indistinguishable from that of the wild-type protein (*27*). Cloning, overexpression, and purification of this protein were carried out as described previously (*28*).

A nitrocellulose filter binding assay in a microtiter format was used to obtain equilibrium dissociation constants (K_ds) (*16*). Each experiment included an unmodified RNA control to account for differences in protein preparation and experimental variability in protein dilution. The variation of the absolute K_d of one given hairpin is less than 3-fold in independent experiments. Relative binding affinities are calculated by dividing the K_d value of the unmodified RNA by that of the modified RNA determined in the same experiment. Each experiment was repeated at least three

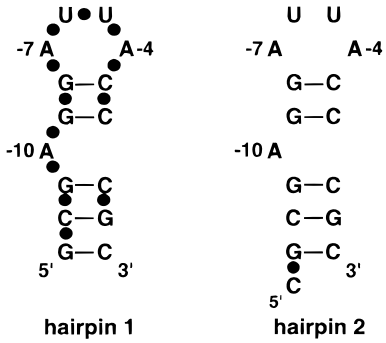

FIGURE 2: RNA oligonucleotides used in this study with the sites of phosphorothioate substitutions denoted with black circles (●).

times, and individual results are within 2-fold of the reported average.

Association rate constants were measured for several of the modified RNAs by mixing trace 3′-^{32}P-labeled RNA with several protein concentrations and filtering aliquots at time intervals (*9*). When the initial rate is plotted against the protein concentration, the slope equals k_{on}. Dissociation rate constants were measured by forming a complex between trace 3′-^{32}P-labeled RNA and near-saturating protein concentrations, adding an excess of unlabeled RNA, and filtering aliquots at different times after the chase was added (*16*). Again, the value of k_{off} of unmodified RNA was divided by that of the modified RNA to give a relative rate to allow for better comparison with the equilibrium binding data. All rate measurements were repeated at least twice, and the individual results are within 2-fold of the reported average.

RESULTS

Preparation of RNA Oligonucleotides. A 15-nucleotide RNA, hairpin 1, was designed (Figure 2) to contain all of the sequence elements required for recognition by the MS2 coat protein. Under standard conditions, hairpin 1 binds the protein with a K_d of 1 nM, which is very similar to that of the natural RNA substrate and several other consensus hairpins (*29*−*31*). Individual phosphorothioate linkages were introduced synthetically at 12 different positions within hairpin 1 (Figure 2). The phosphates are numbered according to the standard numbering system based on the position of the first nucleotide of the start codon of the replicase gene (*5*). To study position −13 as an internal phosphate, hairpin 2 was used, which has an extra cytidine residue at the 5′-terminus compared with hairpin 1 (Figure 2). In each case, the short length of the oligoribonucleotides permitted the purification of the two phosphorothioate stereoisomers by reversed phase or anion exchange HPLC. The resulting 26 modified RNA hairpins incorporate sulfur into all the positions where phosphate contacts are seen in the crystal structure of the complex as well as into several control positions where no contacts are observed.

Measurement of the Dissociation Constant. To evaluate the effect of a phosphorothioate substitution at a given position on the binding affinity, a nitrocellulose filter binding assay was used to determine the K_ds of ^{32}P-labeled RNA hairpins. An unmodified control was included in each experiment to ensure the reproducibility of the assay. The value of the K_d was found to depend on whether the RNA was 5′-^{32}P-labeled or 3′-^{32}P-labeled. This unusual situation is most likely due to a contact between phosphate −13 and ArgA49, which was not detected in the initial electron density maps (*32*), but later observed at higher resolution (*8*). As a result of this contact, hairpin 1 has a 10-fold higher affinity for the protein when it is 5′-^{32}P-labeled than when it is 3′-^{32}P-labeled where it lacks the contact between ArgA49 and the phosphomonoester. This hypothesis was confirmed by the finding that the binding affinity of hairpin 2 is the same when it is 5′-^{32}P-labeled or 3′-^{32}P-labeled. While it is unlikely that the presence or absence of this phosphate contact influences the outcome of phosphorothioate modifications at other positions within the RNA, K_ds were obtained for all 26 variants of hairpins 1 and 2 labeled at either their 3′- or 5′-ends, yielding two independent measurements.

Rate Measurements. Since the equilibrium data showed that most effects on K_d due to phosphorothioate modification are quite small, it was desirable to have a second, independent method for determining the binding affinity. Rate measurements are usually more reproducible and more sensitive than K_d measurements, because they do not rely on performing accurate dilutions of the protein (*12*). Previous experiments established that the k_{on} for RNA binding to the MS2 coat protein is diffusion-controlled (*9*), making it unlikely that backbone modifications would have any impact on the association rate constant. To be certain, k_{on} was determined for several of the RNAs. The association rate constants were found to be $(0.7 \pm 0.1) \times 10^8$ M^{-1} min^{-1}, which agrees very well with the values of 0.6×10^8 and 1×10^8 M^{-1} min^{-1} previously reported for different consensus hairpins (*9*, *33*). Thus, differences in binding affinity are expected to be solely due to differences in the dissociation rate constants. Therefore, k_{off} measurements were obtained for each 3′-^{32}P-labeled phosphorothioate-containing RNA using a chase protocol (*16*) to give a third independent measurement.

Effects of Sulfur Substitution on Complex Stability. Table 1 summarizes the data of the experiments described above. All the rate and equilibrium data are reported as relative values obtained by dividing the number of the unmodified RNA by that of the modified one. Effects that are less than 2-fold are considered within the experimental error and therefore not regarded as significant. The three data sets are in excellent agreement with one another. The only exception is the RNA containing the R_P phosphorothioate at position −6, which binds 17-fold tighter than the control RNA when both are 3′-^{32}P-labeled, but only 3-fold tighter than the control RNA when both are 5′-^{32}P-labeled. It is likely that the latter data are incorrect since very tight K_d values encountered with some 5′-^{32}P-labeled MS2 hairpins are difficult to determine accurately due to denaturation of the MS2 coat protein at the very low protein concentrations needed to determine such a tight K_d (*33*). This conclusion is supported by the fact that the dissociation rate of the R_P −6 oligomer is 23-fold slower than the control when 3′-^{32}P-labeled and 30-fold slower when 5′-^{32}P-labeled. The close agreement between the three independent measurements for all the other positions supports the view that the presence or absence of the contact of phosphate −13 with ArgA49 does not significantly affect the thermodynamic contribution of protein−phosphate contacts at other sites to the overall binding affinity and that the association rate is unaffected by phosphorothioate substitutions.

Table 1: Effects of the Phosphorothioate Modifications at Different Positions on K_d and k_{off}

position of phosphate[a]	stereoisomer	relative K_d of 5′-labeled RNA[b]	relative K_d of 3′-labeled RNA[b]	relative k_{off} of 3′-labeled RNA[b]
+1	R_P	*1.0*	*1.0*	*1.3*
	S_P	*0.6*	*0.6*	*0.9*
−2	R_P	*1.9*	*1.0*	*0.9*
	S_P	*1.1*	*0.9*	*0.9*
−3	R_P	*1.2*	*0.6*	*0.7*
	S_P	*1.4*	*1.3*	*0.7*
−4	R_P	**2.4**	**3.2**	**2.5**
	S_P	**2.3**	**2.9**	**2.0**
−5	R_P	**0.2**	**0.5**	**0.3**
	S_P	**0.5**	**0.5**	**0.5**
−6	R_P	**2.8**	**16.8**	**23.8**
	S_P	*1.4*	*1.6*	**2.3**
−7	R_P	**2.3**	**5.7**	**3.6**
	S_P	**0.3**	**0.3**	**0.3**
−8	R_P	**0.2**	**0.2**	**0.4**
	S_P	**2.4**	**4.4**	**5.4**
−9	R_P	*0.6*	**0.5**	**0.5**
	S_P	*0.8*	*0.6*	*0.9*
−10	R_P	**3.0**	**3.1**	**2.3**
	S_P	**0.3**	**0.2**	**0.2**
−11	R_P	**0.5**	**0.4**	*0.6*
	S_P	*0.9*	*0.9*	*0.9*
−12	R_P	*0.8*	*1.2*	*0.9*
	S_P	*0.9*	*1.8*	*1.2*
−13	R_P	*1.0*	*1.0*	*1.0*
	S_P	*1.2*	**2.2**	**2.5**

[a] Phosphorothioates at all positions with the exception of phosphate −13 were introduced into hairpin 1. [b] Relative numbers are obtained by dividing the value of the unmodified RNA by that of the modified one. A typical value for the K_d of unmodified hairpin 1 is 1 nM when 5′-^{32}P-labeled and 10 nM when 3′-^{32}P-labeled and for k_{off} is 0.4 min^{-1} when 3′-^{32}P-labeled. The corresponding values for the unmodified hairpin 2 are 3 nM, 3 nM, and 0.1 min^{-1}. Effects that are less than 2-fold, shown in italics, are within the experimental error and therefore not regarded as significant. This cutoff is somewhat arbitrary for positions where the effects are just on the borderline of detection.

In most cases, the effect of a single phosphorothioate substitution on the overall protein binding affinity is relatively small (2−5-fold). The replacement of a given nonbridging oxygen by sulfur can either strengthen or weaken the binding of the RNA to the coat protein. At some positions, only one phosphorothioate isomer affects the binding affinity while the other isomer does not. In cases where both the R_P and S_P isomers affect the binding affinity, the effect can be either in the same or in the opposite directions. The individual results of this phosphorothioate modification study will be interpreted in the next section and compared to information obtained by examining the crystal structure of the complex.

DISCUSSION

One of the motivations for this work was to establish whether phosphorothioate substitution experiments would accurately identify all the protein−phosphate contacts in the MS2 system and thus be useful for identification of such contacts in RNA−protein complexes of unknown structure. Figure 1 shows a schematic representation of the phosphorothioate substitution data compared with the position of protein−phosphate contacts as observed in the crystal structure of the complex (*8*). In a general way, the agreement between the structure and the modification data is excellent. At every one of the eight phosphates where an RNA−protein contact is observed in the crystal structure, the affinity of at least one of the two phosphorothioate isomers is affected by the substitution. At all four control phosphates where no contact is present in the crystal structure, the K_d of both phosphorothioate stereoisomers is unchanged. These results strongly suggest that all the protein−phosphate contacts observed in the crystal structure contribute thermodynamically to the total free energy of complex formation. Figure 3 is a three-dimensional representation of the data shown in Figure 1. It can be clearly seen that the phosphates that contact the protein, as identified both by the crystal structure and by the modification data, nicely define the RNA docking site onto the protein. Therefore, phosphorothioate substitution is a useful method for identifying RNA surfaces involved in the binding of a protein.

The MS2 system has been used previously for a study with phosphorothioate-containing RNAs (*10*), in which the phosphorothioate linkages were introduced into the hairpin by in vitro transcription using one specific NTP(αS) at a time. As a result, each RNA usually contained several modifications, and only the R_P isomers were present. By using different NTP(αS)s and hairpin sequences, phosphorothioate positions that changed the binding affinity were deduced using the assumption that the thermodynamic effects of several phosphorothioate substitutions per RNA hairpin are additive. It was concluded that only R_P phosphorothioates at positions −3, −7, −8, and −9 affected affinity. In contrast, we observe no effect at position −3 and find additional effects of R_P sulfurs at positions −4, −5, −6, −10, and −11. Presumably, the discrepancies are due to offsetting effects at several positions and the lack of thermodynamic additivity for RNAs containing multiple phosphorothioate substitutions (*12*, *34*, *35*). While it is possible that some of the additional R_P sites could have been identified by performing ensemble experiments (*36*, *37*) involving partial substitution with NTP(αS)s, the relatively small effect of phosphorothioate modifications on K_d would make such an experiment difficult especially since modification can both decrease and increase affinity. Thus, the greater sensitivity of detection and availability of the S_P isomer make the chemical synthesis approach the preferred one.

Introduction of a phosphorothioate for a phosphodiester linkage is unlikely to alter the overall RNA structure. However, the phosphorothioate substitution is expected to cause small local structural effects and minor changes in the electrostatic potential. The net negative charge at the modified position is redistributed such that it is mainly localized on the sulfur atom (*38*, *39*), and accordingly, the π-bond is rearranged to preferentially form a double bond between the phosphorus and the unsubstituted oxygen. In the crystal structure of uridine 2′,3′-*O*,*O*-cyclophosphorothioate (*40*), the P−S bond length is about 0.4 Å longer and the P═O bond is slightly shorter than the corresponding P−O bonds in the phosphodiester. Distortion of the O−P−S angle due to the larger size of the sulfur could not be detected. These differences listed above are not great. Indeed, in the 2.17 Å resolution crystal structure of a B-DNA oligomer containing alternating phosphodiester and R_P phosphorothioate linkages (*41*), no significant differences between the modified and unmodified phosphates were observed. Furthermore, NMR data of DNA and DNA·RNA hybrid duplexes containing phosphorothioate linkages in their DNA moiety revealed only minor proton chemical shift differences

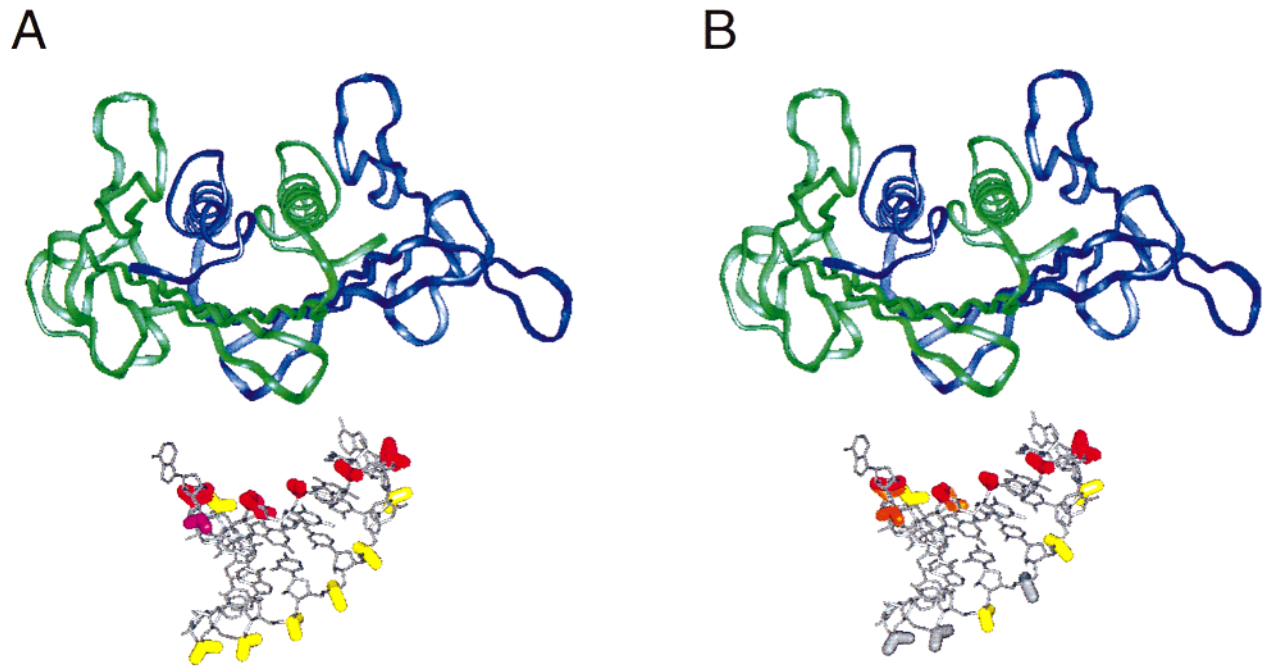

FIGURE 3: (A) Crystal structure of the complex with the phosphates that contact the protein in red and the phosphates that do not contact the protein in yellow. Position −9 that forms an intramolecular hydrogen bond is shown in purple. (B) The results of our modification study superimposed on the structure. Positions where a sulfur substitution either strengthens or weakens binding are shown in red, positions which show minor effects in orange, positions which can be modified without changing the binding affinity in yellow, and untested positions in gray.

adjacent to the sulfur atoms. It was concluded that, within the accuracy of the experimental method, the modified duplexes containing either R_P or S_P phosphorothioate linkages have the same structural features as the unmodified duplex (*42*−*45*). It is possible, however, that the phosphorothioate modification may alter the flexibility of the RNA backbone without changing the structure. This could therefore modulate the protein affinity without altering a direct contact made by that specific phosphate position. The data presented in this study make this possibility less likely, since the phosphorothioates at position −12 or −3, which lie either within or at the edge of the binding site but do not interact with the protein, have no effect on the affinity. Even modification of phosphate −9, which is directly involved in a structural rearrangement during complex formation by serving as the acceptor for an intramolecular hydrogen bond with the hydroxyl group at position −11 (*8*, *46*), decreases binding only 2-fold in the case of one isomer and not at all in the other one. Therefore, phosphorothioate linkages do not appear to greatly affect either the structure or the flexibility of the unbound RNA.

One of the questions that this study tried to answer is whether a phosphorothioate substitution would have different effects at phosphate positions that form ionic contacts with the protein than at positions that only form hydrogen bonds. One might expect, for example, that because hydrogen bonds are highly dependent on the distance and bond angle of the interacting atoms they would be greatly affected by a phosphorothioate substitution. In contrast, since ionic contacts observed in MS2 always involve the relatively flexible lysine and arginine side chains, they may not be as sensitive to the change of the charge distribution in the phosphorothioate linkage as the side chain could readjust to accommodate the sulfur. However, among the five contacts involving charged amino acids and the four contacts involving uncharged amino acids present in the MS2 complex, there is no obvious correlation between the size of the effect and the type of interaction. The largest effect is seen at position −6 where the *pro*-R_P oxygen forms a single hydrogen bond with AsnB55 and the R_P phosphorothioate-containing RNA binds about 20-fold tighter than the unmodified RNA. However, similar introduction of a sulfur for three other nonbridging oxygens that serve as hydrogen bond acceptors with uncharged amino acids (TyrA85 and phosphate −5, and SerB52 and AsnB55 and phosphate −7) results in much smaller (2−6-fold) changes in the binding affinity, either strengthening or weakening the interaction. Introducing phosphorothioates at positions that form ionic contacts (phosphates −4, −7, −8, −10, −11, and −13) either reduces or increases binding between 2- and 5-fold, which is similar to the efforts seen with the hydrogen bonding sites. Therefore, phosphorothioate modification studies will be helpful in predicting the sites of protein−phosphate contacts, but it does not seem possible to conclude, on the basis of such biochemical data, which type of interaction is formed at that particular site.

Since the environment of each individual protein−phosphate contact is different, the corresponding thermodynamic contribution to the overall binding energy is expected to vary. It would be valuable to use the phosphorothioate data to obtain an estimate of individual contributions and to rationalize them in terms of the crystal structure of the wild-type RNA−protein complex. Such an analysis can be anticipated to be quite difficult for at least four reasons. First, substitution of a sulfur for an oxygen that interacts with the protein changes several parameters in the local region of the modification that may have offsetting effects on the binding affinity. While the greater length of the P−S bond and the poor ability of sulfur to act as a hydrogen bond acceptor are expected to be unfavorable in many cases, sulfur is more

hydrophobic and less solvated than oxygen which may permit an increase in the binding energy in other cases. The slightly larger size of the sulfur atom could also cause steric clashes that would reduce the binding affinity (*47*). Second, substitution of one nonbridging oxygen with sulfur is expected to influence the binding properties of the other nonbridging oxygen at this phosphate as well. Since sulfur receives the majority of the net charge, the oxygen will be less negative and the P═O bond length shorter. Again, this charge redistribution may either strengthen or weaken the interaction. Third, in the MS2 system many of the protein−phosphate contacts are quite complicated. Often, several amino acids contact the same phosphate (as seen at position −7), or the same amino acid interacts with two or three nonbridging oxygens of one phosphate or two adjacent phosphates (e.g., positions −8 and −10/−11). Therefore, introduction of a phosphorothioate can potentially disrupt several contacts. Finally, interpretation may be complicated by the fact that the hydrogen bonding network and side chain positions could potentially rearrange to accommodate the phosphorothioate substitution just as structural rearrangements in response to specific mutations have been seen in crystal structures of several mutants of T4 lysozyme (*48*).

Despite these caveats, the phosphorothioate data can be rationalized on the basis of the crystal structure when the protein−phosphate contact is relatively isolated. One example is the hydrogen bond that is formed between the *pro*-R_P oxygen at position −5 and the hydroxyl group of TyrA85 (Figure 4A). Substitution of either nonbridging oxygen with sulfur reduces binding about 2−3-fold each, suggesting that in both cases the hydrogen bond formed with TyrA85 is partially disrupted. If it is assumed that the modified RNAs bind the protein in the same way as the unmodified one, different molecular explanations for the reduced affinity are needed for each isomer. In the R_P isomer, the lower binding affinity may be due to the fact that sulfur is a poorer hydrogen bond acceptor than oxygen. In the S_P isomer, the hydrogen bond acceptor has changed in several ways that may lower the free energy of binding. First, the uncharged oxygen is a weaker hydrogen bond acceptor than the partially charged oxygen in the unmodified RNA. Furthermore, the P═O distance is shorter, and therefore, the hydrogen bond may be somewhat lengthened in the S_P phosphorothioate. Thus, in this case, reasonable explanations can be given for why both phosphorothioate substitutions reduce the binding affinity.

When both nonbridging oxygens at one position are involved in a protein−phosphate contact, interpretation of the phosphorothioate substitution data becomes more difficult. This is illustrated by the two examples where an arginine in the MS2 coat protein contacts a phosphate in the RNA target. One of the amino groups of Arg49 in the A subunit of the MS2 coat protein dimer interacts with both oxygens of phosphate −13 (Figure 4B), while both the imino group and one of the amino groups of Arg49 in the B subunit each interact with one of the two oxygens at position −8, forming a "fork" (Figure 4C). Despite the fact that both bonds formed at position −13 are very similar in length and geometry, the effects of the two phosphorothioate isomers on binding are nevertheless different. Substitution of the *pro*-S_P oxygen leads to a 2-fold stronger binding, while substitutions of the *pro*-R_P oxygen leaves binding affinity unchanged. As the arginine side chain is flexible and this phosphate position lies at the very 5′-terminus of the hairpin, it seems possible that slight structural rearrangements of the protein or RNA may occur to accommodate the sulfur, making the effects small and less predictable. Interestingly, in a complex structure of the MS2 coat protein and an RNA hairpin with an extra two base pairs extending the lower stem (*49*), the distance between ArgA49 and phosphate −13 was greater than 4 Å and therefore very long for a salt bridge, implying that this interaction is not crucial for complex formation.

Introduction of either phosphorothioate isomer for phosphate −8 changes binding affinity to a greater extent than substitution at position −13 and in opposite directions for the two stereoisomers. The R_P isomer has a 3-fold decreased affinity for the coat protein, and the S_P isomer binds 4-fold tighter. If one assumes that the "fork" interaction and nearby GluB89 do not allow the side chain of ArgB49 to rearrange, these results can be somewhat understood from the structure. The S_P isomer should allow the formation of a strong ionic contact between the charged sulfur and the imino group, while the oxygen seems to be ideally positioned for a hydrogen bond with the amino group. In contrast, in the R_P isomer, the nonbridging oxygen is hydrogen bonding with the imino group, which is a weaker hydrogen bond donor than the terminal amino groups, presumably reducing the binding energy of the complex. However, it should be emphasized that this is just one possible explanation and other factors may contribute to the reported differences seen in the binding affinities.

When a complex set of hydrogen bonds is formed between the protein and the RNA backbone, it does not seem possible to rationalize the substitution data based on the structure of the wild-type complex. This is illustrated by the multidentate interaction that LysB61 forms with phosphates at positions −10 and −11 (Figure 4D). According to the crystal structure, the distances between the amino group of LysB61 and the *pro*-S_P oxygens at positions −10 and −11 are very similar (2.8 and 3.0 Å), and therefore, the two contacts would be expected to contribute similarly to the binding affinity. In contrast, the *pro*-R_P oxygen at position −10 is 3.6 Å from the amino group of LysB61, and the *pro*-R_P oxygen at position −11 is even further away (4.5 Å). Modification of phosphates −10 and −11 affects complex stability very differently. Introducing the sulfur into the *pro*-R_P position at phosphate −11 weakens binding slightly, presumably because the S_P oxygen is now uncharged and, therefore, makes a weaker bond with LysB61. Surprisingly, the S_P phosphorothioate at position −11 has no effect on binding. This may be explained by the offsetting effects of a weaker hydrogen bond to a sulfur acceptor but a stronger electrostatic interaction with the more negative sulfur. However, using a similar rationale for position −10, it is hard to explain why the R_P phosphorothioate strengthens binding 3-fold and the S_P phosphorothioate weakens binding 5-fold. The opposite would have been expected from the change in charge distribution caused by the sulfur and the distances of the two nonbridging oxygens from the amino group in the crystal structure of the unmodified complex. This seems to imply that other factors such as the hydrophobicity and electronegativity of the local environment or small local structural rearrangements also influence the outcome of a phosphorothioate modification.

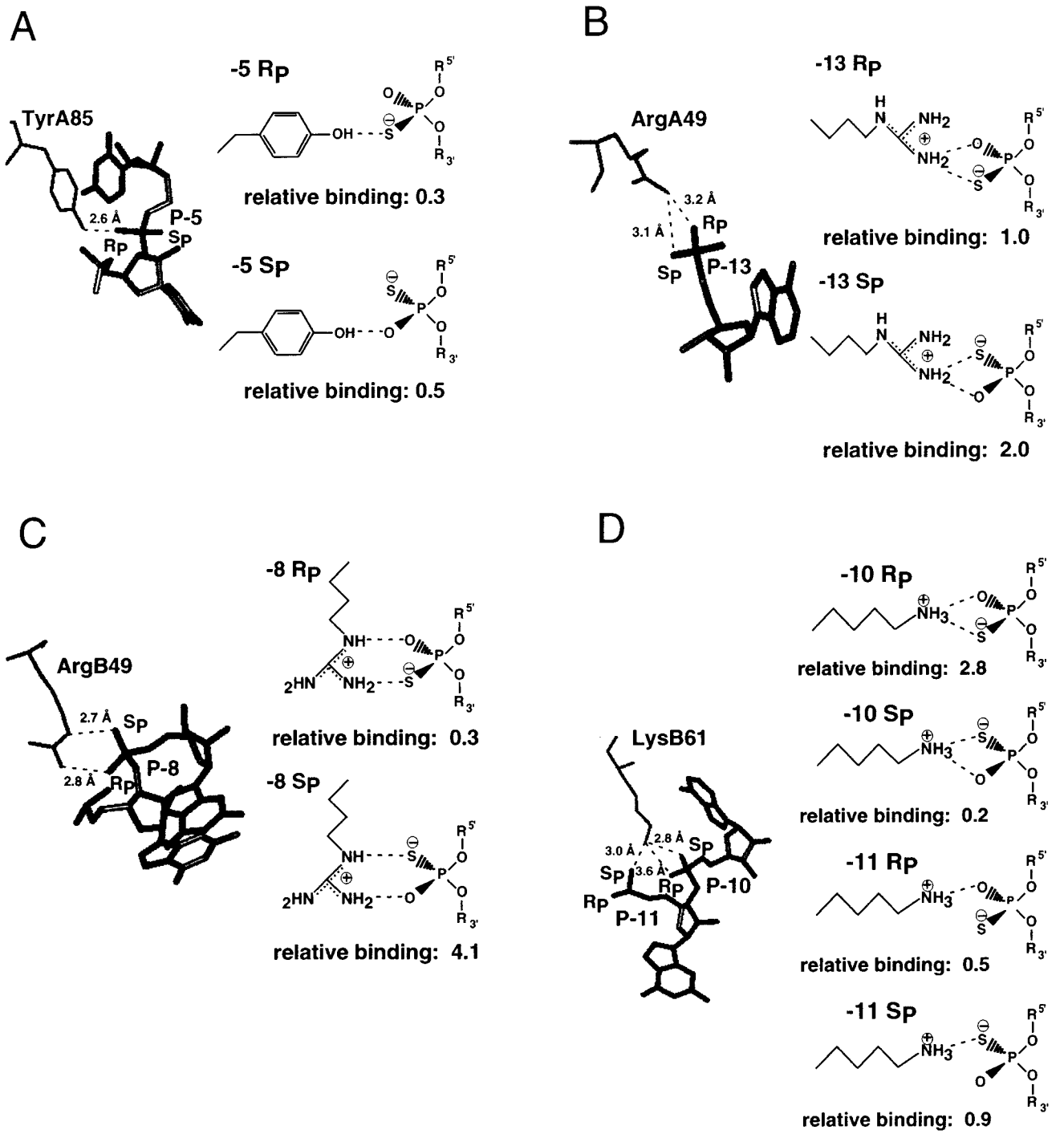

FIGURE 4: Details of the interactions and modification results of the phosphates at positions −5 (A), −13 (B), −8 (C), and −10/−11 (D). The relative binding is an average of all three data sets shown in Table 1. For better comparison, the two nonbridging oxygens (O1 and O2) as identified by the crystal structure are labeled as the corresponding phosphorothioate isomers R_P and S_P.

This is the first time that such a phosphorothioate modification "walk" has been applied to a protein−RNA complex for which a cocrystal structure is available. However, the data can be compared with a similar study in which the interaction of *Eco*RV restriction endonuclease with a double-stranded DNA containing its GATATC recognition sequence was probed (*11*). In this study, phosphorothioates were introduced at nine positions within the DNA binding site that were revealed by the available crystal structure of the complex (*50*, *51*) and the stereoisomers separated. The steady state parameters for DNA cleavage were obtained at those positions where cleavage could be observed. Although it is possible that a change in K_M due to phosphorothioate substitution does not directly reflect changes in binding affinity, a correlation between K_M and K_d was observed in another study (*52*). The effects of phosphorothioate substitutions on K_M in *Eco*RV are very similar in magnitude to those seen in the MS2 coat protein system. For most positions, modification seems to strengthen DNA binding slightly, but for three isomers, it seems to weaken binding. In contrast to the MS2 system, the correlation between the modification data and the protein−phosphate contacts observed in the crystal structure is not very good for *Eco*RV. For example, phosphate −3 does not interact with the protein, but the K_M

of one of the phosphorothioates was changed 5-fold. Furthermore, both isomers at position −2 affect K_M less than 2-fold, although this phosphate interacts with several amino acids. The poor correlation between modification data and the cocrystal structure in *Eco*RV is likely due to several reasons. First, the number and complexity of interactions with a given phosphate are much greater in the *Eco*RV complex than in the MS2 complex, increasing the opportunity for offsetting effects. Second, unlike MS2 coat protein, *Eco*RV requires at least two divalent metal ions for specific DNA binding and subsequent cleavage. Part of the effect of phosphorothioate substitution at position −2 could be due to the binding of a divalent metal ion rather than due to a specific amino acid contact (*11*). Other phosphates may be similarly involved in metal ion binding. Finally, the crystal structure might not represent the Michaelis complex, as the substrate is not cleaved in the crystal. Some rearrangement of the protein−phosphate contacts may have to occur to allow for DNA cleavage (*51*).

While most of the protein−phosphate contacts in the *Eco*RV system are different in detail from those of MS2, there is one case where the same contact is observed in both systems. DNA phosphate −5 forms an ionic contact with ArgB229 of the endonuclease that is identical to the MS2 contact of ArgA49 with phosphate −13 depicted in Figure 4B. Interestingly, the response to phosphorothioate substitutions at these positions is quite different in the two systems. The R_P phosphorothioate at position −5 of the *Eco*RV recognition sequence seems to strengthen binding 4-fold, while the S_P phosphorothioate leaves the K_M unchanged. Both effects are therefore opposite to what is seen for position −13 in the MS2 hairpin, where the S_P phosphorothioate strengthens the binding 2-fold while the R_P phosphorothiaote has no effect. This not only suggests that the local environment around a contact contributes to the effect of a phosphorothioate substitution and that some local rearrangements of the protein may occur but also underscores the difficulty of rationalizing binding effects at the molecular level.

Protein−phosphate contacts have been shown to contribute to both the specificity and affinity of DNA−protein complex formation. To a first approximation, the structure of a DNA double helix can be considered to be uniform, presenting the phosphates in a similar spatial arrangement relatively independent of the sequence. Phosphate contacts with the protein therefore contribute to a large extent to the affinity of those interactions. This nonspecific binding is often used for the site recognition mechanism by allowing the protein to bind electrostatically to the DNA and then to slide along the strand to find its recognition site (*53*−*55*). However, specific DNA binding can also depend on the ease with which the protein can distort the DNA helix, for example, by introducing kinks or bending the backbone (*56*−*58*). This phenomenon, in which a protein senses the base sequence of DNA through the DNA's backbone conformation or its flexibility, is usually referred to as "indirect readout" (*59*). The same mechanism should also be very important for RNA−protein complexes, as the secondary structure of RNA in cells is both conformationally diverse and quite flexible. This study suggests that all the protein−phosphate contacts in the MS2 system contribute to the binding affinity, and that many phosphates are probably also involved in specificity. A clear case of indirect readout at phosphates −6 and −7 has been reported for the MS2 coat protein (*16*), and it will be interesting to see how widespread this phenomenon is in other RNA−protein systems.

As long as an accurate binding assay is available, it can be concluded from this work that phosphorothioate substitutions can be used to accurately identify protein−phosphate contacts in nucleic acid−protein complexes of unknown structure. Even when a crystal structure is available, it solely provides information about which phosphates interact with the protein, while only biochemical studies can also provide insight into the thermodynamic contribution of each individual contact to the overall binding affinity. Both hydrogen bonds and ionic contacts seem to be affected similarly by the substitution of a phosphorothioate for a phosphodiester at the site of interaction, increasing or reducing the free energy of binding by 0.4−1.6 kcal/mol. A larger data set will be helpful in better understanding the thermodynamic changes caused by the introduction of a sulfur at a given position.

ACKNOWLEDGMENT

We thank D. Wuttke, S. Schultz, and members of the Uhlenbeck laboratory for helpful advice and critical evaluation of the manuscript and L. Lilias and K. Valegård for providing coordinates for the refined complex crystal structures prior to publication. D.D. gratefully acknowledges R. Seckler for ongoing support and stimulating discussions.

REFERENCES

1. Jen-Jacobson, L. (1997) *Biopolymers 44*, 153−180.
2. Mattaj, I. W. (1993) *Cell 73*, 837−840.
3. Kenan, D. J., Query, C. C., and Keene, J. D. (1991) *Trends Biochem. Sci. 16*, 214−220.
4. Johansson, H. E., Liljas, L., and Uhlenbeck, O. C. (1997) *Semin. Virol. 8*, 176−185.
5. Witherell, G. W., Gott, J. M., and Uhlenbeck, O. C. (1991) *Prog. Nucleic Acid Res. Mol. Biol. 40*, 185−220.
6. Peabody, D. S. (1997) *Mol. Gen. Genet. 254*, 358−364.
7. Lago, H., Fonseca, S. A., Murray, J. B., Stonehouse, N. J., and Stockley, P. G. (1998) *Nucleic Acids Res. 26*, 1337−1344.
8. Valegård, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G., and Liljas, L. (1997) *J. Mol. Biol. 270*, 724−738.
9. Carey, J. C., and Uhlenbeck, O. C. (1983) *Biochemistry 22*, 2610−2615.
10. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Biochemistry 28*, 2849−2855.
11. Thorogood, H., Grasby, J. A., and Connolly, B. A. (1996) *J. Biol. Chem. 271*, 8855−8862.
12. Lesser, D. R., Grajkowski, A., Kurpiewski, M. R., Koziolkiewicz, M., Stec, W. J., and Jen-Jacobson, L. (1992) *J. Biol. Chem. 267*, 24810−24818.
13. Olsen, D. B., Kotzorek, G., and Eckstein, F. (1990) *Biochemistry 29*, 9546−9551.
14. Usman, N., Ogilvie, K. K., Jiang, M.-Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
15. Iyer, R. P., Phillips, L. R., Egan, W., Regan, J. B., and Beaucage, S. L. (1990) *J. Org. Chem. 55*, 4693−4699.
16. Johansson, H. E., Dertinger, D., LeCuyer, K. A., Behlen, L. S., Greef, C. H., and Uhlenbeck, O. C. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9244−9249.
17. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedle, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
18. Slim, G., and Gait, M. J. (1991) *Nucleic Acids Res. 19*, 1183−1188.

19. Moore, M. J., and Sharp, P. A. (1993) *Nature 365*, 364−368.
20. Burgers, P. M., and Eckstein, F. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4798−4800.
21. Loverix, S., Winquist, A., Stroemberg, R., and Steyaert, J. (1998) *Nat. Struct. Biol. 5*, 365−368.
22. Greef, C. H., Seeberger, P. H., and Caruthers, M. H. (1996) *Tetrahedron Lett. 37*, 4451−4454.
23. Bruce, A. G., and Uhlenbeck, O. C. (1978) *Nucleic Acids Res. 5*, 3665−3677.
24. England, T., and Uhlenbeck, O. C. (1978) *Nature 275*, 560−561.
25. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Donis-Keller, H., Maxam, A. M., and Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527−2538.
27. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1995) *Biochemistry 34*, 10600−10606.
28. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1996) *EMBO J. 15*, 6847−6853.
29. Carey, J. C., Cameron, V., deHaseth, P. L., and Uhlenbeck, O. C. (1983) *Biochemistry 22*, 2601−2610.
30. Beckett, D., and Uhlenbeck, O. C. (1988) *J. Mol. Biol. 204*, 927−938.
31. Lim, F., and Peabody, D. S. (1994) *Nucleic Acids Res. 22*, 3748−3752.
32. Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994) *Nature 371*, 623−626.
33. Lowary, P. T., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 10483−10493.
34. Szwajkajzer, D., and Carey, J. (1997) *Biopolymers 44*, 181−198.
35. Mayer, A. N., and Barany, F. (1994) *J. Biol. Chem. 269*, 29067−29076.
36. Vörtler, C. S., Fedorova, O., Persson, T., Kutzke, U., and Eckstein, F. (1998) *RNA 4*, 1444−1454.
37. Knöll, R., Blad, R., and Fürste, J. P. (1997) *RNA 3*, 132−140.
38. Frey, P. A., and Sammons, R. D. (1985) *Science 228*, 541−545.
39. Liang, C., and Allen, L. C. (1987) *J. Am. Chem. Soc. 109*, 6449−6453.
40. Saenger, W., and Eckstein, F. (1970) *J. Am. Chem. Soc. 92*, 4712−4718.
41. Cruse, W. B. T., Salisbury, S. A., Brown, T., Cosstick, R., Eckstein, F., and Kennard, O. (1986) *J. Mol. Biol. 192*, 891−905.
42. González, C., Stec, W., Reynolds, M. A., and James, T. L. (1995) *Biochemistry 34*, 4969−4982.
43. LaPlanche, L. A., James, T. L., Powell, C., Wilson, D. W., Uznanski, B., Stec, W. J., Summers, M. F., and Zon, G. (1986) *Nucleic Acids Res. 14*, 9081−9093.
44. González, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., and James, T. L. (1994) *Biochemistry 33*, 11062−11072.
45. Furrer, P., Billeci, T. M., Donati, A., Kojima, C., Karwowski, B., Sierzchala, A., Stec, W., and James, T. L. (1999) *J. Mol. Biol. 285*, 1609−1622.
46. Kerwood, D. J., and Borer, P. N. (1996) *Magn. Reson. Chem. 34*, S136−S146.
47. Brautigam, C. A., and Steitz, T. A. (1998) *J. Mol. Biol. 277*, 363−377.
48. Alber, T., Dao-pin, S., Wilson, K., Wozniak, J. A., Cook, S. P., and Matthews, B. W. (1987) *Nature 330*, 41−46.
49. Grahn, E., Stonehouse, N. J., Murray, J. B., Van Den Worm, S., Valegård, K., Fridborg, K., Stockley, P. G., and Liljas, L. (1999) *RNA 5*, 131−138.
50. Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., and Wilson, K. S. (1993) *EMBO J. 12*, 1781−1795.
51. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry 34*, 683−696.
52. Wenz, C., Jeltsch, A., and Pingoud, A. (1996) *J. Biol. Chem. 271*, 5565−5573.
53. Ohlendorf, D. H., and Matthew, J. B. (1985) *Adv. Biophys. 20*, 137−151.
54. Jeltsch, A., and Pingoud, A. (1998) *Biochemistry 37*, 2160−2169.
55. Berg, O. G., and Ehrenberg, M. (1982) *Biophys. Chem. 15*, 41−51.
56. Anderson, J. E., Ptashne, M., and Harrison, S. C. (1987) *Nature 326*, 846−852.
57. Liu-Johnson, H.-N., Gartenberg, M. R., and Crothers, D. M. (1986) *Cell 47*, 995−1005.
58. Kumar, S., Duan, Y., Kollman, P. A., and Rosenberg, J. M. (1994) *J. Biomol. Struct. Dyn. 12*, 487−525.
59. Otwinowski, S., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., and Sigler, P. B. (1988) *Nature 335*, 321−329.

BI991769V