The Tertiary Structure of a DNA Aptamer Which Binds to and Inhibits Thrombin Determines Activity†

Ke Yu Wang,‡ Steven H. Krawczyk,§ Norbert Bischofberger,§ S. Swaminathan,*,§ and Philip H. Bolton‡

Chemistry Department, Wesleyan University, Middletown, Connecticut 06459, and Gilead Sciences, 346 Lakeside Drive, Foster City, California 94404

Received May 27, 1993; Revised Manuscript Received August 3, 1993®

ABSTRACT: The solution-state three-dimensional structure of the DNA aptamer d($G_1G_2T_3T_4G_5G_6$-$T_7G_8T_9G_{10}G_{11}T_{12}T_{13}G_{14}G_{15}$) which binds to and inhibits thrombin has recently been determined by NMR methods (Wang et al., 1993). This DNA adopts a highly compact, highly symmetrical structure which consists of two tetrads of guanosine base pairs and three loops. The basic features of this three-dimensional structure are preserved when the aptamer binds to thrombin. The three-dimensional structure can be used as a basis for interpreting the relative activities of modified aptamers as well as for proposing a model for the aptamer–thrombin complex. This investigation also provides a demonstration of a novel approach to medicinal chemistry in which a wide range of molecules are synthesized, a lead molecule is identified, and the structural information on the lead compound allows for rational design of additional compounds of potential therapeutic value.

During the past few years a set of new approaches based on the mimicry of biological processes for the discovery of potential therapeutic agents have been developed (Joyce, 1992). These approaches are based on the synthesis of a very large number of molecules which are then screened for activity against a particular molecular target, and the selected molecules are then amplified (Joyce, 1992; Tuerk & Gold, 1990; Beaudry & Joyce, 1992). Typically, many rounds of selection are used to obtain a set of molecules with the desired activity. These procedures have been referred to as directed evolution. The common feature of these in vitro systems with natural biological evolution is that selective pressure is applied to a diverse population of molecules and only those molecules which have the proper characteristics are reproduced.

A successful application of a directed evolution methodology has been the discovery of a class of DNA molecules which bind to and inhibit thrombin (Bock et al., 1992). An RNA molecule which binds to a specific molecular target is referred to as an aptamer (Ellington & Szostak, 1990), and the aptamer term has been extended to DNA (Bock et al., 1992; Wang et al., 1993). The inhibition of thrombin has therapeutic value in cardiovascular surgery and other settings in which anticoagulation is desired (Salzman, 1992). Acute vascular diseases such as myocardial infarction and cerebral infarction also constitute major health risks in which anticoagulation is desired. Thrombin is a protease which is central to the blood coagulation cascade and hence a major target for anticoagulation and vascular disease therapy.

It was found that d($G_1G_2T_3T_4G_5G_6T_7G_8T_9G_{10}G_{11}$-$T_{12}T_1{}^3G_{14}G_{15}$) is the DNA aptamer with the highest reported affinity for thrombin (Bock et al., 1992). This DNA aptamer was found to significantly increase the thrombin-catalyzed clotting times of both purified fibrinogen and human plasma and does not compete with known active site inhibitors of thrombin (Bock et al., 1992). The aptamer is also active in animal models. A comparison of the sequences of the DNA aptamers which bind most strongly to thrombin indicated that there is a consensus sequence of GGTTGGTNTGGTTGG.

A report on the determination and features of the tertiary structure of d($G_1G_2T_3T_4G_5G_6T_7G_8T_9G_{10}G_{11}T_{12}T_{13}G_{14}G_{15}$) in solution based on NMR methods was recently presented (Wang et al., 1993). This DNA exhibits a number of long-range NOEs between residues which are not adjacent in sequence, which allowed the determination of the tertiary structure. This DNA adopts a highly compact, highly symmetrical structure which consists of two tetrads of guanosine base pairs and three loops. The residues of the tetrads alternate anti–syn–anti–syn. This novel structural motif for DNA may also be relevant to the structure of telomere DNA (Wang et al., 1993) and has some features in common with a structure of a DNA dimer determined by crystallography (Kang et al., 1992). The tertiary structure of this molecule is depicted in Figure 1.

The tertiary structure of the aptamer, unlike the duplex structure of B-DNA, is highly sequence specific. This implies that if the aptamer tertiary structure is important for thrombin binding and inhibition, then the activity of DNA molecules should correlate with their having a sequence compatible with the tertiary structure of the aptamer. It is noted that the aptamer tertiary structure is both thermodynamically stable with a T_m of about 50 °C and is kinetically stable with the imino protons of the G quartets having exchange lifetimes on the order of months. These physical characteristics suggest that the tertiary structure of the aptamer is likely to be preserved upon binding to thrombin.

This investigation demonstrates the next step in the development of potential therapeutic agents based on directed evolution approaches. This next step is the incorporation of structural information about the potential therapeutic agent in the presence and absence of the molecular target to guide further design.

Molecular modeling and other studies described below suggest that the aptamer binds to the exo site of thrombin. Aptamer binding to thrombin does not inhibit the activity

† The research at Wesleyan was supported, in part, by the Patrick and Catherine Weldon Donaghue Medical Research Foundation.

‡ Wesleyan University.

§ Gilead Sciences.

® Abstract published in *Advance ACS Abstracts*, October 1, 1993.

0006-2960/93/0432-11285$04.00/0

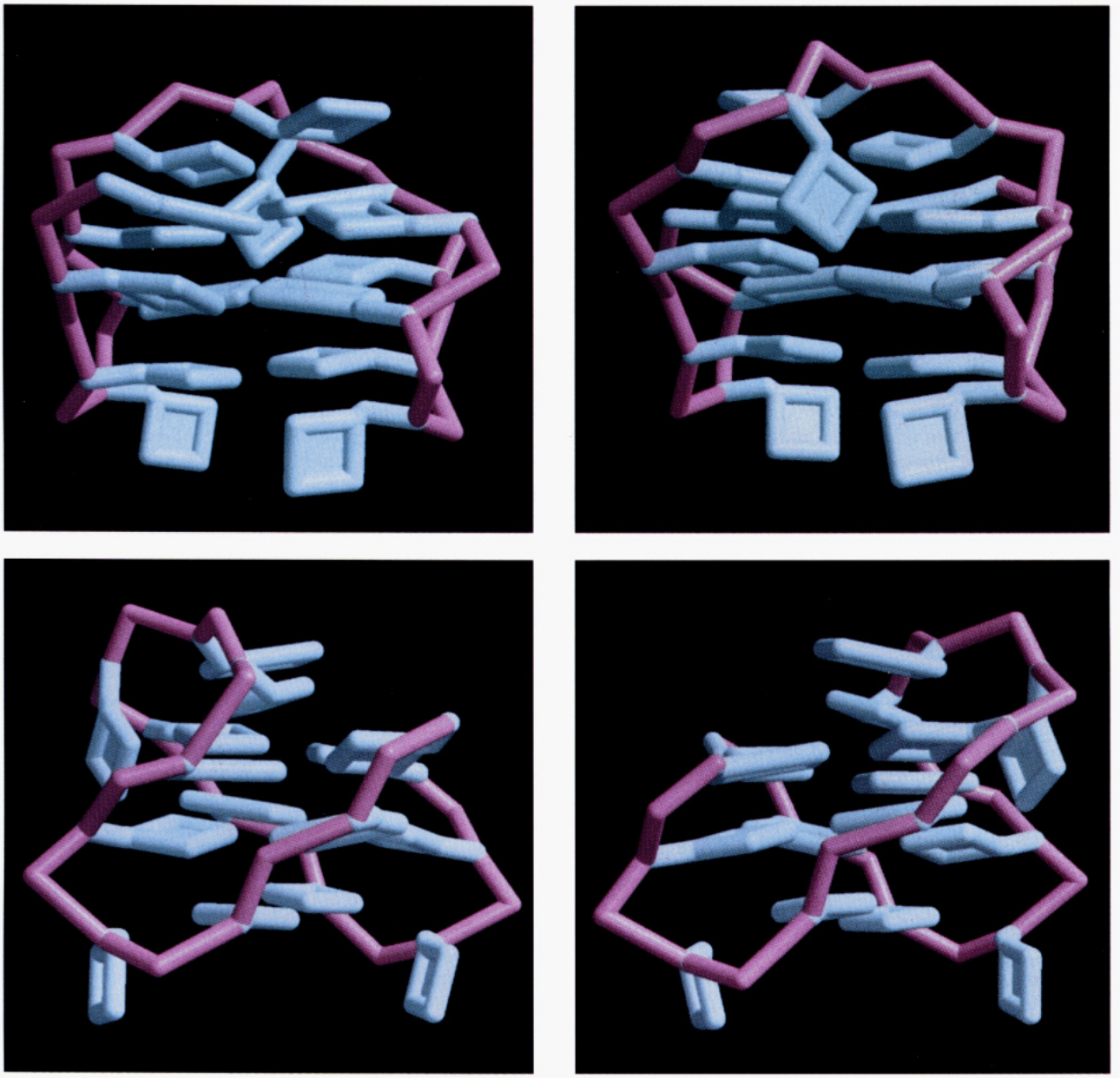
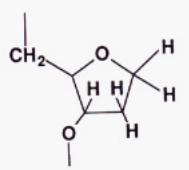


FIGURE 1: The structure of the aptamer is shown with the bases in blue and the backbone in pink. In panel A (top left) the "front" view of the aptamer is shown with residue 1 in the upper right and residue 15 in the upper left and the TGT loop on top. In panel B (top right) the back view of the aptamer is shown to illustrate the position of the $T_7G_8T_9$ loop relative to the $\underline{G}_1$-G_6-G_{10}-G_{15} tetrad. In panel C (bottom left) a side view of the aptamer is shown with residue 1 to the front of residue 15. In panel D (bottom right) the complementary side view is shown with residue 15 to the front of residue 1.

toward small substrates nor does aptamer binding compete with known active site inhibitors. Thus, it appears that the aptamer inhibits the proteolysis of fibrin by blocking the exo site. This investigation is aimed at examining the conformation of the aptamer bound to thrombin, to determine if the conformation of the aptamer undergoes significant change upon binding to thrombin and how this structural information can be used to interpret the activities of various analogues of the aptamer. It will be shown that the structural information obtained on the aptamer, free in solution and bound to thrombin, can be used to propose a mode of action for the aptamer, to interpret the results on analogues of the aptamer and to guide future drug design.

MATERIALS AND METHODS

The DNA aptamer was synthesized by phosphoramidite methods and purified by HPLC. All NMR samples of the aptamer were at a concentration of 150 OD_{260} in the selection buffer containing 0.14 M NaCl, 20 mM perdeuteriated Tris, and 5 mM potassium chloride at a pH of 7.4. All of the NMR experiments reported here were obtained with the sample at 10 °C. Phosphorothioate-labeled samples (Eckstein & Jovin, 1983) were also synthesized by phosphoramidite methods and purified by HPLC. The phosphorothioate samples contained both diastereomers. The DNA samples containing tetrahydrofuran abasic sites (I), 7-deazaguanosine, inosine, 8-methylguanosine, or uracil were also synthesized by phosphoramidite methods and purified by HPLC.



tetrahydrofuran abasic site (I)

NMR spectra were obtained using Varian XL-400 and VXR-400 spectrometers as well as a Bruker AM-600 located at the NIH regional resource at the University of Wisconsin, Madison. Phase-sensitive 600-MHz NOESY experiments at 100- and 250-ms mixing times in D_2O were collected into 2048 complex points in t_2 and 800 increments of t_1 with a spectral width of 6024.1 Hz and a recycle delay of 1.3 s, and 48 transients were obtained for each increment of the evolution time following two dummy scans. The 250-ms 600-MHz NOESY in H_2O were recorded with a spectral width of 13 157.9 Hz, 2048 complex points in t_2, and 490 increments of t_1, and 64 transients were obtained for each increment of the mixing time. TOCSY experiments at 600 MHz were acquired into 2048 complex points in t_2 and 774 increments of t_1 with a mixing time of 51 ms, and 48 transients were obtained for each increment of the mixing time.

The free induction decays of the two-dimensional NOESY and TOCSY data were multiplied by a 90°-shifted sine square bell weighting in both dimensions. The two-dimensional spectra were zero-filled to 2048 by 2048 real points. FELIX and Varian VNMR software were used for the data processing.

PECOSY data were collected on a Varian VXR-400 spectrometer at 10 °C. The spectral widths were 3240.4 Hz with 2048 complex points in t_2 and 512 FIDs in t_1. A total of 64 scans were averaged for each FID after two dummy scans, and a 1.0-s recycle delay was used. The data were processed with a Gaussian weighting in both dimensions and a 2048 by 2048 real point matrix using Varian VNMR data processing.

One-dimensional NOE experiments were carried out by the standard difference spectra approach with selective saturation for 0.5 s. These experiments were carried out on a Varian VXR-400 NMR spectrometer. The spectral width was 4000 Hz with 16 000 complex points. A total of 256 scans were collected after two dummy scans with a recycle delay of 0.9 s.

The ^{31}P spectrum of the aptamer is the same over the temperature range of 10–20 °C except in terms of resolution. The ^{31}P data were collected with 6016 complex points, a spectral window of 3271.2 Hz, and a delay time of 0.5 s, and 1024 transients were obtained.

In the case of the phosphorothioate-labeled samples, 1600 transients were obtained with 10 944 complex points, a spectral width of 11 001.1 Hz, and a recycle delay of 0.5 s. The ^{31}P resonances of the two stereoisomers of the phosphorothioate were observed at about 55 ppm downfield from the corresponding phosphates (Eckstein & Jovin, 1983). The spectra were processed by VNMR software with 1-Hz line broadening.

Aptamer–Thrombin Experiments. The covalent active site inhibitor D-Phe-Pro-Arg chloromethyl ketone, 0.65 mg (Skrzypczak-Jankun et al., 1989; Bode et al., 1989), was added to 12 mg of human thrombin (MW = 36 700; Haematologic Technologies Inc.) followed by dialysis for 1 week against a buffer solution of 46.7 mM NaCl, 1.67 mM KCl, 6.67 mM perdeuteriated Tris, and 1 mM EDTA at pH 5.3 and 0 °C to remove the glycerol and phosphate present in the original thrombin sample as well as excess inhibitor. The volume of the sample was reduced from 1.2 to 0.5 mL by blowing nitrogen over the solution in a 0.5-mm NMR tube at 0 °C. The sample was checked by ^1H and ^{31}P NMR at 10 °C before addition of DNA to thrombin to make sure that there was no phosphate or other impurities in the thrombin sample.

The sample of the aptamer–thrombin complex contained 0.34 μmol of thrombin and 0.13 μmol of the aptamer, 14 OD_{260}. The one-dimensional ^{31}P NMR data were collected into 27 648 complex points with a spectral width of 13 831.3 Hz and a recycle delay of 2.6 s on a Varian VXR-400 spectrometer at 10 °C, and a total of 22 016 scans were averaged.

Optical melting curves were obtained using a Hewlett-Packard UV/vis diode array spectrophotometer equipped with the Hewlett-Packard variable temperature accessory. The optical melting curves were obtained at 260 nm with the DNA sample in the same buffer as used for the NMR experiments.

NOE Cross-Peak Quantification. NOE cross-peak volumes from the 600-MHz 250-ms NOESY were quantified using the program FELIX. A rectangular box was used to cover each cross-peak, and the shape and size of each box were chosen manually of best fit an individual cross-peak. The results are summarized in Table I.

Structure Determination Methods. The structure for the aptamer was generated using the NOE volumes as distance constraints. The NOE cross-peak volumes were categorized as strong, medium, and weak to represent distance ranges (Clore & Gronenborn, 1989). These constraints were used to construct the structures for the G-G, T-T, and T-G-T fragments. These fragments were then connected to obtain an initial structure for the aptamer. A Dreiding II force field (Mayo et al., 1990) was then set up for the molecule, which included 2.5 kcal/Å^2 constraints for the hydrogen bonds of the G quartets. Thus, each G quartet was modeled with eight hydrogen bond and eight torsional angle constraints. There were 34 additional distance constraints which were each assigned a force constant of 5 kcal/Å^2. The initial structure was then subjected to 100 steps of conjugate energy minimization, followed by 25 ps of microcanonical dynamics and by an additional 100 steps of conjugate energy minimization.

The final structure, which is shown in Figure 1, was then used to predict the experimental NOE cross-peak volumes by solving the complete relaxation matrix assuming a 3-ns overall motion correlation time. The predicted and experimental NOE cross-peak volumes were found to be in good general agreement, as were the predicted and experimental scalar couplings.

Aptamer–Thrombin Complex Structure. The aptamer–thrombin complex model was built on the basis of the interactions between the negatively charged phosphate groups of the aptamer and the positively charged amino acids at the exo site of thrombin. The docking utilized the NMR structure of the aptamer (Wang et al., 1993) and the crystal structure of thrombin (Rydel et al., 1990). The docking procedure involved attempts using several orientations of the aptamer in the general region of the exo site. The model for the aptamer–thrombin complex is the one which maximized the favorable interactions between the complementary charged groups of the aptamer and the thrombin. The model also allowed close contact between the aptamer and thrombin without unfavorable steric interactions and with minimal movement of the thrombin from the starting crystal structure. The model for the aptamer–thrombin complex is shown in Figure 8.

Aptamer Activity and Binding Constant Determinations. The clotting time assays and inhibition constants of the modified aptamers were performed as described elsewhere (Bock et al., 1992).

RESULTS AND DISCUSSION

The basic structural characterization of the aptamer has been previously described (Wang et al., 1993), and additional details are provided here. The imino proton spectrum indicated the presence of eight imino protons which are slowly exchanging with water. The examination of the imino proton

Table I: Interresidue and Intraresidue NOEs

(A) Interresidue NOEs[a]

strong	G2 H8–G1 H2′ G2 H8–G2 H2″ ***G2 H1′–T4 CH_3*** ***G2 H2″–T4 CH_3***		G6 H8–G5 H2′ G6 H8–G5 H2″	T9 CH_3–G8 H1′ ***T9 CH_3–G1 H1*** ***T9 CH_3–G6 H1*** ***T9 CH_3–G10 H1*** ***T9 CH_3–G15 H1***	G11 H8–G10 H2″ ***G11 H1′–T13 CH_3*** ***G11 H2″–T13 CH_3***		G15 H8–G14 H2′ G15 H8–G14 H2″
medium	G2 H1′–T3 CH_3 G2 H2″–T3 H6 ***G2 H2′–T4 CH_3***	T4 CH_3–T3 H6	G6 H8–G5 H1′ G6 H8–G5 H3′ G6 H1′–T7 H2′ G6 H1′–T7 H3′	T9 CH_3–G8 H2″	G11 H8–G10 H2′ G11 H1′–T12 CH_3 G11 H2″–T12 H6 ***G11 H2′–T13 CH_3***	T13 CH_3–T12 H6	
weak	G2 H8–G1 H1′ G2 H8–G1 H3′ G2 H1′–T3 H6 G2 H2′–T3 H6 ***G2 H8–T4 CH_3*** ***G2 H8–T4 H2′*** ***G2 H8–T4 H2″*** ***G2 H8–T4 H6***	T4 CH_3–T3 H2′ T4 CH_3–T3 H2″ T4 CH_3–T3 H1′ T4 CH_3–T3 CH_3 T4 H6–T3 H2′ T4 H6–T3 H2″	G6 H1′–T7 H6 ***G6 H1′–G8 H8*** ***G6 H2′–G8 H8*** ***G6 H2″–G8 H8***	T9 H6–G8 H1′ T9 H6–G8 H2″	G11 H8–G10 H1′ G11 H1′–T12 H6 G11 H2′–T12 H6 ***G11 H8–T13 CH_3*** ***G11 H8–T13 H2′*** ***G11 H8–T13 H2″*** ***G11 H8–T13 H6***	T13 H6–T12 H2′ T13 H6–T12 H2″ T13 CH_3–T12 H1′ T13 CH_2–T12 H2′ T13 CH_3–T12 H2″ T13 CH_3–T12 CH_3	G15 H8–G14 H1′ ***G15 H8–T9 H1′***

(B) Intraresidue NOEs

	G1	G2	T3	T4	G5	G6	T7	G8	T9	G10	G11	T12	T13	G14	G15
strong	H8–H1′ H1′–H2″ H2′–H2″	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2′ H6–H2″ H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2′ H6–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′ H3′–H4′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2″ H6–H2′ H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′
medium	H1′–H2′ H1′–H4′ H2′–H3′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H6–H1′ H6–H2′ H1′–H4′	H6–H1′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H8–H1′ H8–H3′ H1′–H4′	H6–H1′ H6–H3′ H1′–H2′ H1′–H4′ H3′–H4′	H8–H1′ H8–H2″ H3′–H4′	H6–H1′ H6–H2′ H1′–H3′	H1′–H2′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H6–H1′ H6–H2′ H1′–H4′	H6–H1′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H8–H3′ H1′–H4′
weak	H8–H2′ H8–H2″ H8–H3′	H8–H1′ H8–H3′	H6–H2″ H6–H3′	H6–H3′ H6–H4′ H1′–H3′ H1′–H4′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H1′–H3′	H6–H4′ H1′–H3′	H8–H3′ H8–H4′ H1′–H3′ H1′–H4′	H6–H2″ H6–H3′ H6–H4′ H1′–H4′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′	H6–H2″ H6–H3′	H6–H3′ H1′–H3′ H1′–H4′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H1′–H3′

[a] The long-range NOEs are in bold italic type. The imino protons have not been assigned to specific residues, only to which quartet, and the NOEs between imino protons and H6, H1′ of T9 and H2′, H2″ of T4 and T13 are not included in this table.

region as a function of temperature showed that the aptamer melts at about 50 °C. By way of contrast, the hypochromicity at 260 nm was very small, in agreement with studies on telomere DNA sequences (Sen & Gilbert, 1990; Henderson et al., 1987). The hypochromicity for the DNA aptamer was too small to allow an accurate optical melting temperature to be determined.

The exchange rates of the imino protons are on the time scale of months. This very slow exchange has been previously noted for nucleic acids containing G tetrads (Smith & Feigon, 1992; Wang & Patel, 1992; Wang et al., 1993). The slow exchange rate is most likely due to a large entropic barrier associated with a cooperative opening of the aptamer tetrad containing strucure in contrast to the breathing of duplex DNA.

The one-dimensional proton NMR spectrum showed that only one form of the aptamer is present in solution since there are 15 signals from the nine G H8 and six T H6 protons in the aromatic region. The proton spectrum of the aptamer was assigned on the basis of NOESY, DQCOSY, PECOSY, and TOCSY data (Van de Ven & Hilbers, 1988; Clore & Gronenborn, 1989; Wüthrich, 1986; Ernst et al., 1987) obtained at 400 and 600 MHz.

The NOESY data demonstrated that four of the G residues are in the syn conformation since the GH8–H1′ intraresidue NOEs of these residues are very strong. Therefore, assignments had to be made for 5′-anti–3′-anti, 5′-anti–3′-syn, and 5′-syn–3′-anti linkages. In 5′-anti–3′-anti linkages there are typically 5′ H1′, H2′ to 3′ H8/6 NOEs, and in 5′-syn–3′-anti there are 5′ H1′, H2′, H3′ to 3′ H6/8 NOEs. Interresidue NOEs are typically weak for 5′-anti–3′-syn linkages (Van de Ven & Hilbers, 1988; Clore & Gronenborn, 1989; Wüthrich, 1986).

Assignment of the Proton NMR Spectrum. The spectra were assigned using NOESY data such as those in Figures 2 and 3. The cross-peaks between the aromatic and H2′/H2″ protons are shown in Figure 2, and the cross-peaks between the aromatic and H1′ protons are shown in Figure 3. There are two strings of 5′-G-G-T-T-3′ connectivities which are $\underline{G}_1$-G_2-T_3-T_4 and $\underline{G}_{10}$-G_{11}-T_{12}-T_{13}. Syn G residues are designated by G. Two additional 5′-G-G-3′ connectivities are shown between $\underline{G}_{14}$-G_{15} and $\underline{G}_5$-G_6. Two stretches of 5′-GTT-3′ residues could also be identified on the basis of anti–anti connectivities. In addition to the sequential NOEs a number of long-range NOEs, some of which are shown in Figure 3 listed in Table I, were identified. The combination of the sequential and long-range connectivies allowed all of the resonances to be assigned with one ambiguity. The NOE data were not sufficient to unequivocally determine which stretch of GGTT arose from $\underline{G}_1$-G_2-T_3-T_4 and which from $\underline{G}_{10}$-G_{11}-T_{12}-T_{13}. There are no clear NOEs from T_4 to $\underline{G}_5$ nor from T_{13} to $\underline{G}_{14}$, and the resonances of T_4 and T_{13} are nearly degenerate as shown in Figure 3. To obtain unequivocal assignment information, two modified aptamers were prepared and examined. One of these contained U at position 4 and the other contained U at position 13. The one-dimensional spectra in Figure 4 show that neither of these substitutions has a significant impact on the chemical shifts of the aromatic protons. The chemical shift of the H6 proton of U is distinct from that of T such that the substitutions allowed the lifting of the near degeneracy of the H6 protons of T_4 and T_{13}. One-dimensional NOEs were obtained for the H6 proton of T_{13} of the U4-labeled sample and the H6 proton of T_4 of the U13 sample, and representative results are shown in Figure 5. The one-dimensional NOE data allowed the determination of the assignments of the $\underline{G}_1$-G_2-T_3-T_4 and $\underline{G}_{10}$-G_{11}-T_{12}-T_{13} sections

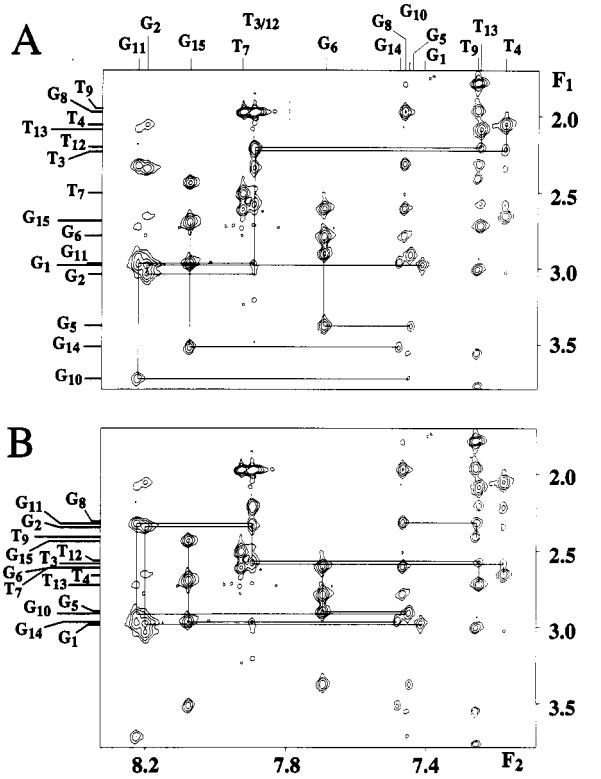


FIGURE 2: The spectrum shown is a portion of the 600-MHz NOESY spectrum obtained with a mixing time of 250 ms at 10 °C. The cross-peaks are between H8/H6 and H2′/H2″ protons. In panel A the anti–anti connectivities are indicated, and in panel B those involving syn residues are indicated. This region also contains long-range connectivities such as that between the H8 of G8 and the H2′ of G6.

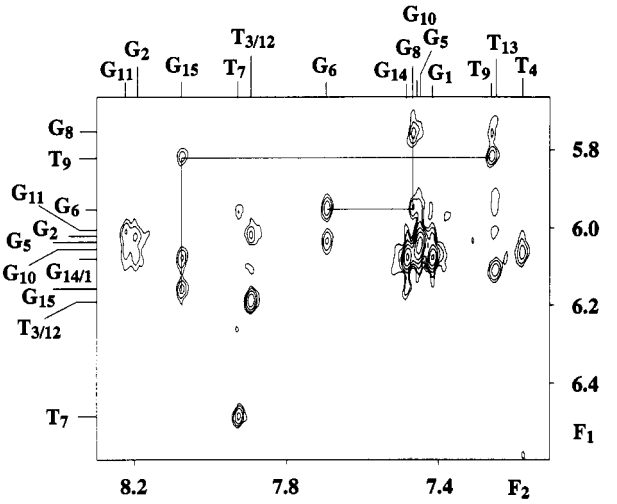


FIGURE 3: The contour plot shown is a portion of the 600-MHz NOESY spectrum obtained with a mixing time of 250 ms. The cross-peaks are between the aromatic and H1′ protons with the resonance positions indicated. This region also contains long-range connectivities such as that between the H8 of G15 and the H1′ of T9.

of the aptamer. The completion of the assignments allowed the tabulation of the NOEs which are given in Table I.

Scalar Couplings and Sugar Conformations of the Aptamer. The DQCOSY and PECOSY spectra of the aptamer have been obtained at 400 MHz. The patterns of the cross-peaks indicate that all of the sugar conformations are predominantly in the 2′-endo conformation, and the scalar couplings are listed in Table II.

Assignment of the ^{31}P Spectrum of the Aptamer. The ^{31}P spectrum of the aptamer was assigned via the use of

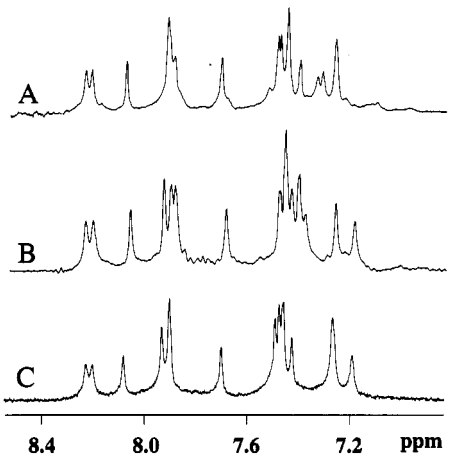


FIGURE 4: The spectra are of the aptamer (C), the U4-containing aptamer (B), and the U13-containing aptamer (A), obtained at 400 MHz.

phosphorothioate labeling of each of the 14 positions with typical data, and the assignments of the ^{31}P spectrum are shown in Figure 6. It is noted that the two most downfield ^{31}P signals are at quartet–two-base-loop junctions, G_2-T_3 and G_{12}-T_{13}.

These coupling constant and ^{31}P results are in contrast to those obtained with Z-DNA. In Z-DNA the ^{31}P chemical shifts of the syn G–phosphate–anti C sites are observed to be downfield shifted relative to those of the anti–syn sites of Z-DNA and of the anti–anti sites of B-DNA (Van de Ven & Hilbers, 1988; Feigon et al., 1985). Also, in Z-DNA the syn G residues have 3′-endo sugar conformations whereas the syn G residues of the aptamer have predominantly 2′-endo conformations. Thus, syn residues in DNA are not necessarily associated with either downfield ^{31}P chemical shifts or 3′-endo conformations.

Structure of the Aptamer. The structure for the aptamer is shwon in Figure 1 from several perspectives. This structure was obtained, as described above, from the combined use of the NOE and scalar coupling data and molecular modeling. The veiw in Figure 1A illustrates the compact nature of the aptamer structure. The view in Figure 1B shows the orientation of the T_7-G_8-T_9 loop relative to the $\underline{G}_1$-G_6-$\underline{G}_{10}$-G_{15} tetrad. The views shown in Figure 1C,D illustrate that the backbone of the aptamer contains essentially four straight segments connected by the three loops. These views also allow the low level of helicity of the backbone to be seen.

This structure for the aptamer is highly compact, has G-G-G-G tetrads which are syn–anti–syn–anti, and has a loop partially folded over the $\underline{G}_1$-G_6-$\underline{G}_{10}$-G_{15} tetrad and T_3, T_4, T_{12}, and T_{13} under the G_2-$\underline{G}_5$-G_{11}-$\underline{G}_{14}$ tetrad. This model is consistent with all of the NMR results, with the consensus sequence for thrombin binding and that GGTTGG also binds to and inhibits thrombin (Bock et al., 1992).

Structure–Activity Correlations of Modified Aptamers. A series of modified aptamers have been synthesized and their inhibition constants determined (Krawczyk et al., 1993). The inhibition constant is the concentration required to double the time for the thrombin-catalyzed polymerization of fibrinogen. The inhibition constants of the modified aptamers were determined with the results listed in Table III. The unmodified aptamer has an inhibition constant of 20 nM (Bock et al., 1992).

The results show that the substitution of an abasic residue for any of the G residues participating in the G-G-G-G tetrads significantly reduces the inhibition of thrombin. The substitution of an abasic residue for either T_4 or T_{13}, which are stacked on the G_2-$\underline{G}$-G_{11}-$\underline{G}_{14}$ tetrad, also significantly reduces inhibition. Substitution of an abasic residue for G_8 or T_9 reduces the activity by about 10-fold whereas substitution for T_7 has little effect on the inhibition. These results on the abasic substitutions are completely consistent with the structure proposed for the aptamer since substitution for any of the residues involved in either G tetrad or stacking on a G tetrad is expected to disrupt the structure and hence the binding to thrombin. The results on substitution for T_9 and G_8 are consistent with the structure having some interactions between these residues and the $\underline{G}_1$-G_6-$\underline{G}\underline{G}_{10}$-$G_{15}$ tetrad. The NMR results and the structure indicate that T_7 has at most weak

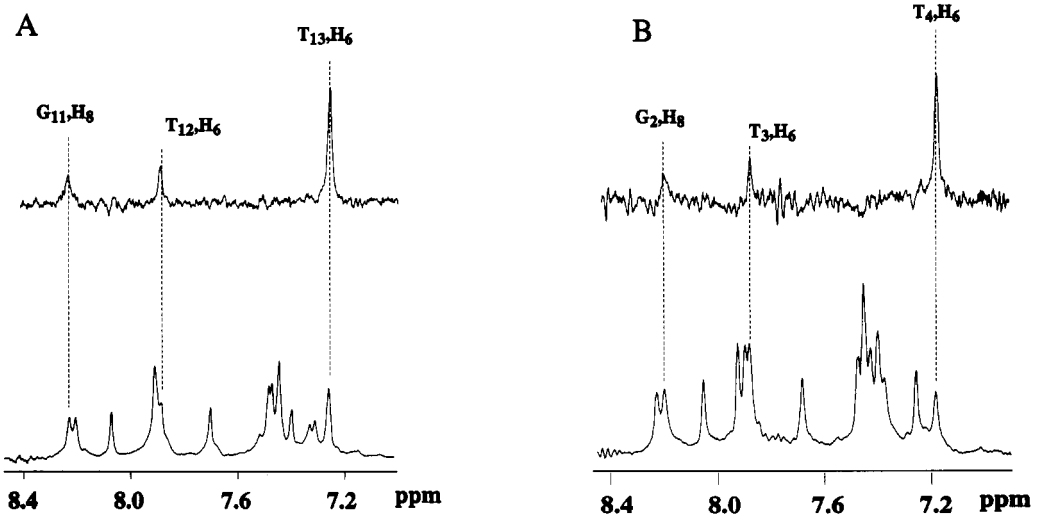


FIGURE 5: NOE difference spectra obtained for the selective saturation of the H6 resonance of T13 in the U4-containing apatmer (panel A) and for the selective saturation of the H6 resonance of T4 in the U13-containing aptamer (panel B) are shown.

Table II: Coupling Constants of the Aptamer

	G1	G2	T3	T4	G5	G6	T7	G8	T9	G10	G11	T12	T13	G14	G15
$J_{1'2'}$		9.5	3–6	9.6		9.6	9.5	9.6	9.5		9.5	3–6	9.5		9.5
$J_{1'2''}$			6.4	6.3	3–6	3.1	6.4	6.4	6.3	3.2	3.3	6.4	6.3	3.2	6.4
$\Sigma 1'$		16.2	13	15.6	15.8	18.9	15.9	15.8	15.8	18.9	16.0	13	15.6	19.3	15.9

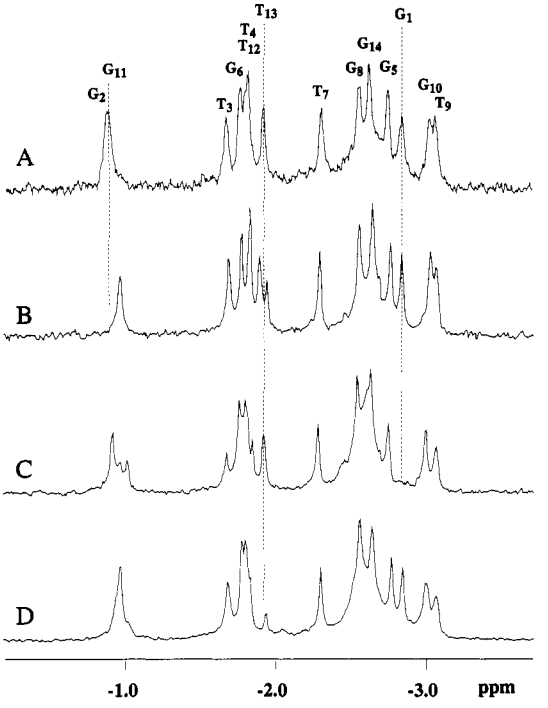


FIGURE 6: The ^{31}P spectra shown are for the aptamer (panel A), for the aptamer labeled with phosphorothioate at position 11 (panel B), for the aptamer labeled with phosphorothioate at position 1 (panel C), and for the aptamer labeled with phosphorothioate at position 13 (panel D).

Table III: Concentration of the Inhibitors Which Gives Rise to a Doubling of the Polymerization of Fibrinogen by Thrombin

	abasic	7-deaza-G	inosine	8-methyl-G
G1	>2 μM	>2 μM	>2 μM	~20 nM
G2	>2 μM	>2 μM	>2 μM	~160 nM
T3	~75 nM			
T4	>2 μM			
G5	>2 μM	>2 μM	>2 μM	~20 nM
G6	>2 μM	>2 μM	>2 μM	~80 nM
T7	~15 nM			
G8	~150 nM	~50 nM	~50 nM	~25 nM
T9	~120 nM			
G10	>2 μM	>2 μM	>2 μM	~20 nM
G11	>2 μM	>2 μM	>2 μM	~80 nM
T12	~75 nM			
T13	>2 μM			
G14	>2 μM	>2 μM	>2 μM	~20 nM
G15	>2 μM	>2 μM	>2 μM	~65 nM

interactions with the rest of the molecule, which is consistent with abasic substitution at this position having little effect on the thrombin inhibition constant.

All of the G residue positions have been individually substituted with 7-deazaguanosine and with inosine. 7-Deazaguanosine lacks the N7 of guanosine, which is the hydrogen bond acceptor of an amino group in G tetrads. Inosine lacks the amino group which is involved as a hydrogen bond donor, to an N7, in the G tetrads. Substitution of either 7-deazaguanosine or inosine for any of the G residues involved in the G tetrads effectively significantly reduces activity, which is consistent with the structure. Substitution of G_8, which is the only G not involved in a G quartet, has a minor effect as evidenced by the approximately 2-fold increase in inhibitiion constant.

The structure has G residues 1, 5, 10, and 14 with syn orientations about the glycosidic bond with all of the other G residues anti. Due to steric reasons 8-methylguanosine has a preference for the syn orientation (Saenger, 1984), and substitution of 8-methylguanosine for any of the syn G residues

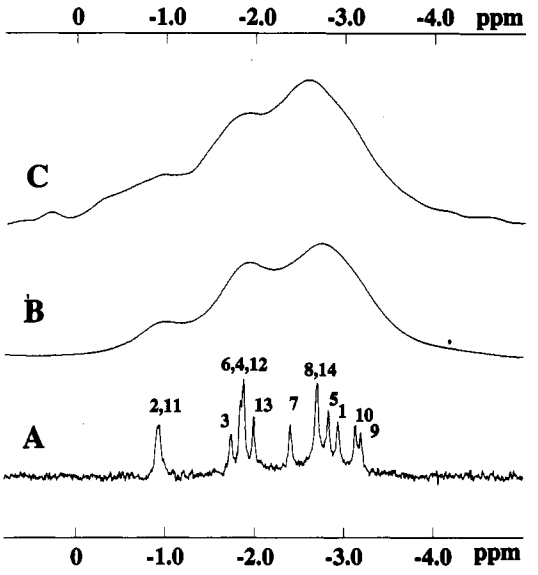


FIGURE 7: Panel A shows the ^{31}P spectrum of the aptamer, and panel B contains the ^{31}P spectrum of the aptamer which has been broadened for comparison with the ^{31}P spectrum of the aptamer–thrombin complex shown in panel C. The small signals at ≈–0.5 and ≈–4.5 ppm are at the noise level of this experiment.

has little or not impact on the thrombin inhibition constant. As expected on the basis of the structure, it was found that substitution of any of the anti G residues which are involved in a G tetrad with 8-methylguanosine led to an increase in the inhibition constant. Substitution for G_8 had little or no effect on the thrombin inhibition constant.

The aptamer structure is not only consistent with all of these substitution results but can be used to predict whether a particular substitution wil lhave no effect on inhibition, or increase the inhibition constant by a factor of about 10 or less, or whether a substitution will significantly reduces inhibition. In addition, these results not only support the aptamer structure, shown in Figure 1, but indicate that this structure needs to be preserved for efficient thrombin inhibition.

Aptamer–Thrombin Complex. The complex of thrombin, a 36.7-kDa protein, with the aptamer can be investigated by means of ^{31}P NMR. The ^{31}P NMR spectrum of the free aptamer, shown in Figure 7, exhibits an almost 2.5 ppm wide range of chemical shifts; the assignments of the ^{31}P resonances have been presented previously (Wang et al., 1993) and are also indicated in Figure 7. This range of chemical shifts is diagnostic for the conformation of the aptamer based on the examination of the ^{31}P spectra of the modified aptamers discussed above.

The binding constant of the aptamer to thrombin is about 10 nM, and hence all of the aptamer is bound to thrombin at the millimolar concentrations used for NMR. Figure 7 contains the spectra of the free aptamer and the aptamer in the presence of thrombin. The line widths of the aptamer increase to about 50–60 Hz when complexed with thrombin. These line widths are consistent with the correlation time of the ≈45-kDa complex of thrombin and aptamer. The spectrum of the free aptamer has been broadened to allow comparison with that of the thrombin complex and is shown in Figure 7. A detailed comparison of the spectra of the free and complexed aptamer is not feasible without information about the internal motion correlation times of the ^{31}P sites in the complex which affect the chemical shift anistropy and dipolar contributions to the observed line widths. However, these results show that the range and distribution of the ^{31}P shifts of the aptamer are not significantly changed upon binding to thrombin. This result supports the notion that the overall fold of the aptamer

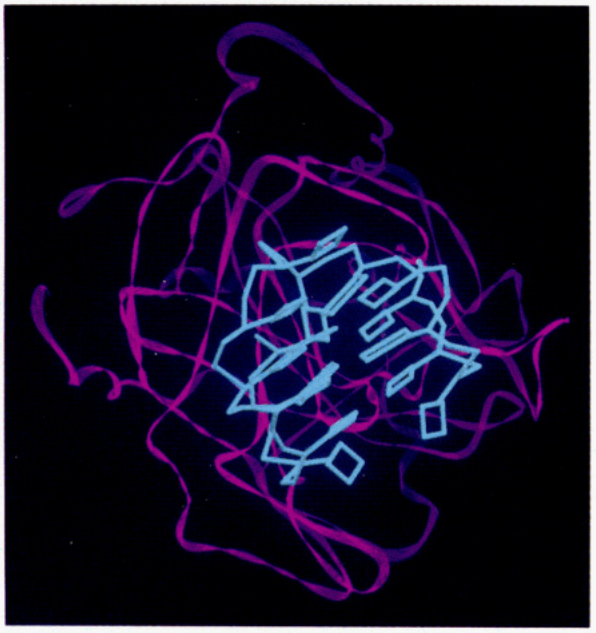


FIGURE 8: The structure of a model for thrombin binding to the exo site of thrombin is shown. The aptamer is shown in blue and the thrombin as a ribbon diagram in magneta.

is not changed upon binding to thrombin. These results do not rule out conformational changes which do not alter the ^{31}P chemical shifts.

Since these results, and those on the substituted aptamers, indicate that the structure of the aptamer does not undergo significant conformational change upon binding to thrombin, a preliminary model of the aptamer–thrombin complex has been built. In this model the solution-state structure of the aptamer has been docked with the crystal structure of thrombin (Rydel et al., 1990). The model structure for the complex is shown in Figure 8 in which the aptamer binds to the exo site of thrombin. The exo site of thrombin is an anion binding site which is distant from the active site and is thought to play a key role in determining the substrate specificity of thrombin (Fenton, 1981, 1986, 1989, Bode et al., 1989; Rydel et al., 1990; Skrzypczak-Jakun et al., 1989). The binding site of the aptamer can be thought of as residues 1, 2, 3 and 13, 14, 15 on the exposed side and residues 4 through 12 on the thrombin side. Small changes in the structure of the T_7-G_8-T_9 loop were made to accommodate complex formation. This binding site is consistent with results which show that the aptamer does not compete with active site inhibitors of thrombin, that the binding of the aptamer does not affect the activity of thrombin toward small substrates, and that the binding of the aptamer to thrombin protects some of the lysines at the exo site from methylation (Bock et al., 1992; Paborsky et al., 1993).

Relation to Drug Design. The results presented here indicate that the structure of the aptamer is required for efficient inhibition of thrombin and that the basic structural features of the aptamer free in solution are preserved upon binding to the exo site of thrombin. The aptamer structure can be used to predict which modified aptamers will inhibit thrombin. This is allowing the design of potential drugs which preserve the key structural features of the aptamer but which may have superior therapeutic value. Therefore, directed evolution-type approaches can be used to find a lead molecule, and the structure of the lead molecule in the presence of the molecular target can be used to predict the activity of potential therapeutic agents. It is anticipated that the determination of the nature of the thrombin–aptamer interactions will provide further clues to drug design.

ACKNOWLEDGMENT

The 600-MNz spectra were obtained with the assistance of Dr. Frits Abildgaard using the NMR facility at the University of Wisconsin, Madison, which is supported by NIH Grant RR02301 with equipment purchased with support from the University of Wisconsin, the NSF (DMB-8415048), the NIH (RR02781), and the USDA.

REFERENCES

Beaudry, A. A., & Joyce G. F. (1992) *Science 257*, 635–641.
Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., & Toole, J. J. (1992) *Nature 355*, 564–566.
Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., & Hofsteenge (1989) *EMBO J. 11*, 3467–3475.
Clore, G. M., & Gronenborn, A. M. (1989) *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.
Eckstein, F., & Jovin, T. M. (1983) *Biochemistry 22*, 4546–4550.
Ellington, A. D., & Szostak, J. W. (1990) *Nature 346*, 818–822.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of NMR in One- and Two-Dimensions*, Oxford University Press, Cambridge.
Feigon, J., Wang, A. H.-J., Van der Marel, G. A., Van Boom, J. H., & Rich, A. (1985) *Science 230*, 82–84.
Fenton, J. W., II (1981) *Ann. N.Y. Acad. Sci. 370*, 468–495.
Fenton, J. W., II (1986) *Ann. N.Y. Acad. Sci. 485*, 5–15.
Fenton, J. W., II, Witting, J. I., Pouliott, C., & Fareed, J. (1989) *Ann. N.Y. Acad. Sci. 556*, 158–165.
Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., & Blackburn, E. H. (1987) *Cell 51*, 899–908.
Joyce, G. F. (1992) *Sci. Am.*, 90–97 (December).
Kang, C., Zhang, X., Ratliff, R., Moyzis, R., & Rich, A. (1992) *Nature 356*, 126–131.
Krawczyk, S., Bischofberger, N., Griffin, L., Law, V., Shea, R., & Swaminathan, S. (1993) *Nucleic Acids Res.* (submitted for publication).
Mayo, S. L., Olafson, B. D., & Goddard, W. A., III (1990) *J. Phys. Chem. 94*, 8897–8909.
Paborsky, L. R., McCurdy, S. N., Griffin, L. C., Toole, J. J., & Leung, L. L. K. (1993) *J. Biol. Chem.* (in press).
Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, H., Roitsch, C., & Fenton, J. W., II (1990) *Science 249*, 277–280.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Salzman, E. W. (1992) *N. Engl. J. Med. 326*, 1017–1019.
Sen, D., & Gilbert, W. (1990) *Nature 344*, 410–414.
Skrzypczak-Jakun, E., Rydel, T. J., Tulinsky, A., Fenton, J. W., II, & Mann, K. G. (1989) *J. Mol. Biol. 206*, 755–757.
Smith, F. W., & Feigon, J. (1992) *Nature 356*, 164–168.
Tuerk, C., & Gold, L. (1990) *Science 24*, 505–510.
Van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Wang, K. Y., McCurdy, S., Shea, R. G., Swaminathan, S., & Bolton, P. H. (1993) *Biochemistry 32*, 1899–1904.
Wang, Y., & Patel, D. J. (1992) *Biochemistry 31*, 8112–8119.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.