Structure Determination and Binding Kinetics of a DNA Aptamer−Argininamide Complex[†,‡]

Stephanie A. Robertson,[§] Kazuo Harada,[∥] Alan D. Frankel,[∥] and David E. Wemmer*[,§]

Department of Chemistry, University of California, Berkeley, California 94720-1460, and Department of Biochemistry & Biophysics, University of California, San Francisco, California 94143

Received June 30, 1999; Revised Manuscript Received October 25, 1999

ABSTRACT: The structure of a DNA aptamer, which was selected for specific binding to arginine, was determined using NMR spectroscopy. The sequence forms a hairpin loop, with residues important for binding occurring in the loop region. Binding of argininamide induces formation of one Watson−Crick and two non-Watson−Crick base pairs, which facilitate generation of a binding pocket. The specificity for arginine seems to arise from contacts between the guanidino end of the arginine and phosphates, with atoms positioned by the shape of the pocket. Complex binding kinetics are observed suggesting that there is a slow interconversion of two forms of the DNA, which have different binding affinities. These data provide information on the process of adaptive recognition of a ligand by an aptamer.

Molecular recognition between nucleic acids and their cognate substrates is essential in numerous biological processes, including transcriptional and translational events. This recognition has been shown to often rely strongly on secondary and tertiary structural elements of the nucleic acids. It is well-established that nucleic acids are not only biologically functional in duplex form but also take on numerous other structural motifs that are capable of interaction with other molecules. An RNA secondary structure has been found to play important roles in translational regulation, protein synthesis, and mRNA splicing (*1*, *2*). Similarly, DNA hairpins, dumbbells, and cruciforms are known to play roles in genetic recombination and regulation (*3*−*6*). Determination of nucleic acid−substrate complex structures is an essential step in understanding the importance of the role nucleic acid structural features have in recognition and regulation in these complex biological processes.

Previously determined structures of nucleic acid complexes, in both duplex as well as noncanonical nucleic acid forms, include examples such as drug−DNA and drug−RNA complexes (*7*−*9*) as well as protein−DNA and protein−RNA complexes (*10*, *11*). These structures provide insight into important elements for molecular recognition by nucleic acids.

An understanding of nucleic acid recognition elements has also broadened as in vitro selection and selective evolution have been used to identify nucleic acid motifs with unique folding and ligand-binding properties, the resulting molecules being termed aptamers (*12*−*14*). The aptamers selected in this manner include RNAs that bind ATP (*15*), protein cofactor FMN (*16*, *17*), arginine and citrulline (*18*−*21*), and DNAs that bind thrombin (*22*) and arginine (*23*). These aptamers have exhibited binding affinities in the micromolar to nanomolar range and are able to discriminate between their target ligand and closely related molecules. These techniques of aptamer derivation have contributed to the understanding of protein−nucleic acid interactions, nucleic acid recognition elements, and investigations in nucleic acid drug development (*13*, *24*, *25*). Structure determination of a number of these aptamers has also lent insight into the folding capabilities and adaptive properties of nucleic acids in order to recognize their cognate substrates (*26*−*29*).

DNA aptamers specific for binding of arginine were discovered (*23*) that have secondary folding motifs unrelated to known arginine-binding RNA sequences (*18*−*21*). These DNA aptamers were specific for the guanidinium group and were selected under conditions identical to those used in identifying RNA aptamers that bind arginine in a manner similar to HIV-1 TAR RNA (*20*). Mutational and chemical modification studies showed that one of the DNA motifs isolated, a stem-loop structure, contained eight residues within the loop as well as two bases at the top of the stem that were crucial to binding of arginine.

Here, we present the structure of argininamide complexed to a DNA containing a 10 base loop. The complex contains several unusual features not seen in previously determined nucleic acid complexes, including three cross-loop base pairs that are not formed in the absence of argininamide. Two of these base pairs are non-Watson−Crick, and the third is formed by a residue that loops out of sequence to form a

[†] This work was supported by the Director, Office of Biological & Environmental Research, Office of Energy Research of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER75281 and the National Science Foundation, DMB 86-09305 and BBS 87-20134 to D.E.W, and by NIH grant GM 47478 to A.D.F.

[‡] Coordinates for the structures have been deposited into the PDB, file name 1DB6.

* Author contact information: Department of Chemistry, University of California, Berkeley, CA 94720-1460. Telephone: (510) 486-4318. Fax: (510) 486-6059. E-mail: dewemmer@LBL.gov.

10.1021/bi9915061 CCC: $19.00 © 2000 American Chemical Society
Published on Web 01/13/2000

Watson−Crick base pair. The complex exhibits intermediate exchange binding kinetics at low ligand/DNA ratios, which have been analyzed using NMR[1] line shape theory to provide information on kinetic parameters. The structure of a closely related sequence has been published previously (*36*), and while both show envelopment of argininamide within a binding pocket, the structures determined differ in significant ways. The structure of a second argininamide-binding DNA aptamer (*23*) has been reported (*30*), which again illustrates the importance of non-Watson−Crick interactions in formation of ligand-binding sites.

MATERIALS AND METHODS

Oligonucleotides were synthesized at 1 or 10 μM scale on an Applied Biosystems, Inc. (Foster City, CA) automated synthesizer using solid phase β-cyanoethylphosphoramidite chemistry and were HPLC purified. NMR samples were prepared by dissolving dried oligonucleotide to give 1.44 or 0.8 mM solutions, in either 90% H_2O/10% D_2O or 99.99% D_2O containing 10 mM sodium phosphate buffer, pH 6.2, 0.05 mM EDTA, and 0−80 mM NaCl. Samples that were used in data collection for the purposes of structure calculations also contained approximately eight to 10 equiv of argininamide. The sequences used in NMR studies were

$$5'\text{-}C_1G_2A_3C_4C_5\mathit{A}_6\mathit{A}_7\mathit{C}_8\mathit{G}_9\mathit{T}_{10}X_{11}X_{12}\mathit{C}_{13}\mathit{G}_{14}\mathit{C}_{15}\mathit{C}_{16}\text{-}\mathit{T}_{17}G_{18}G_{19}T_{20}C_{21}G_{22}\text{-}3'$$

where $X_{11}X_{12}$ is AT, GT, or AG and letters in italics represent residues that were shown to be critical to binding of arginine by chemical modification and mutagenesis studies (*23*). $X_{11}X_{12}$ are GT in the structures calculated.

All NMR data were collected on either a Bruker AMX-600 or a GE GN/Ω500 spectrometer. ^{1}H Nuclear Overhauser Effect Spectroscopy (NOESY) spectra were collected in H_2O with the standard pulse sequence in which the last pulse was replaced by a 11 spin−echo pulse sequence (*31*). Spectra were collected at 5 and 8 °C with 125 and 150 ms mixing times, and 512 complex points in t_1 and 2048 in t_2. ^{1}H NOESY spectra were collected in D_2O at 20 °C using 60, 100, and 200 ms mixing times, as well as at 15 and 25 °C using 200 ms mixing times, and 512 by 1024 complex points. The removal of the residual HOD signal was accomplished using presaturation. ^{1}H TOCSY (80 ms mixing time) spectra and ^{1}H COSY spectra were collected at 20 °C with 512 by 1024 or 2048 complex points. One-dimensional experiments used in kinetic analyses were collected at 5, 20, and 25 °C with 8192 complex points. All spectra were processed using the program FELIX (Molecular Simulations, Inc.).

Distance restraints were obtained using relaxation rate matrix calculations (*32*) with experimental NOE cross-peak volumes calibrated to the cytosine H5−H6 and thymidine H6-methyl cross-peaks. Input structures used in relaxation rate matrix calculations were obtained from restrained molecular dynamic calculations in the program DISCOVER (Molecular Simulations, Inc.) using strong, medium, and weak NOE intensities (1.8−3, 1.8−4, 1.8−5 Å) to fold the initially randomized DNA oligomer in a procedure analogous to the structure calculation methodology described below. Several starting structures calculated this way, as well as correlation times of 2, 3, and 4 ns, were used in the relaxation rate matrix calculations to obtain final distance constraints used in structure calculations.

The final structures were calculated using the program DISCOVER (Molecular Simulations, Inc.). Since NOE patterns in the stem region of the DNA (bases 1−6 and 17−22) were consistent with canonical B-form DNA, an initial structure containing a standard B-form stem and randomized loop nucleotides was used. The argininamide was initially positioned >20 Å away from the DNA. Structures were minimized using steepest descents, followed by a conjugate gradient minimization to ensure good covalent geometry prior to restrained molecular dynamics. Restrained molecular dynamics were initiated at 600 K for 4 ps, then the temperature was reduced to 100 K over 10 ps. The temperature was then reduced from 100 to 0 K over 10 ps, and the final structures were conjugate gradient minimized. Thirty final structures were calculated and the mean and standard deviation for each of the energy terms were calculated. Structures that contained any energy terms greater than two times the standard deviation plus the mean were considered suspect and were dropped (two structures). Any remaining structures that had at least one constraint violation greater than 0.5 Å were also dropped (seven structures). The remaining converged structures were ranked based on constraint energies, total energy, and number of violations. On the basis of calculations (3000 permutations) of the RMS and standard deviation as the number of structures is increased to the total number of structures (21 total structures), 10 structures were determined to be representative of the family of structures based on the criterion that further addition of structures to the family had no effect on the RMS standard deviation.

Line shape analyses were performed using modified versions of the Bloch equations (*33*) and were calculated in Mathematica (Wolfram Research).

RESULTS

Complex Formation and Kinetics. To obtain satisfactory data for a structure determination of the DNA−argininamide complex, it was necessary to find conditions under which the complex is the predominant species. Using the dissociation constant of 10^{-4} M calculated from CD measurements (*23*), a 2 mM DNA sample would be expected to require approximately 2 equiv of ligand to be 96% complexed based on simple equilibrium binding behavior. The appearance of the 1D NMR spectra of samples containing a 2:1 argininamide/DNA ratio seemed reasonably consistent with this expectation. However, the NOESY spectrum of this complex showed significant effects from conformational exchange of the DNA, as indicated by substantial line broadening for a number of loop residue hydrogens. For example, the T10 methyl hydrogens were broadened from 18 Hz in the free DNA to approximately 60−80 Hz (which appear to be at least three peaks in intermediate exchange) in a 2:1 argininamide−DNA complex (1.6 mM DNA, 25 °C), and sugar

[1] Abbreviations: NMR, nuclear magnetic resonance; HPLC, high performance liquid chromatography; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; RMS, root-mean-squared; CD, circular dichroism.

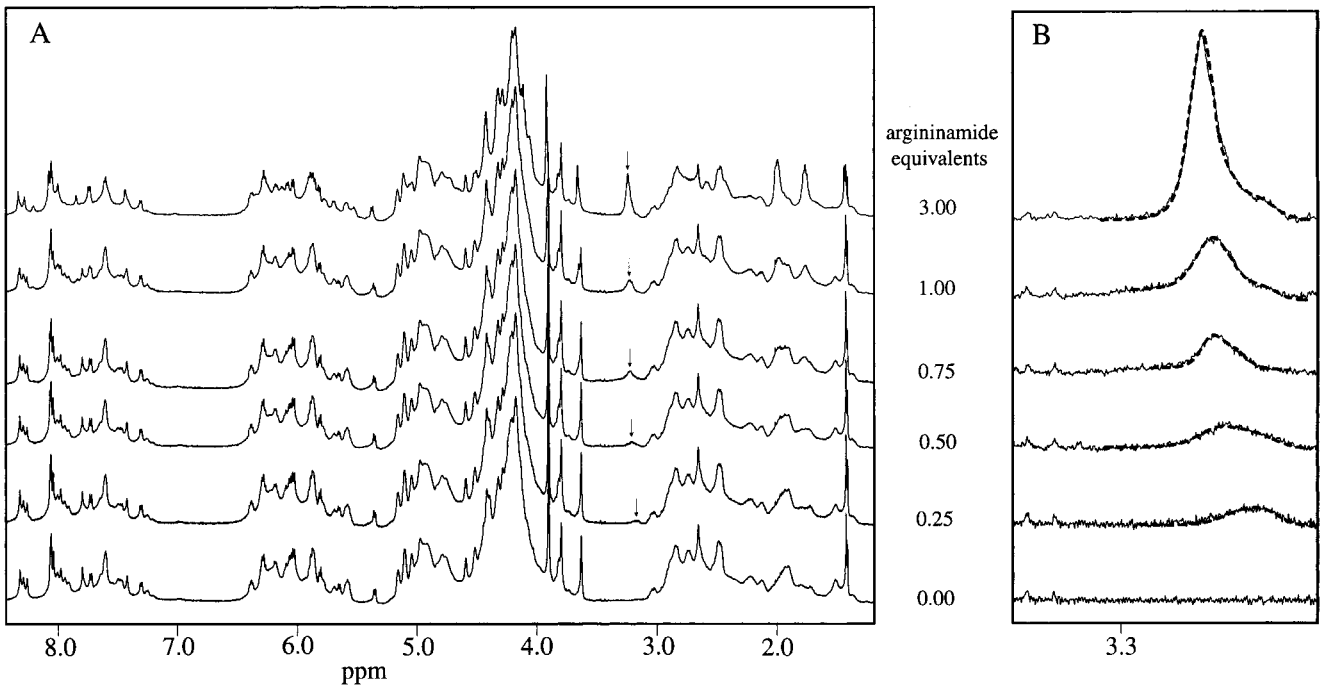

FIGURE 1: Proton NMR spectra taken in D_2O solution, 25 °C, are shown at several points in a titration of the DNA aptamer 5′-CGACCAACGTGTCGCCTGGTCG-3′ with argininamide, with the mole ratios of argininamide to DNA indicated. (A) Shows the full spectrum with the position of the argininamide δH resonances indicated by an arrow. (B) Shows an expansion of the argininamide δH resonance region, with dashed curves representing spectra simulated using the exchange model discussed.

protons for C16 were virtually undetectable. Addition of excess argininamide (5+ equiv) significantly sharpened these lines (the T10 methyl resonance sharpens to a single line of approximately 30 Hz).

The need for addition of excess ligand to form a fully complexed DNA suggested that the aptamer might contain more than one argininamide-binding site or that complicated kinetics might be involved. Because the complex is in intermediate to fast exchange, only one set of lines is seen for both the DNA and ligand and thus information on the order of complexation is not immediately apparent. However, the aptamer was selected under conditions designed to yield molecules containing a single binding site by chromatographing DNA using arginine resins with varying arginine densities (covering 3 orders of magnitude (*23*)). CD measurements further indicated that only one conformational change occurred in the DNA as ligand was added. While the selection methods and CD data suggested that the DNA had only one argininamide-binding site, careful titration of argininamide showed unanticipated binding behavior (based on the argininamide NMR lines) that was not consistent with simple intermediate to fast binding behavior for a complex with a dissociation constant of 10^{-4} M.

At low argininamide/DNA ratios, the argininamide lines were broadened to the extent that they were almost undetectable. With the measured dissociation constant of 10^{-4} M, 94% of ligand is expected to be bound at 0.25 equiv in a 2 mM DNA sample, and a relatively sharp argininamide line is expected at the bound state chemical shift.

NOESY spectra of the DNA in the absence of argininamide showed some exchange cross-peaks near the diagonal, indicating two conformations of the DNA that slowly interconvert. NOEs are clear for the major form, but cross-peaks for the second form are missing, indicating that it is present at relatively low concentration. One possible explanation for the two conformers is that the out of sequence stacking (C16 below G14) requires major rearrangement of bases and hence is slow. An alternative is that the anti and syn base conformations at G14 interconvert slowly, with the syn form in the final argininamide-bound structure, but anti was preferred in the free DNA. It is likely that these two forms, regardless of the precise difference, have different binding affinities for argininamide. Line shapes for arginine were calculated using the McConnell equations (*33*), assuming that binding occurred only to one form, and kinetic parameters were determined; k_{off} was estimated to be approximately 120 s^{-1} and the exchange rate at low argininamide/DNA ratios was approximately 240 s^{-1}. The bound argininamide frequency was calculated to be 50 Hz lower than the free argininamide for the δH resonance at 500 MHz with a K_d of 1.2 mM. Figure 1 shows an example of the titration behavior seen, and curve fits for the argininamide δ-proton resonance. While the calculated fits are reasonable, the derived binding constant is not consistent with that determined by other methods, which is an indication that the binding model used is oversimplified. A full model would require different binding constants for the two states, as well as kinetic parameters for each, and would be undetermined with the presently available data, necessitating other independent measurements.

Assignments. With the exception of some of the H5′ resonances, all of the assignments of the hairpin were readily obtained using standard procedures for DNA assignment (*34*) and are tabulated in the Supplementary Information. Expected sequential NOEs consistent with B-form duplex DNA were seen for residues 1−10 and 17−22. Some connectivities were seen between residues 10 and 11, and between residues 12 and 13, which allowed for eventual assignment of these residues. However, the region between residues 14 and 17 had a number of NOEs that were unusual and made

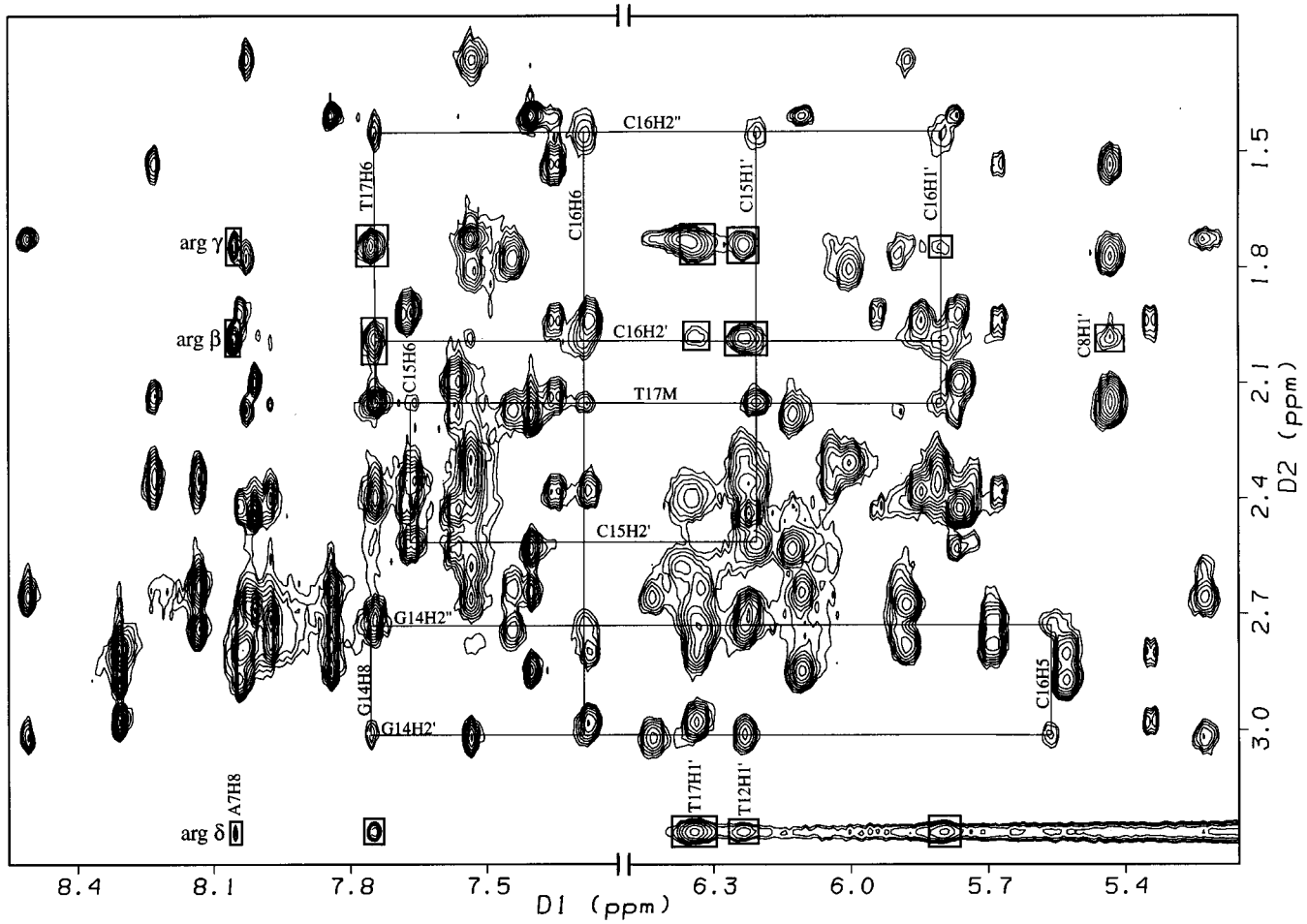

FIGURE 2: Parts of a NOESY spectrum of the DNA aptamer 5′-CGACCAACGTATCGCCTGGTCG-3′ in the presence of excess argininamide are shown, 5 equiv argininamide, ca. 2 mM DNA, 25 °C, 200 ms mixing time. The regions plotted show cross-peaks between aromatic and H1′ DNA protons and the H2′, H2″ of DNA as well as argininamide protons. Nonsequential NOEs between C16 and G14, and between T17 and C15 are highlighted by lines. Contacts between DNA and argininamide are indicated by boxes around the cross-peaks.

identification of C15 and C16 ambiguous. 5-methylcytosine (5mC) was substituted at position 16 and clearly indicated that C16 loops out of sequence to make a number of nonsequential contacts. C16H6 and -H5 (or H6 and methyl in the 5mC variant) make contacts with G14H2′ and -H2″ as well as with C15H1′. C15H1′ makes contacts with both its own H2′ and H2″ as well as with C16H2′/H2″ and with the T17 methyl group. The contact from C15H1′ to the H2′/H2″ of C16 is notably unusual, since in canonical DNA interresidue H1′ to H2′/H2″ distances are generally well out of NOE range (>5 Å), and normally only intra-sugar H1′ to H2′/H2″ contacts are seen. T17H6 also has contacts to C16H2′ and -H2″. Figure 2 shows the aromatic and H1′ to H2′/H2″ region of a NOESY spectrum for the complex, and highlights a number of the contacts in the loop region that define the unusual behavior of this region.

The 1D H_2O spectra show the appearance of new imino resonances upon addition of argininamide that can be assigned to several residues in the loop that are indicative of the new loop base pairings formed upon ligand complexation. Complete analysis of NOESY spectra, including spectra of the sequence containing inosine substituted at position 14 and 5mC substituted independently at positions 15 and 16, showed that loop residues form an argininamide bound conformation that includes extensive cross-loop connectivities, establishing the formation of three new base pairs. C16 base pairs in a normal Watson−Crick fashion with G9, indicated by NOEs from the G9 imino to C16 amino protons. NOEs from the C15 amino protons to A7H8 and amino resonances establish reverse Hoogsteen base pair formation between these residues. The sharpness and chemical shifts of the A7 and C15 amino protons are also indicative of hydrogen bond formation. An exceptionally strong NOE between G14H8 and its own H1′ establishes that G14 is in a syn conformation. G14 forms a reverse wobble base pair with T10, established by observation of an NOE cross-peak from the T10 imino to G14 imino and by chemical shifts for these iminos typical of G−T base pairs. The G14 imino proton also shows a cross-loop NOE to the G9 imino proton. Figure 3A shows part of the loop region from one representative calculated structure that highlights the unusual interactions made by bases 7−10 and 14−16.

Numerous contacts are seen between the argininamide nonexchangeable side chain protons and the DNA. As expected, most of the contacts are to important residues as well as to one of the two important bases at the top of the stem. An additional contact to the sugar of nonessential residue 12 is also seen. While the identity of residue 12 is not critical to argininamide binding, a small increase in binding affinity was seen with T over G (*23*), possibly reflecting steric preferences at this position. No direct NOEs were seen between any of the exchangeable protons on the

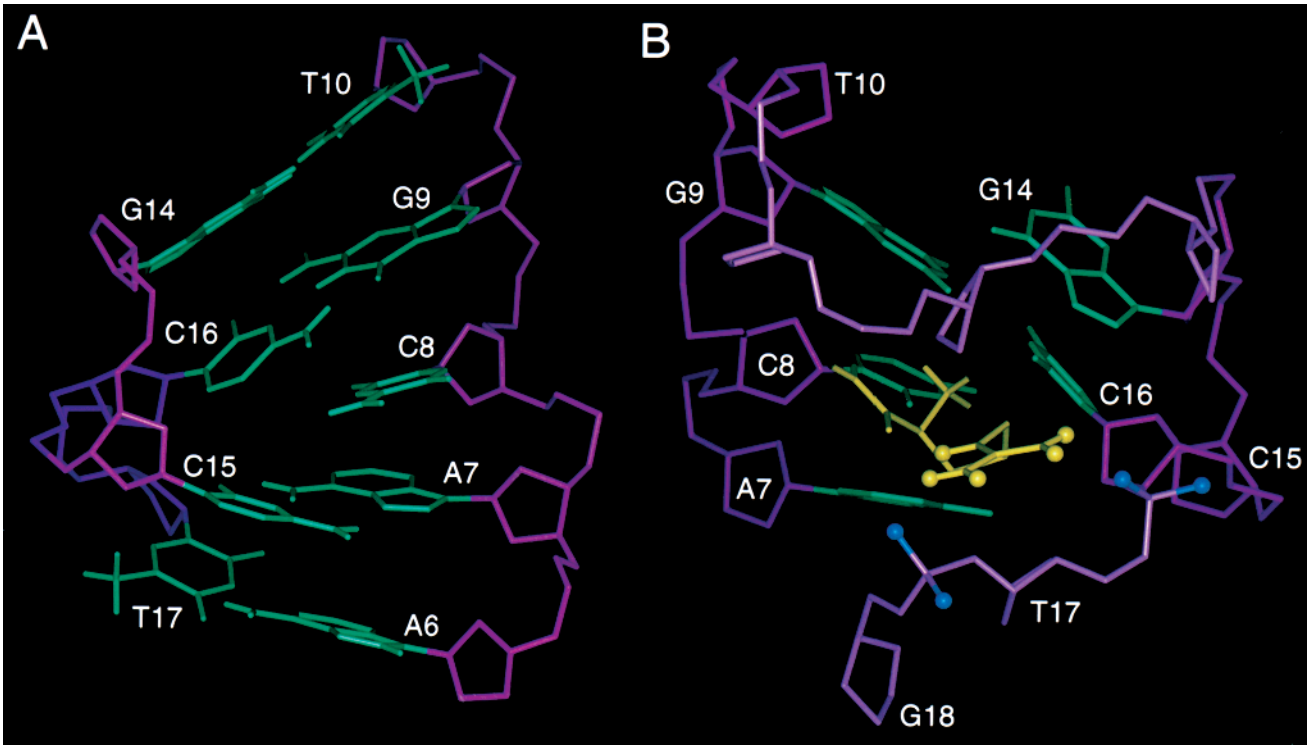

FIGURE 3: The loop region of the argininamide−DNA complex, residues A6−T10 and G14−T17, from a single representative structure is shown. (A) Bases are shown with imino, amino, and methyl hydrogens, while other hydrogens and all phosphate oxygens are omitted for clarity. Bases are green and backbone atoms are purple. (B) A different view of the same region is shown. Argininamide amide, amino, guanidinium hydrogens, the phosphate oxygens of T17 and G18 are displayed, with the guanidinium protons and phosphate oxygens indicated by spheres. The argininamide is shown in yellow. Only the backbone atoms of G11, T12, and C13 are shown but are not labeled.

Table 1: Intermolecular NOE Contacts in the Argininamide−DNA Aptamer Complex

argininamide atom	direct DNA contacts
Hα	A7H2, C8H1′
Hβ	A7H2, T12H1′, G14H8, T17H6, T17H3
Hγ	A7H2, G14H8, G9H1, T12H1′, T17H1′, T17H3′, T17H4′, C16H1′, C16H3′
Hδ	G9H22, G9H1, C16H1′, C16H4′, T17H1′, T17H3′, T17H4′

Table 2: NMR Refinement Statistics for the Argininamide-DNA Aptamer Complex

NMR distance restraints in the entire complex	
total number distance restraints	433
intraresidue restraints	244
interresidue restraints	143
hydrogen bond restraints	23
DNA−argininamide restraints	23
number of NOE violations >0.3 Å	5 ± 1
pairwise rmsd (Å) among 10 refined structures	
rmsd (Å) - entire complex	1.32 ± 0.35
rmsd (Å) - complex core (A6−T10, T12, G14−T17, G18 sugar)	0.92 ± 0.27

argininamide and the DNA. A summary of intermolecular NOE contacts is shown in Table 1.

DNA Structure. Table 2 summarizes NMR refinement statistics for the argininamide−DNA aptamer complex. The stem of the hairpin is in a canonical B-form conformation, showing the expected NOE patterns. As none of the residues in the stem is implicated in binding, with the exception of the closing AT base pair at the base of the loop, this region was not expected to show unusual characteristics. In contrast, the loop region shows many interesting features. Figure 4 presents a stereoview of 10 superimposed structures of the loop region.

The loop bases show almost normal B-form stacking from A6−T10 (Figure 3). A slight kink is seen in the backbone at C8, which remains the only residue unpaired in the stretch from A6−T10. Residues G14, C15, and C16 make the critical pairings across the loop with T10, A7, and G9, respectively (Figure 3A). A very unusual backbone conformation is observed, with C16 looped out of sequence to form the Watson−Crick G9−C16 basepair sandwiched between the C15−A7 reverse Hoogsteen and the G14−T10 reverse wobble. This configuation of bases results in a large bend on one side of the loop to form part of the argininamide-binding pocket. Figure 5 shows one representative structure where the deep bend in the loop and formation of the binding pocket are apparent. Residues in the top of the loop, particularly G11 and C13, are poorly defined in the structures, with few NOE connectivities observed between T10 and G11, and between T12 and C13. The location of T12 is slightly better defined due to interactions with the ligand, and this base helps pull the top of the loop down around the ligand to close the argininamide-binding pocket (Figure 3B).

DISCUSSION

Structure determinations of aptamer complexes have contributed to the understanding of protein−nucleic acid interactions, nucleic acid recognition elements, and of the folding capabilities and adaptive properties of nucleic acids

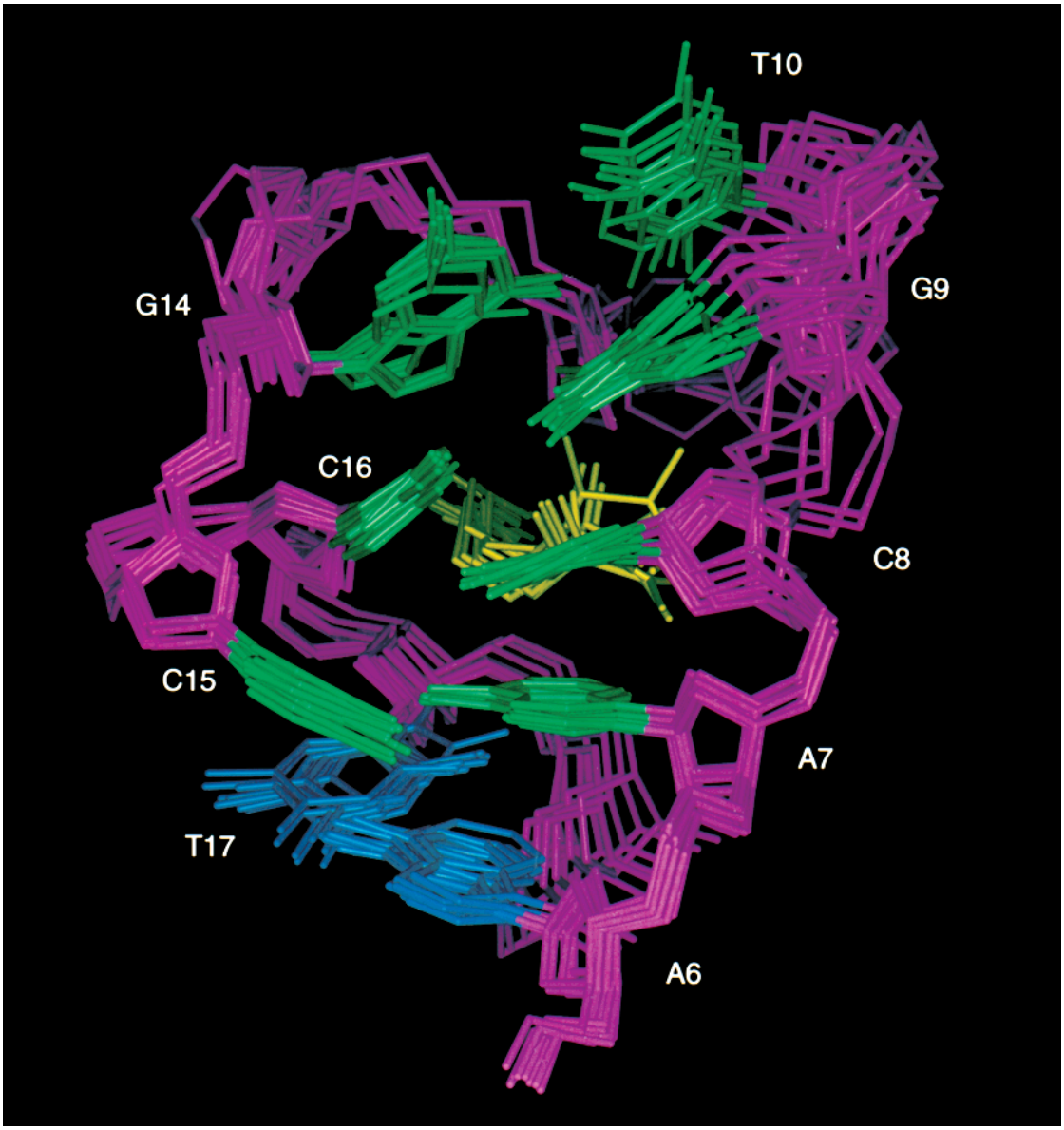

FIGURE 4: A superposition of 10 representative structure of the argininamide−DNA complex is shown. The superposition was done using heavy atoms for residues 6−17, the sugar and backbone atoms of 18 and the argininamide. Phosphate oxygens and bases of residues 11−13 and 18 are not shown for clarity. Backbone atoms are purple, stem bases 6 and 17 are blue, loop bases 7−10 and 14−17 are green, and argininamide is yellow.

in order to recognize their cognate substrates (*25*−*28*). These structural studies have also contributed greatly to development of drugs that target DNA (*35*). The structure of the argininamide−DNA complex presented here contributes to the understanding of how molecules adapt to recognize their substrate.

Binding of argininamide is critical in the formation of the structured loop conformation of this DNA aptamer. The binding pocket is formed by unusual interactions that include the C16 residue looping out of sequence to form a Watson−Crick basepair across the loop with G9, and formation of reverse G14−T10 wobble and reverse A7−C15 Hoogsteen pairs. These interactions are definitive elements in forming the deep kink in the DNA structure that encloses the argininamide ligand. The argininamide is extended in the binding pocket so that the guanidinium group hydrogens are above and between the T17 and G18 backbone phosphates. These probably form hydrogen bonds with these phosphate groups, consistent with chemical interference experiments (*23*), which indicate that at least the T17 phosphate is critical in binding of the argininamide ligand. The amino and amide groups are directed toward the A7−C8 step across the binding pocket from the guanidinium group and likely make hydrogen bond contacts to the C8 base, based on proximity in the structures. It is reasonable to conclude that the amino group has electrostatic interactions with the phosphate groups in the vicinity as well. The presence of such hydrogen bonds and electrostatic interactions is supported by the reduction in binding constant over an order of magnitude upon the loss of the C-terminal argininamide amino and/or the amide group (*23*). This significant reduction in binding constant emphasizes the importance of the amide and amino groups in stabilizing the bound loop conformation.

Both the binding pocket of the structure presented here, and the interaction of the guanidinium group are quite different from the structure of a very similar aptamer reported during the course of this work (*36*). This previous structure aligns the guanidinium group along the Watson−Crick edge

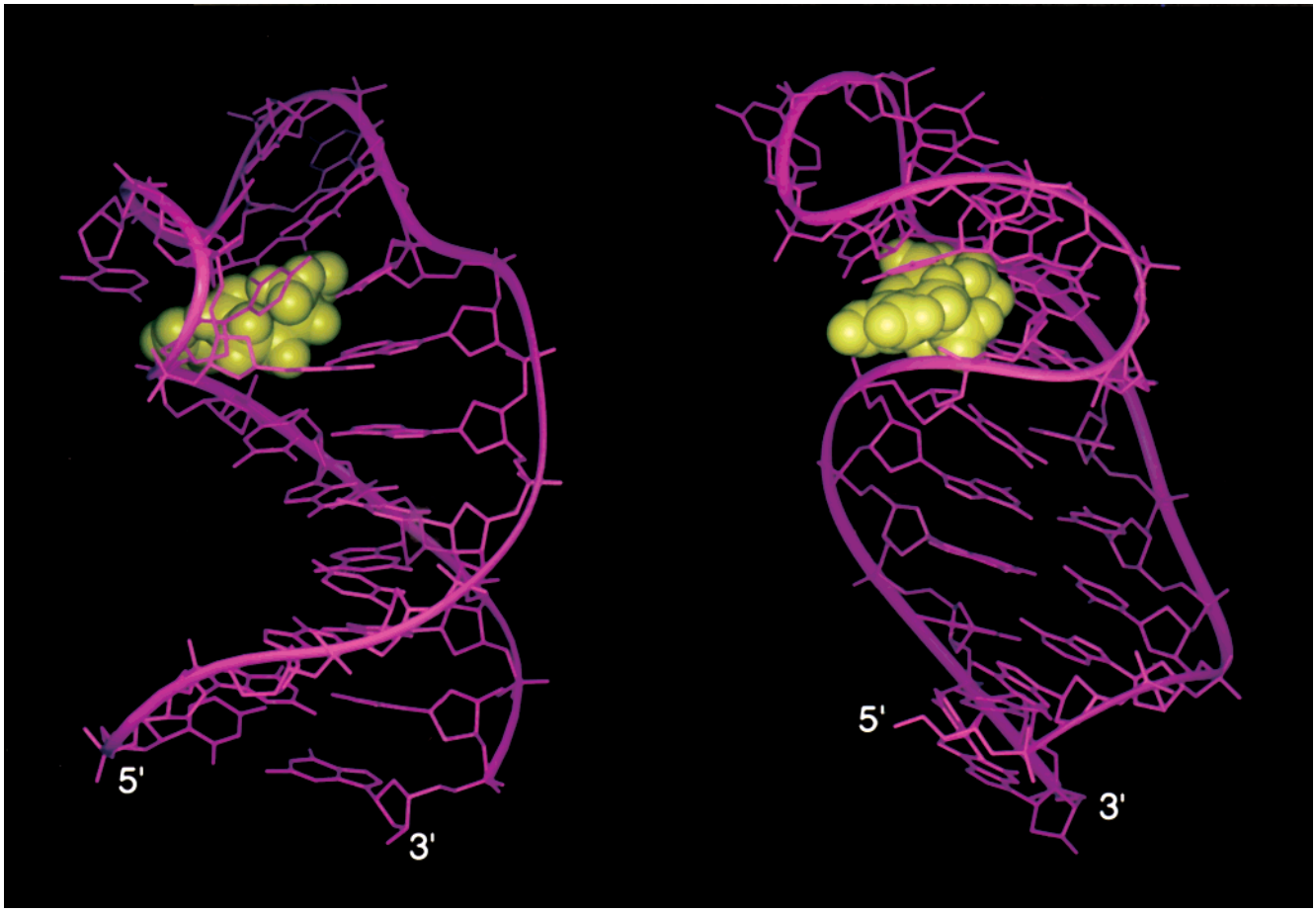

FIGURE 5: Two overall views of the complex of the aptamer with argininamide bound are shown. All heavy atoms are drawn, with the backbone indicated by a ribbon and the argininamide shown in a space filling representation.

of the C8 base in a binding pocket that does not have the C16 base-pairing out of sequence. Instead, this structure has the C16 and C15 roles reversed so that C16 is stacked normally above T17 and is the residue that forms the reverse Hoogsteen with A7, while C15 forms the Watson−Crick base pair with G9. The substitution of 5-methylcytosine for residue C16 in the current studies shows that this previously published structure has incorrectly assigned C16 and C15 and thus has the roles reversed for these two nucleotides. As a result, the binding pocket calculated is formed by pulling the G14−T10 base pair perpendicular to the rest of the loop base-pairs. The kink in the sequence between T17 and G14 that is clearly a dominant feature in the binding pocket formation in the current structures is missing. This paper also reported NOE contacts to the guanidinium protons that position the guanidinium group parallel to the C8 Watson−Crick edge. Since no NOEs were detected in the current complex between DNA and the ϵ or guanidinium hydrogens of argininamide, the orientation was not directly determined, and a number of structures did have the guanidinium group oriented in the opposite direction, facing toward the Watson−Crick edge of the C8 base (but not parallel to and in hydrogen bond position with respect to the base). In all cases, these structures significantly violated the restraints from the side chain of argininamide to the G14H8 (0.7−1.2 Å). Furthermore, this argininamide position in the structures caused the G14−T10 base pair to pull perpendicularly and away from the rest of the base-paired loop residues, resulting in a number of violations of restraints between G14 and C16 (which would also be violated in the previously published structures). Because of these numerous and significant violations, the structures with the guanidinium group facing the C8 base in the current study were discarded.

The structures presented here indicate that recognition of argininamide by the DNA aptamer is apparently strongly determined by contacts to the phosphate backbone, and by a binding pocket that can readily form to enclosed the argininamide side chain, amide, and amino groups. Most of the chemical modification and mutagenesis data on the aptamer (*23*) can be explained by necessity of each of the conserved bases for forming the structured loop conformation. Some sequence variants, such those altering T10 and C13, were studied by NMR and were shown to have some propensity to form alternate structures in the absence of argininamide (data not shown). Formation of a more stable unliganded structure would lower the binding affinity. Substitution of C8 with 5-methylcytosine showed significantly reduced binding affinity as judged by CD and lowered affinity for an arginine−agarose column (K. Harada, unpublished data). This calls into question the argument by Lin and Patel (*36*) that the C8 to T mutation lowers binding due to the loss of the C-type Watson−Crick edge. Instead, it is likely that the methyl group provides some steric hindrance in folding of the DNA into the structured, ligated form, and thus has a lowered propensity toward binding of the ligand.

Line shape analysis of the argininamide δ resonance under different argininamide−DNA ratios using simple free-bound equilibrium behavior did not accurately model the behavior seen. The calculated lines, while fitting the data moderately well, did not completely explain the broadening seen at low argininamide to DNA ratios and suggested a much weaker dissociation constant than that measured by CD (*23*). These discrepancies suggest the need for a more complicated model to describe the binding behavior of this complex. However, the currently available data are not sufficient to determine kinetic parameters of interest in a more complex model. Nonetheless, the NMR data still give insight into the unusual binding kinetics for this aptamer complex that are not apparent by measuring the DNA conformational change by CD.

CONCLUSIONS

Oligonucleotides selected for specific binding to amino acids are an interesting minimalistic approach to understanding protein−nucleotide recognition elements and can lend important insight into recognition of functional groups in small molecules by oligonucleotides. Most of the chemical modification and mutagenesis data for the argininamide-binding DNA aptamer that implicated the identity of the 10 bases critical in argininamide binding (*23*) can be explained by necessity for those specific residues to form the bound DNA conformation, rather than for specific argininamide−DNA contacts. This observation argues for selection of a DNA sequence with strong propensity toward a conformation that is further induced and stabilized by argininamide through electrostatic interactions. On the basis of selective criteria, clearly the guanidinium group is critical for binding of the ligand to the DNA. The structures presented here show phosphate−guanidinium group interactions as one of the defining elements for DNA recognition of the argininamide. These phosphate−guanidinium group interactions are common to many protein−nucleotide complex structures (*37*) and have been proposed as a means to specifically recognize RNA tertiary structure (*38*). The argininamide also appears to provide important electrostatic stabilization in a region that otherwise has numerous backbone phosphate groups in closer proximity than normally seen in nucleotide structures. The lack of obvious base-specific interactions with the argininamide may mean that the positioning of hydrogen bond elements of the ligand for interaction with backbone phosphates defines selection of the guanidinium group. The low-binding affinity of argininamide to this DNA aptamer as compared to some RNA aptamers that recognize arginine via important base-specific contacts (10 μM affinity) (*39*) is likely a result of these differences in modes of binding and the selection conditions chosen.

The structures here add to a growing database of knowledge on the variety of conformations available to oligonucleotides and lend insight into important roles in ligand−nucleotide interactions for both biological and pharmaceutical applications. Clearly, in some cases, stacking interactions and hydrogen bonds contribute to specificity and stability in unexpected ways and will provide useful insights as additional structural concepts used in design strategies in aptamer selection experiments.

ACKNOWLEDGMENT

S.A.R. would like to thank Randy Ketchnem, Jarrod Smith, and Patty Fagan for help with various aspects of the calculations done in this work, and for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

One table, S1, giving chemical shifts for resonances in the complex, two pages. This information is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gesteland, R. F., and Atkins, J. F. (1993) *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Nagai, K., and Mattaj, I. W. (1994) *RNA-Protein Interactions*. IRL Press, New York.
3. Dai, X., and Rothman-Denes, L. (1998) *Genes Dev. 17*, 2782−2790.
4. Darlow, J. M., and Leach, D. R. (1998) *J. Mol. Biol. 275*, 17−23.
5. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (1998) *J. Mol. Biol. 280*, 61−72.
6. Wadkins, R. M., Vladu, B., and Tung, C. S. (1998) *Biochemistry 37*, 11915−11923.
7. Geierstanger, B. H., and Wemmer, D. E. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 463−493.
8. Ramos, A., Gubser, C. C., and Varani, G. (1997) *Curr. Op. Struct. Biol. 7*, 317−323.
9. Wemmer, D. E., and Dervan, P. B. (1997) *Curr. Op. Struct. Biol. 7*, 355−361.
10. Chang, K. Y., and Varani, G. (1997) *Nature Struct. Biol.* Supplement, 854−858.
11. Wahl, M. C., and Sundaralingam, M. (1995) *Curr. Op. Struct. Biol. 5*, 282−295.
12. Ellington, A. D., and Szostak, J. W. (1990) *Nature 346*, 818−822.
13. Szostak, J. W. (1992) *Trends Biochem. Sci. 17*, 89−93.
14. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
15. Sassanfar, M., and Szostak, J. W. (1993) *Nature 364*, 550−553.
16. Burgstaller, A. T., and Ramulok, M. (1994) *Angew. Chem., Intl. Ed. Eng. 33*, 1084−1087.
17. Lauhon, C. T., and Szostak, J. W. (1995) *J. Am. Chem. Soc. 117*, 1246−1257.
18. Famulok, M. (1994) *J. Am. Chem. Soc. 116*, 1698−1706.
19. Connell, G. J., Illangesekare, M., and Yarus, M. (1993) *Biochemistry 32*, 5497−5502.
20. Tao, J., and Frankel, A. D. (1996) *Biochemistry 35*, 2229−2238.
21. Knight, R. D., and Landweber, L. F. (1998) *Chem. Biol. 5*, R215−220.
22. Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., and Toole, J. J. (1992) *Nature 355*, 564−566.
23. Harada, K., and Frankel, A. D. (1995) *EMBO J. 14*, 5798−5811.
24. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.
25. Klug, S. J., and Famulok, M. (1994) *Mol. Biol. Rep. 20*, 97−107.
26. Breaker, R. R. (1997) *Curr. Op. Chem. Biol. 1*, 26−31.
27. Feigon, J., Dieckmann, T., and Smith, F. W. (1996) *Chem. Biol. 3*, 611−617.
28. Patel, D. (1997) *Curr. Opin. Chem. Biol. 1*, 32−46.
29. Patel, D. J., Suri, A. K., Jiang, F., Jiang, L., Ran, P., Kumar, R. A., and Nonin, S. (1997) *J. Mol. Biol. 272*, 645−664.
30. Lin, C. H., Wang W., Jones, R. A., and Patel, D. J. (1998) *Chem. Biol. 5*, 555−572.
31. Sklenar, V., and Bax, A. (1987) *J. Magn. Res. 74*, 469−479.

32. Borgias, B. A., and James, T. L. (1990) *J. Magn. Res. 87*, 475−487.
33. McConnell, J. (1958) *J. Chem. Phys. 28*, 430−431.
34. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley & Sons, Inc., New York.
35. Chaires, J. B. (1998) *Curr. Opin. Struct. Biol. 8*, 314−320.
36. Lin, C. H., and Patel, D. J. (1996) *Nature Struct. Biol. 3*, 1046−1050.
37. Gronenborn, A. M., and Clore, G. M. (1995) *Crit. Rev. Biochem. Mol. Biol. 30*, 351−385.
38. Calnan, B. J., Tidor, B., Biancalanas, S., Hudson, D., and Frankel, A. D. *Science 252*, 1167−1171.
39. Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E., and Famulok, M. (1996) *Science 272*, 1343−1347.

BI9915061