Acta Cryst. (1996). D**52**, 272–282

An Ambiguous Structure of a DNA 15-mer Thrombin Complex

K. PADMANABHAN AND A. TULINSKY*

Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA

(Received 5 July 1995; accepted 9 October 1995)

Abstract

The structure of a complex between thrombin and a GGTTGGTGTGGTTGG DNA 15-mer has been analyzed crystallographically. The solution NMR structure of the 15-mer has two stacked G-quartets similar to that found in the previous X-ray structure determination of the 15-mer–thrombin complex [Padmanabhan, Padmanabhan, Ferrara, Sadler & Tulinsky (1993). *J. Biol. Chem.* **268**, 17651–17654]; the strand polarity, however, is reversed from that of the crystallographic structure. The structure of the complex here has been redetermined with better diffraction data confirming the previous crystallographic structure but also indicating that the NMR solution structure fits equally well. Both 15-mer complex structures refined to an R value of about 0.16 presenting a disconcerting ambiguity. Since the two 15-mer structures associate with thrombin in different ways (through the TGT loop in the X-ray and TT loop in the NMR model), other independent lines of physical or chemical evidence are required to resolve the ambiguity.

1. Introduction

Aptamer was initially used to refer to RNA molecules that bind to a specific molecular target (Ellington & Szostak, 1990; Szostak, 1992). The term has now been extended to DNA molecules (Bock, Griffin, Latham, Vermaas & Toole, 1992; Wang, McCurdy, Shea, Swaminathan & Bolton, 1993). Selection of DNA aptamers that bind to thrombin and inhibit thrombin-catalyzed fibrin-clot formation *in vitro* has been achieved using combinatorial methods (Bock *et al.*, 1992). A 15-mer consensus sequence G1G2T3T4G5G6T7G8T9-G10G11T12T13G14G15 was found to have the highest affinity for thrombin ($IC_{50} \simeq 100$ nm) and this sequence significantly increased the thrombin-catalyzed clotting times of both purified fibrinogen and human plasma.

Guanine bases in DNA are unique in that they can align in a planar G-tetrad arrangement in the presence of monovalent cations (Guschlbauer, Chantot & Thiele, 1990; Sundquist, 1991; Sen & Gilbert, 1991). Such G-quartet pairing alignments have been observed and characterized in systems ranging from guanine monophosphates (Pinnavaia, Miles & Becker, 1975; Pinnavaia *et al.*, 1978; Borzo, Detellier, Lazlo & Paris, 1980) to polyguanylic acid (Gellert, Lipsett & Davies, 1962). More recent interest has been in the direction of higher order structures adopted by single-stranded telomeric overhangs at the ends of chromosomes (Blackburn & Szostak, 1984; Blackburn, 1991).

In an effort to identify the region of thrombin with which the single-stranded DNA aptamer interacts, the inhibition of fibrinogen-clotting activity by the consensus sequence was studied using two recombinant mutant forms of thrombin (Arg75 → Glu and Lys149E → Ala)† (Wu, Tsiang & Sadler, 1992). The results with the Glu mutation suggest that the single-stranded DNA binding site is located in the thrombin fibrinogen-recognition exosite and overlaps the thrombin platelet receptor and thrombomodulin binding sites. The inhibition of thrombin-induced platelet activation by the aptamer has also been observed (Griffin, Tidmarsh, Bock, Toole & Leung, 1993). Binding at this anion-binding exosite of thrombin by platelet thrombin-receptor peptide (Mathews *et al.*, 1994) has been shown to be involved in platelet activation (Vu, Wheaton, Hung, Charo & Coughlin, 1991).

The aptamer binding site on thrombin was also examined by modeling, solid-phase plate binding assays and by chemical modification studies (Wang, Krawczyk, Bischofberger, Swaminathan & Bolton, 1993; Paborsky, McCurdy, Griffin, Toole & Leung, 1993). These showed that the thrombin aptamer bound specifically to α-thrombin but not to γ-thrombin and that hirudin competed with aptamer binding; γ-thrombin is a proteolylic cleavage product of α-thrombin with one cleavage in the fibrinogen-binding exosite (Fenton *et al.*, 1977; Rydel *et al.*, 1994). The results of these studies complement the results from the mutation experiment suggesting again that the thrombin anion-binding exosite is important for aptamer–thrombin interactions. The chemical-modification studies (Paborsky *et al.*, 1993) additionally identified Lys36 and Lys70 as integral components; both are located in the anion-binding exosite of thrombin (Skrzypczak-Jankun *et al.*, 1991). In addition, it is known that the aptamer did not compete with known active-site inhibitors so it is not involved with the catalytic site of the enzyme (Bock *et al.*, 1992).

† Chymotrypsin numbering (Bode *et al.*, 1989).

To determine the nature of the association in this novel, but non-physiological inhibition of thrombin activity, we determined the crystal structure of thrombin complexed with the consensus 15-mer (Padmanabhan *et al.*, 1993). The oligonucleotide folds into a unimolecular quadruplex in the DNA–thrombin complex containing two G-quartets linked by two TT loops at one end and a TGT loop at the other end. (Fig. 1). Solution NMR structures of the 15-mer have also been determined (Wang *et al.*, 1993, Macaya, Schultze, Smith, Roe & Feigon, 1993; Schultze, Macaya & Feigon, 1994) where the quartet structure is essentially the same as that found in the complex. However, the TT loops span the narrow grooves while the TGT loop spans the wide groove in the NMR structure while the opposite occurs in the complex (Fig. 1). The difference thus appears to be a reversal in strand polarity between the two* (Schultze *et al.*, 1994).

Other quadruplex structures have been determined by NMR (Henderson, Hardin, Walk, Tinoco & Blackburn, 1987; Wang *et al.*, 1991; Smith & Feigon, 1992; Wang & Patel, 1992, 1993). The first X-ray crystal structure determination of G-quadruplexes was a four-stranded dimer of two hairpin GGGGTTTTGGGG sequences, which had four G-quartets with a potassium ion at their center (Kang, Zhang, Ratliff, Moyzis & Rich, 1992). The consensus 15-mer is the first crystallographic structure of a G-quadruplex formed by a single-stranded DNA.

The high-resolution NMR structure of the sodium form of GGGGTTTTGGGG (Smith & Feigon, 1992) is different from the structure of the potassium form of the same oligonucleotide determined crystallographically (Kang *et al.*, 1992). In the former, a head-to-tail dimer has thymine loops connected diagonally across the G-quartet while in the latter, two antiparallel hairpins dimerize and thymidine residues form loops between adjacent strings of guanines and the two loops are at opposite ends of the G-quartet core. The result is that there are major differences in the phosphodiester backbone and in the mode of dimerization between the two structures even though the arrangement of the guanine bases of the two is nearly identical. A somewhat similar situation is observed between the DNA 15-mer structure complexed with thrombin and the solution NMR structure of the aptamer (Fig. 1).

A preliminary description of the crystal structure of the aptamer–thrombin complex indicates the dominant interactions in the fibrinogen-recognition site between the DNA and thrombin involve the TGT loop (Padmanabhan *et al.*, 1993). Since then, we have examined the NMR structure of the aptamer (Schultze *et al.*, 1994) with respect to the crystallographic diffraction data. The coordinates of the NMR structure of the 15-mer were fitted to the initial electron-density map of the complex without aptamer in the structure-factor calculations. By rotating the NMR model of the aptamer 180° about an axis between the two G-quartets of the X-ray structure and then 90° around an axis perpendicular to them (T7G8T9 and T12T13 of NMR superimposing approximately on T3T4 and T7G8T9 of X-ray, respectively) (Fig. 1), a reasonably good fit could be achieved of the NMR structure with the electron density. Even more surprisingly, the new model of the DNA, bound to the fibrinogen exosite of thrombin through its TT loops, rather than the TGT loop, also refined well to an R value of about 0.16 ($R = 0.159$ for the X-ray structure). Although this ambiguity presented a disturbing problem, precedence for conformational differences between crystal and solution structures of oligonucleotides with quartets and loops had already been established by the different crystal and NMR GGGGTTTTGGGG structures described above (Kang *et al.*, 1992; Smith & Feigon, 1992), which has also been critically reviewed by Williamson (1993). In an effort to resolve or confirm the ambiguity of the two models of the DNA aptamer with thrombin, the structure of the complex has here been redetermined at 2.8 Å resolution using a larger and better crystal. We describe the refinement (2.8 Å resolution) and comparison of both DNA models, which confirm the original observation that both structures fit the crystallographic data equally well. The manner of association of the DNA–thrombin complex remains in doubt and will have to be appealed to other independent lines of physical and/or chemical evidence.

* We would like to thank Dr Juli Feigon for pointing this out to us. The two aptamer structures will hereinafter be referred to as X-ray and NMR structures, respectively.

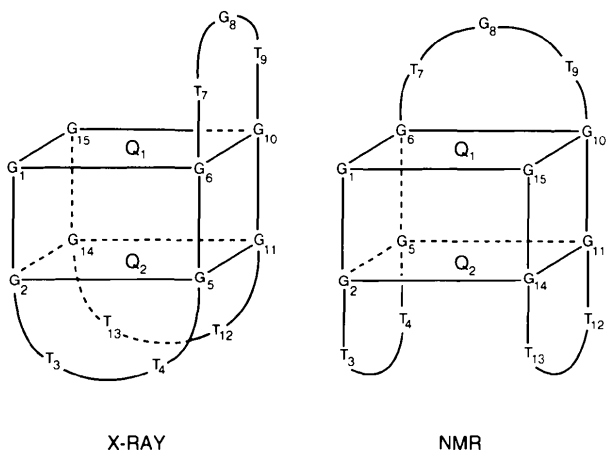

Fig. 1. Schematic drawings of the X-ray (left) and NMR (right) structures of the aptamer. TT loops span wide grooves, TGT spans narrow groove in X-ray structure and *vice versa* in NMR structure.

2. Experimental

2.1. *Crystallization*

The 15-mer oligonucleotide was synthesized using phosphoramidite chemistry on an Applied Biosystems DNA synthesizer model 380B. The thrombin–aptamer

complex was prepared by placing a twofold excess of 15-mer solution over a frozen 1 ml sample of human D-Phe-Pro-Arg chloromethyl ketone (PPACK)-thrombin (Fenton, Fasco *et al.*, 1977) at a protein concentration of about 0.91 mg ml^{-1} in 0.75 *M* NaCl. This was then diluted with an equal volume of 0.1 *M* sodium phosphate buffer (pH 7.3). The solution of the complex was concentrated to about 4 mg ml^{-1} using a Centricon-10 miniconcentrator in a refrigerated centrifuge. Crystals were grown by the hanging-drop method with 2 mg ml^{-1} protein complex, 0.025 *M* sodium phosphate buffer, 0.05 *M* Hepes buffer (pH 7.5), 0.05 *M* ammonium sulfate, 0.19 *M* NaCl, 10%(*w*/*v*) PEG 4000 and 4%(*v*/*v*) isopropanol in 10 µl drops equilibrated against 1 ml well solution of 0.1 *M* ammonium sulfate, 0.1 *M* Hepes buffer (pH 7.5), 20%(*w*/*v*) PEG 4000 and 8%(*v*/*v*) isopropanol. The size of the crystal used for the original intensity data collection was 0.4 × 0.1 × 0.02 mm (Padmanabhan *et al.*, 1993), while the crystal used for the structure redetermination here was twice the size (0.4 × 0.1 × 0.05 mm).

Both crystals belong to the orthorhombic space group $P2_12_12_1$, four complexes per unit cell but the unit-cell dimension along the *a* direction was about 2 Å longer in the larger crystal with $a = 58.28$, $b = 77.61$, $c = 100.02$ Å. The larger crystal diffracted X-rays marginally to about 2.5 Å resolution; however, only data to 2.8 Å were used for analysis due to the weak nature of the reflections beyond 2.8 Å resolution [1232 reflections observed (27%) but only $282 > 3\sigma$ in this range].

2.2. *Data collection*

X-ray diffraction intensity data were collected with an R-AXIS II imaging-plate detector system using graphite-monochromatized Cu $K\alpha$ radiation from a 0.3 × 3.0 mm source of a Rigaku RU200 rotating-anode X-ray generator operating at 50 kV and 100 mA. The crystal-to-detector distance was set at 120 mm, the oscillation range was 3.0° and each frame was exposed for 75 min. The unit cell was determined by autoindexing (Higashi, 1990). The data frames were reduced and scaled using the processing software for the R-AXIS II system. A total of 21 631 observations were averaged to give 9225 unique reflections with $|F|^2 > \sigma$ with an $R_{\rm merge}$ of 9.7%. This compares with only 12 408 observations to 2.9 Å resolution, 6990 unique reflections, with $R_{\rm merge} = 11.6\%$ for the original crystal (Padmanabhan *et al.*, 1993). The number of present observed reflections corresponds to 68% of the total number possible up to 2.8 Å (38% between 3.0 and 2.8 Å resolution).

2.3. *Structure determination and refinement*

Since the present data set was of better quality than the previous one and the *a* axis of this crystal was about 2 Å longer, the structure was determined *ab initio*. Rotation-translation searches for thrombin were carried out using the program *X-PLOR* (Brünger, 1992*a*). The 3000 largest Patterson vectors between 20.0 and 3.0 Å of the model map were chosen for the rotation search. The highest solution had a peak height of 6.6, 5.0σ above the mean in the resolution range 8.0–4.0 Å. The ten highest solutions were subjected to a Patterson correlation refinement (Brünger, 1990). The correlation for the highest solution was 0.14 but only 0.06 for the next highest solution. The translation search using the best rotation gave a peak with a correlation of 0.51 (15σ above the mean) in the 8.0–4.0 Å resolution range. Rigid-body refinement of this thrombin structure in the same resolution range resulted in $R = 0.37$. All of the foregoing was practically the same as that obtained with the original crystal (Padmanabhan *et al.*, 1993).

The rigid-body model was refined using the restrained least-squares program *PROFFT* (Finzel, 1987). Overall *B* followed by individual *B* refinement in the 10.0–2.8 Å resolution range gave $R = 0.27$. The $(2F_o - F_c)$ electron density and $(F_o - F_c)$ difference maps were calculated and the thrombin molecule was fitted examining the maps. The PPACK inhibitor in the catalytic site was found and also fitted to the map. The region between the fibrinogen-binding exosite and the heparin-binding site (Karshikov, Bode, Tulinsky & Stone, 1992; Gan, Li, Chen, Lewis & Shafer, 1994; Ye, Rezaie & Esmon, 1994; Sheehan, Tollefsen & Sadler, 1994) of a neighboring thrombin molecule showed density corresponding to the DNA, as was the case with the previous structure (Padmanabhan *et al.*, 1993). This model, still without DNA, was subjected to further refinement and the *R* value reduced to 0.25.

An examination of the $(2F_o - F_c)$ electron density and $(F_o - F_c)$ difference maps at this stage revealed several noteworthy features. The density for the two G-quartets of the DNA model reported previously was large and clear. However, it was also possible to fit the NMR structure of the 15-mer (Schultze *et al.*, 1994) in the density, where the quartets of both DNA models overlap closely, even though the two DNA structures have different folding (Figs. 1 and 2). Aside from the chain direction and orientation between the two models, another major difference was in the positioning of the loops connecting the quartets. In the original (and the present) X-ray model, the TGT loop interacts with the fibrinogen recognition site of the thrombin molecule, whereas the TGT loop of the NMR model is close to the heparin binding site of the neighboring symmetry-related thrombin molecule (Fig. 3). Thus, both the aptamer structures bridge neighboring thrombin molecules in the crystal structure to produce infinite chains of thrombin–aptamer molecules.

In order to avoid biasing the structure by refinement using the complete DNA structure of either model, a more cautious and systematic approach was pursued. The bases of six guanines of the quartets were well defined in the initial map. Refinement including these bases reduced the *R* value to 0.24 and the resulting map

Table 1. *R values at different stages of refinement*

% Aptamer	No. of waters	X-ray model		NMR model	
		R	Free *R*	*R*	Free *R*
0	0	0.254	0.260	0.254	0.260
100	68	0.193	0.202	0.188	0.194
100	131	0.162	0.172	0.157	0.161
100	151 (149)*	0.161	0.170	0.155	0.160

* NMR model.

showed better overall density for the DNA and the density for the quartets became outstanding, except for one guanine (G10, X-ray structure). In addition, 68 prominent water molecules could be located in the map of the crystal structure. This model with seven guanines of the quartets and the water molecules refined to an *R* value of 0.22 for data between 10.0 and 2.8 Å. Since the loops are positioned differently between the X-ray and NMR models, further refinement was carried out separately.

The electron-density map at this stage was used to fit both DNA structures completely. From this point, least-squares refinement was carried out using the program *NUCLIN/PROFFT* (Westhof, Dumas & Moras, 1985) and the two DNA models were refined independently along with the thrombin molecule.

2.4. *Refinement using the X-ray structure model*

The original model of the DNA (Padmanabhan *et al.*, 1993) was positioned to match the positions of the quartets that had been determined. The structure was refined progressively including more water molecules. With 151 water molecules, the refinement converged to an *R* value of 0.161 in the resolution range 10.0–2.8 Å.

2.5. *Refinement using the NMR model of the DNA*

The NMR model of the DNA (Schultze *et al.*, 1994) was fitted in the map at $R = 0.22$ to match the quartets of guanines. Refinement of this model progressively including water molecules (149) converged to an *R* value of 0.155, with 90 of the water molecules within 1.0 Å of water molecules in the X-ray structure. From an examination of the *R* values alone of the two different DNA–thrombin complexes, it is not possible to fix with certainty which of the models is correct. Free *R* values randomly omitting 10% of the reflections (Brünger, 1992*b*) were calculated for both the models at several stages of refinement (Table 1). This also did not help in distinguishing between the two DNA models.

The r.m.s. deviation of bond distances from target values is 0.015 and 0.016 Å for the X-ray and NMR structures, respectively; the corresponding deviation in bond angles is 3.9° for both structures. The *R* value of a thrombin complex including both DNA structures simultaneously at half occupancy was a surprising 0.16. Removing the G-quartets from structure-factor calculations increased *R* by about 4.5% while removing the loop structures produced a much smaller affect (~ 2.7%). The *R* value between the calculated structure amplitudes of the two different DNA–thrombin structures was 11.8%; this is practically within expected error limits ($R_{merge} = 9.7\%$ for the observed diffraction data). Thus, the ambiguity between the two DNA models persists throughout all examined aspects of the structure analysis.

3. Results

The DNA aptamer has a folded structure in solution and in the thrombin complex. In both the models, the DNA has two stacked G-quartets, with the T7G8T9 loop on

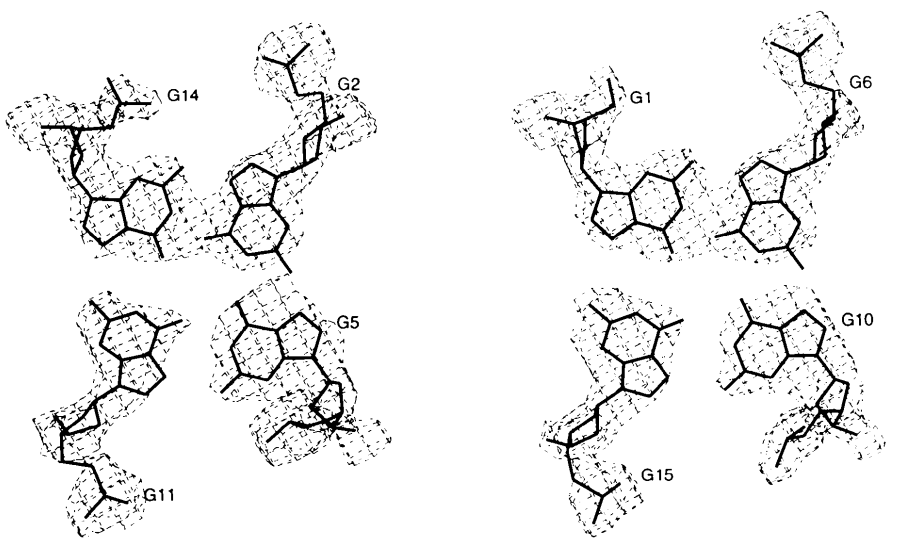

Fig. 2. Fit of Q2 (see Fig. 1) of X-ray (left) and Q1 of NMR (right) structures in ($2F_o - F_c$) electron density contoured at 1σ level.

one side of the quartets and the T3T4 and T12T13 loops on the other side (Fig. 1). The two quartets G1-G6-G10-G15 and G2-G5-G11-G14 will be hereafter referred to as Q1 and Q2, respectively (Figs. 1 and 2). A major difference between the two models in the complex is that, in the NMR structure, the two TT loops interact with the fibrinogen-recognition site of one thrombin molecule and the TGT loop interacts with the heparin-binding site of the neighboring thrombin, whereas in the X-ray structure the opposite occurs. The electron density for the quartets in both the models is well defined, even though they differ by a 90° rotation (Figs. 1 and 2); however, G10 in the X-ray and G14 in the NMR structure are weaker compared to the others. In the NMR model, density for the both thymidines of the TGT loop is poor, whereas the thymidines of the two TT loops are fairly well defined. In

Fig. 3. Stereoview of the relative orientation of the two thrombin molecules in the crystal structure which interact with the DNA. X-ray (top), NMR (bottom); DNA (bold) interacts with fibrinogen exosite of top thrombin and heparin site of bottom thrombin molecule; catalytic triad also bold; only bases with well defined electron density shown.

the X-ray model, T7 and G8 of the TGT loop and T3 and T12 of the TT loops, which extend into the solvent region, have little electron density. In the NMR structure of the aptamer, the two TT loops span the narrow grooves at one end of the quartets and the TGT loop spans the wide groove at other end, whereas the opposite occurs in the X-ray model. Even though there are differences in the loop positions and conformations, the quartets in both the models are practically in the same position at this resolution (r.m.s.$\Delta = 0.8$ Å for 170 common atoms of the eight guanidines).

In the X-ray structure of the DNA, T4 and T13 tend to form a TT base pair (Fig. 4). Moreover, the TT pair stacks with Q2 and T13 is near a hydrophobic patch of a neighboring thrombin molecule involving residues His91, Pro92 and Trp237 of its heparin binding site (Fig. 5). The other thymidine bases (T3,T12) extend toward solvent into weak electron density and cannot interact intramolecularly in the DNA or with thrombin. The G8 base of the TGT loop is questionable but could have a stacking interaction with Q1 while T7 could be buried in a pocket of hydrophobic residues of the fibrinogen exosite of thrombin made by Ile24, His71, Ile79 and Tyr117 (Fig. 5). Although T9 cannot stack within the DNA, it is close to the hydrophobic residues of the putative T7 interaction site.

Except for the TGT and the two TT loops, the quartets of the NMR solution structure of the aptamer are similar to those of the complex (Figs. 2 and 6). The T3 of one TT loop is adjacent to residues Tyr76 and Ile82 of thrombin and T4 has an edge on interaction with Tyr76 (Fig. 7). A TT base pair is formed between T4 and T13, which stacks with Q2 similar to the X-ray structure. However, the atoms to be potentially involved in hydrogen bonds in such a base pair are far apart (~ 5 Å). In the solution NMR structure, in contrast to the thrombin complex, the distances between these atoms are close to hydrogen bonds (Schultze *et al.*, 1994). The other thymidine (T12) of the T12T13 loop is buried at the same hydrophobic pocket formed by thrombin residues Ile24, His71, Ile79

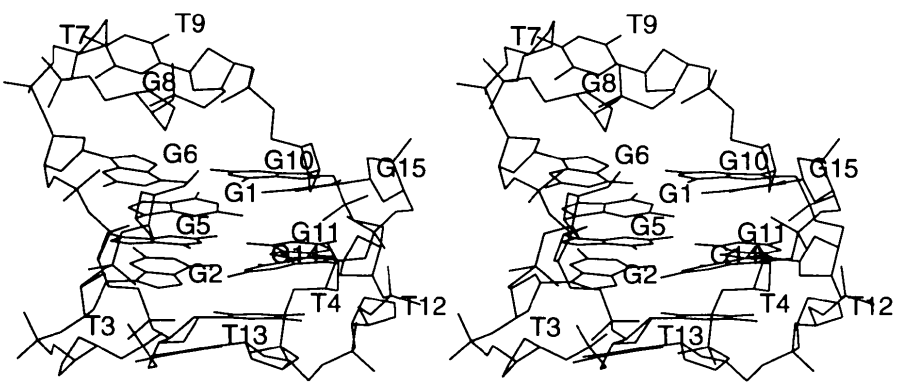

Fig. 4. Stereoview of the X-ray model of the aptamer. Only bases corresponding to well defined electron density shown.

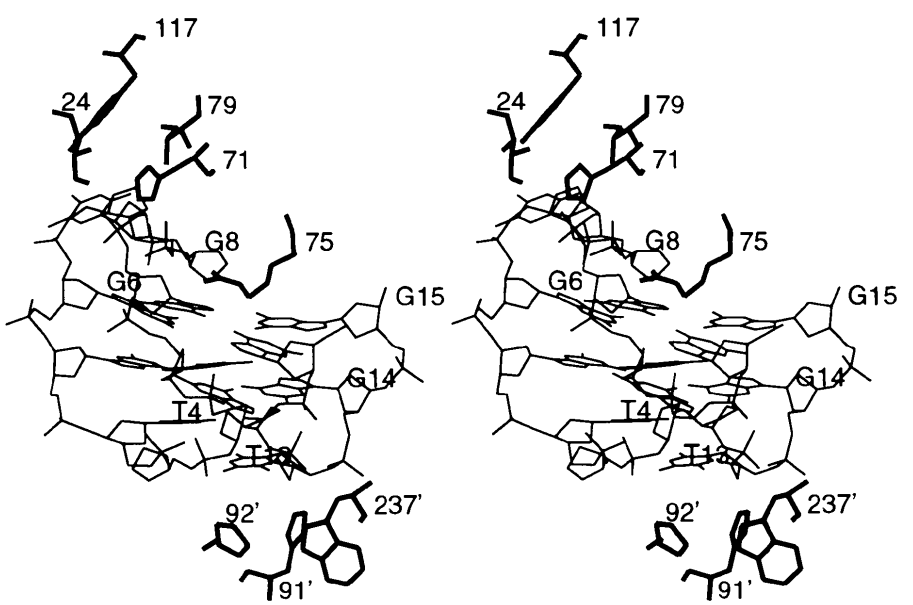

Fig. 5. Stereoview of the hydrophobic interaction site of T13 of the X-ray model of the aptamer. The DNA is shown in thin lines and thrombin residues His91′, Pro92′ and Trp237′, which form the hydrophobic pocket (in bold); primed residues of thrombin indicate the heparin binding-site region of a neighboring crystallographic symmetry-related molecule. Only bases corresponding to good density are shown.

Table 2. *Interactions between thrombin and the DNA 15-mer*

Primes denote symmetry-related molecules.

DNA	Thrombin	Distance (Å)	Interaction
(i) X-ray model			
(*a*) Fibrinogen-recognition exosites			
G8 O2P	His71 NE2	3.9	Ion pair
G8 O1P	Arg75 NH1	3.8	Ion pair
(*b*) Heparin binding site			
T3 O2P	Lys240′ NZ	4.2	Ion pair
T12 O2P	Arg175′ NH2	4.2	Ion pair
T13 O1P	Arg101′ NH1	2.9	Hydrogen bond, ion pair
T13 O2P	Arg101′ NH2	2.6	Hydrogen bond, ion pair
T13 O2P	Asn179′ ND2	2.9	Hydrogen bond
T13	His91′, Pro92′, Trp237′		Hydrophobic
G15 O1P	Arg233′ NE	4.2	Ion pair
(ii) NMR model			
(*a*) Fibrinogen-recognition exosite			
T3	Tyr76, Ile82		Hydrophobic
T12	Ile24, His71, Ile79, Tyr117		Hydrophobic
(*b*) Heparin binding site			
G2 O2P	Arg233′ NE	3.5	Ion pair
G2 O2P	Arg233′ NH2	3.8	Ion pair
G5 O1P	Lys236′ NZ	3.8	Ion pair
G5 O2P	Lys236′ NZ	3.9	Ion pair
T7 O1P	Lys240′ NZ	3.5	Ion pair
T9 O1P	Arg93′ NH2	2.9	Hydrogen bond, ion pair

and Tyr117 that could harbor T7 of the X-ray structure (Fig. 7). Of note is that one thymidine of each TT loop (T3 and T12) does not stack with other DNA bases. The G8 of the TGT loop is also in a hydrophobic pocket formed by His91, Pro92 and Trp237 of a neighboring symmetry-related thrombin molecule, which also can contain T13 of the X-ray structure (Figs. 5 and 7). Finally, G8 and T9, with poor density, of the TGT loop probably stack with Q1, whereas ill defined T7 extends into weak solvent density. The r.m.s. Δ between the solution NMR structure (Schultze *et al.*, 1994) and the NMR model in the DNA–thrombin complex determined here is 1.6 Å for 315 atoms. Thus, if the NMR structure determined here truly represents the aptamer–thrombin complex, then considerable and significant conformational changes occur in the loops of the aptamer accompanying thrombin binding.

4. Discussion

Thrombin catalysis of blood coagulation can be inhibited by either blocking the active site or the fibrinogen-recognition anion-binding exosite (Arg35–Leu41, Lys70–Lys81). Examination of the crystal structure of the DNA complex shows that, in both models, the aptamer interacts with two distinct regions of two thrombin molecules. One of these is the fibrinogen exosite while the other is the heparin binding site near the carboxylate terminal helix of a neighboring thrombin (Fig. 3). The *a* axis of the present crystal is about 2 Å longer than the original crystal (Padmanabhan *et al.*, 1993). This results in a 1.5 Å larger separation of the infinite chains of thrombin–aptamer–thrombin molecules along the *a* direction (Fig. 3).

In the X-ray structure of the DNA–thrombin complex, two residues from the fibrinogen exosite make ion-pair interactions with the 15-mer while four salt bridges form in the heparin site (Table 2). In addition, whereas T7 and T9 may be involved in a hydrophobic cluster in the fibrinogen exosite (Ile24, His71, Ile79 and Tyr117), T13 associates with His91–Pro92 and Trp237 of the heparin site of a symmetry-related thrombin molecule (Fig. 5). From the distribution of the interactions, it would seem that the heparin site would be preferred by the DNA in the X-ray structure.

In the complex with the NMR model of the aptamer, very significantly, there are no ion-pair interactions between the DNA and thrombin in the fibrinogen-binding site, whereas four residues from the heparin-binding site are again involved in good salt bridges with the DNA (Table 2). The major binding interaction in the fibrinogen-binding site appears to be hydrophobic involving T12 (Table 2). To a lesser extent, there are also hydrophobic interactions in the heparin-binding site (Fig. 7). Thus, the

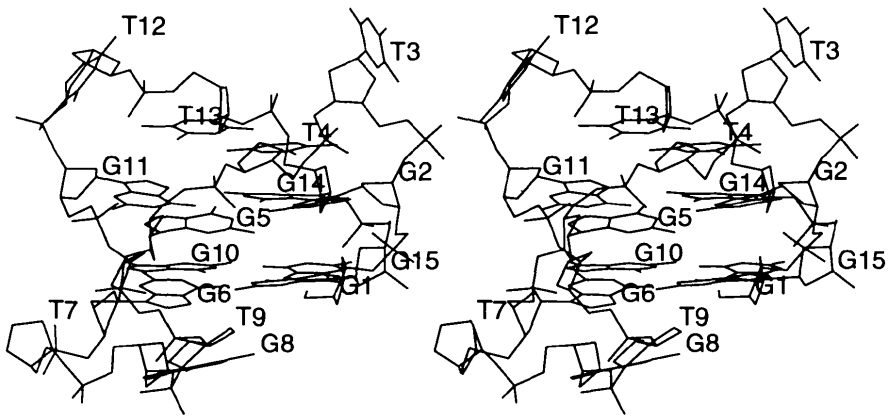

Fig. 6. Stereoview of diffraction data refined NMR model of the aptamer. Orientation with the TT loops at top also the orientation in the crystal compared to the X-ray model (Fig. 4) where the loops are at the bottom. Only bases corresponding to good density shown.

crystallographic result of the NMR structure also indicates that the heparin-binding site should be preferred by the DNA so that both models appear to be somewhat in disagreement with the ancillary evidence indicating the fibrinogen exosite as the primary binding site.

The fibrinogen anion-binding exosite is an important determinant for thrombin recognition of substrates, cofactors and inhibitors. Mutation of Arg73, which is part of the thrombin fibrinogen-recognition exosite, affects interaction with fibrinogen, thrombin platelet receptor and thrombomodulin (Wu *et al.*, 1991). In contrast, mutation of Arg75 affects thrombomodulin and aptamer binding but not fibrinogen clotting and platelet activation (Wu *et al.*, 1991, 1992). In hirudin–thrombin and hirugen–thrombin complexes, Arg75 makes an ion pair with Glu58 of the inhibitor molecule that is bound in a neighboring thrombin but could conceivably achieve the interaction within the same complex in solution (Skrzypczak-Jankun *et al.*, 1991). Mutation of Arg75 reduces the ability of thrombin to bind to single-stranded DNA 15-mer (Wu *et al.*, 1992). Since Arg75 of the X-ray structure interacts with an aptamer phosphate group in the exosite (Table 2) and since the aptamer competes with fibrinogen, His71 and Arg75 must be residues involved in the primary bimolecular binding association in solution. Therefore, both the 15-mer and hirugen inhibit thrombin activity toward macromolecular substrates utilizing the fibrinogen-binding exosite. They do so with similar binding constants (aptamer, 0.1 μM, hirugen, 0.5 μM) (Maraganore, Chao, Joseph, Jablonski & Ramachandran, 1989; Jakubowski & Maraganore, 1990; Bock *et al.*, 1992) but have little effect toward small chromogenic substrates.

In the Arg75 mutation, thrombin was altered to a glutamic acid, which reduced the ability of the aptamer to bind to thrombin (Wu *et al.*, 1992). This could be the result of the disruption of the ion pair between the DNA and Arg75 of thrombin in the X-ray model (Table 2). It could also be related to the possibility that Glu75 could form a salt bridge with Arg77A in the mutated enzyme, which modeling shows would obstruct the bases of either DNA model from reaching the hydrophobic pocket formed by Ile24, His71, Ile79 and Tyr117 (Figs. 5 and 7). The Arg75 mutation is also in agreement with its effects on thrombomodulin binding where Asp416 of thrombomodulin makes a hydrogen-bonded ion pair with Arg75 because the peptide chain runs in the opposite direction (Mathews, Padmanabhan, Tulinsky & Sadler, 1994) to that of hirudin/hirugen and platelet-receptor peptide (Mathews *et al.*, 1994). The other mutation (Lys149EAla) had no effect on the DNA binding. This is most likely because Lys149E is about 15 Å away from Arg75 where DNA binding must occur. It is noteworthy that Arg77A of the region is also the β-autolytic cleavage site of thrombin.

Modeling studies (Wang *et al.*, 1993) assume that the region around the TGT loop of the aptamer interacts with the anion-binding exosite of thrombin and this is consistent with the crystallographic results using the X-ray structure of the DNA. Although chemical modification implicated residues Lys36 and Lys70 to be involved in the aptamer–thrombin interaction (Paborsky *et al.*, 1993), no such interaction has been observed with either model of the aptamer in the crystal structure of the complex. Residues Lys36 and Lys70 participate in fibrinogen–thrombin interaction and these along with four other lysines were shown to be protected from chemical modification following complex formation with hirudin (Chang, 1989). The side chain of Lys70 points towards the center of the loop formed by residues

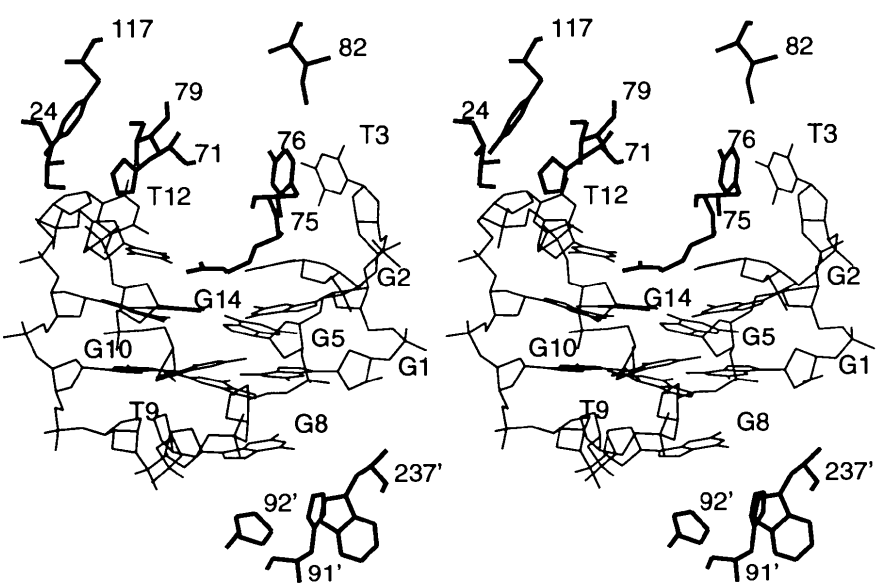

Fig. 7. Stereoview of the hydrophobic interaction sites of T3 and T12 of the NMR model of the aptamer. The DNA is shown in thin lines and thrombin residues, which form the hydrophobic pockets located in the fibrinogen exosite at top (in bold). Heparin binding-site interactions at bottom (primed residues). Only bases with good density shown.

Lys70–Glu80 and occupies a site that is equivalent to the calcium ion position in pancreatic digestive proteinases (Bode & Schwager, 1975). Although there is no direct interaction between the DNA and Lys70, chemical modification may disrupt the Lys70 NE Glu80 OE1 hydrogen-bonded salt bridge (Bode, Turk & Karshikov, 1992) thereby causing a change of the exosite region, which could prevent DNA binding. The observation that the two lysine residues are involved in the aptamer binding agrees generally with the X-ray crystallographic results that both the DNA structures bind to the fibrinogen-recognition exosite where Lys70 is adjacent to His71 (Table 2). The latter forms an ion pair in the X-ray structure of the aptamer or is adjacent to a potentially hydrophobic binding site in the NMR structure.

Additional evidence for the aptamer binding in the fibrinogen-recognition exosite comes from solid-phase plate binding assays (Paborsky *et al.*, 1993). This was carried out using γ-thrombin, an autoproteolyzed derivative of α-thrombin (Fenton, Landis *et al.*, 1977; Rydel *et al.*, 1994), which was assayed for its ability to compete with immobilized α-thrombin for aptamer binding. Even at 2000-fold excess concentration, γ-thrombin, which contains a cleavage within the anion-binding exosite, did not compete with α-thrombin for aptamer binding. More recently, functional mapping of the charged and polar surface residues of thrombin using site-specific mutagenesis showed that the aptamer interacted with Lys70, His71, Arg75, Tyr76, Arg77A (Tsiang *et al.*, 1995). The crystallography implicates His71, Arg75 in the X-ray structure and His71, Arg75, Tyr76 in the NMR structure. All results indicate that the fibrinogen-recognition exosite is of paramount importance for aptamer–thrombin interactions. The DNA 15-mer, therefore, like hirudin carboxylate terminal peptides, inhibits thrombin activity towards macromolecular substrates by blocking the fibrinogen-recognition anion-binding exosite with the binding-site specificity of the DNA apparently related to the folded DNA structure. This display may yet be another example of the tolerance by thrombin, given the opportunity, to bind differently with different substrates (Tulinsky & Qiu, 1993).

The conformation of the DNA aptamer in the crystal structure and its binding to thrombin remains ambiguous since both the X-ray and NMR structures are indistinguishable by usual crystallographic criteria and since the only other physical or chemical evidence related to aptamer–thrombin binding does not implicate DNA components but only those of thrombin and the fibrinogen-recognition site. This is all the more astonishing, and somewhat sobering, since the final R value of the two independent refinements is about 0.16 with only a moderate amount of water structure included. The two different DNA–thrombin structures are not homometric (Pauling & Shappell, 1930; Patterson, 1939, 1944) but rather, cannot be distinguished within the error and extent (resolution) of the present experimental measurement of the diffraction pattern. This is partly due to the limited resolution of the diffraction data (68% complete at 2.8 Å, 38% in the 3.0–2.8 Å shell), which is related to the crystal size and the flexibility of the loops of the aptamer. The congruence of the expected diffraction of the two structures is also related to: (*a*) the large contribution of the thrombin molecule to diffraction, which is practically the same in both cases and (*b*) the near equivalence of the G-quartets in the two structures and their relatively large contribution to diffraction, especially compared with that of the flexible loops of the phosphodiester chain. Possible resolutions of the dilemma could be: identification by chemical or biochemical methods of a base(s) of the aptamer interacting with thrombin or improving the diffraction resolution with larger crystals and/or low-temperature synchrotron measurements.

The NMR structure was marginally better defined in the electron density, having less blatantly missing density (T3, T7, G8 and T12 in X-ray structure) but having poor electron density for T7 and T9. Neither structure, however, can be eliminated from consideration so the binding ambiguity at the fibrinogen site remains: the X-ray structure is consistent with electrostatic interactions in the anion-binding exosite with His71 and Arg75 while the NMR structure, with no charge interactions, is more consistent with hydrophobic recognition by Ile24, His71 and Ile79. Importantly, precedence for the latter occurs in the interaction of the fifth epidermal growth factor (EGF) domain of thrombomodulin with thrombin revealed by alanine-scanning mutagenesis (Nagashima, Lundh, Leonard, Morser & Parkinson, 1993) and by crystallography (Mathews *et al.*, 1994), and to a lesser extent by the Phe56 and Ile59 interactions of hirudin/hirugen and other exosite-binding peptides with thrombin (Qui, Yin, Padmanabhan, Krstenansky & Tulinsky, 1993), which strongly suggest that the specificity of the fibrinogen exosite is principally mediated by hydrophobic association. Whether this will also prove to be the case with 15-mer DNA–thrombin remains to be determined. Another relevant factor of the NMR structure is that the TT loops bind at the exosite rather than the TGT loop. The latter is not conserved among aptamer sequences with binding activity toward thrombin.

The second thrombin site with which DNA interacts in crystals is the heparin-binding exosite of a symmetry-related molecule (Fig. 3). From consensus-sequence considerations (Cardin & Weintraub, 1989), the heparin site of thrombin resides near the carboxylate terminal helix of the B-chain of thrombin, confirmed by site-specific modification of lysyl residues (Church *et al.*, 1989) and mutant studies (Gan *et al.*, 1994; Ye *et al.*, 1994; Sheehan *et al.*, 1994). In addition, on the basis of electrostatic potential-energy calculations (Bode *et al.*, 1992; Karshikov *et al.*, 1992), it has been inferred that

the site includes the 12 positively charged residues between Arg93–Arg97, Arg126, Arg165–Arg175 and His230–Lys240. These charged groups are generally uncompensated and give rise to a large positive electrostatic field, which can associate with highly sulfated negatively charged carbohydrates like heparin. In the crystal structure of the 15-mer thrombin complex, arginine and lysine residues belonging to this region form hydrogen bonds and/or ion pairs with phosphate groups of the aptamer in both models of the DNA (Table 2). Thus, the DNA 15-mer is squeezed between two highly electropositive binding sites of thrombin molecules in the crystal structure giving rise to infinite chains of connected aptamer–thrombin molecules. If it were not for the thrombin mutation experiments, the competition of the 15-mer for the fibrinogen exosite (Wu *et al.*, 1992) and the solid-phase plate binding assay and chemical modification evidence (Paborsky *et al.*, 1993), the solution binding site would most likely have been inferred incorrectly from the crystallographic results alone. In fact, since there is no evidence to the contrary, it is possible that the complex in solution may show association by having some non-systematic secondary interactions with the heparin site of another molecule.

Although the 15-mer sequence had the largest effect on the thrombin clotting time of fibrinogen (prolonged $\times 7$ at 100 n*M* 15-mer), a 6-mer with sequence GGTTGG also inhibited thrombin but at slightly higher concentration (Bock *et al.*, 1992). No inhibition was observed in either case using scrambled sequence controls. We have also grown crystals of the 6-mer sequence complexed with thrombin: monoclinic, $a = 63.1$, $b = 118.1$, $c = 63.1$ Å, $\beta = 104.7°$, space group $P2_1$ (pseudo $C222_1$), four complexes per unit cell, two complexes per asymmetric unit. The crystals, however, were of relatively poor quality and only diffracted X-rays to about 3.8 Å resolution ($R_{\text{merge}} = 8.6\%$, 1/3 observed between 3.8 and 3.5 Å resolution); moreover, the quality could not be improved. Such problems notwithstanding, a rotation/translation search was carried out that gave the orientation and position of the thrombin molecules in the unit cell. The packing arrangement proved to be similar to that of the 15-mer complex with the fibrinogen exosite of one molecule cofacial to the heparin-binding site of a neighboring molecule. Although there was $(2F_o - F_c)$ and $(F_o - F_c)$ electron density between the two sites, it was difficult to interpret in terms of structure. After the structure of the 15-mer complex was solved, it became apparent that in the 6-mer thrombin complex, two DNA 6-mer G-quartets might be cofacially stacked so that the two 6-mers approximate a '15-mer structure' that has two rather than three connecting loops. It was not possible, however, to refine this structure of the complex. A pertinent aspect of such a structure is that the conserved TT loop would bind at the fibrinogen exosite like the NMR structure of the 15-mer.

The formation of G-quartets is generally favored by the presence of monovalent cations (Guschlbauer *et al.*, 1990; Sundquist, 1991; Sen & Gilbert, 1991). The crystal structure of *Oxytricha* telomeric DNA reveals the presence of an axial potassium ion between two G-quartet pairs (Kang *et al.*, 1992). The NMR structures of the 15-mer were determined in solutions containing 0.14 *M* NaCl and 5 m*M* KCl (Wang *et al.*, 1993), and 110 m*M* KCl (Macaya *et al.*, 1993), which may be responsible for the formation of the G-quartet structures. In the present study, the mother liquor of the crystals contained about 0.19 *M* NaCl that is generally used to stabilize thrombin. However, no evidence of a cation between the G-quartets was found in the electron-density maps corresponding to either aptamer structure. The ionic interactions of the complex were generally confined to cationic charges of positively charged side chains of thrombin associating with negatively charged phosphate groups of the DNA 15-mer in the crystal structure. Some of the water molecules in the vicinity of phosphates may be sodium ions, especially since occupancy factors could not be refined with the *NUCLIN/PROFFT* program.*

This work was supported by NIH grant HL 43229. We would like to thank Dr Juli Feigon for providing us with the coordinates of the NMR solution structure, Dr J. Evan Sadler for a sample of the DNA aptamer and Dr Peggy Vanderhoff-Hanaver for growing larger crystals of the thrombin complex.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory [Reference: 1HAO R1HAOSF (X-ray structure) and 1HAP R1HAPSF (NMR structure)]. Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0462).

References

Blackburn, E. H. (1991). *Nature (London)*, **350**, 569–573.
Blackburn, E. H. & Szostak, J. W. (1984). *Ann. Rev. Biochem.* **53**, 163–194.
Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. (1992). *Nature (London)*, **355**, 564–566.
Bode, W. & Schwager, P. (1975). *J. Mol. Biol.* **98**, 693–717.
Bode, W., Turk, D. & Karshikov, A. (1992). *Protein Sci.* **1**, 426–471.
Borzo, M., Detellier, C., Lazlo, P. & Paris, A. (1980). *J. Am. Chem. Soc.* **102**, 1124–1134.
Brünger, A. T. (1990). *Acta Cryst.* A**46**, 46–57.
Brünger, A. T. (1992*a*). *X-PLOR Manual*, Version 3.1, Yale University, New Haven, CT, USA.
Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–475.
Cardin, A. D. & Weintraub, H. J. R. (1989). *Arteriosclerosis*, **9**, 21–32.
Chang, J.-Y. (1989). *J. Biol. Chem.* **264**, 7141–7146.
Church, F. C., Pratt, C. W., Noyes, C. M., Kalayanamit, T., Sherril, G. B., Tobin, R. B. & Meade, J. B. (1989). *J. Biol. Chem.* **264**, 18419–18425.

Ellington, A. D. & Szostak, J. W. (1990). *Nature (London)*, **346**, 818–822.

Fenton, J. W. II, Fasco, M. J., Stackrow, A. B., Aronson, D. L., Young, A. M. & Finlayson, J. S. (1977). *J. Biol. Chem.* **252**, 3587–3598.

Fenton, J. W. II, Landis, B. H., Waltz, D. A. & Finlayson, J. S. (1977). *Chemistry and Biology of Thrombin*, edited by Lundblad, R. L., Fenton II, J. W. & Mann, K. G, pp. 43–70. Ann Arbor Scientific Publishers.

Finzel, B. C. (1987). *J. Appl. Cryst.* **20**, 53–55.

Gan, Z.-R., Li, Y., Chen, Z., Lewis, S. D. & Shafer, J. A. (1994). *J. Biol. Chem.* **269**, 1301–1305.

Gellert, M., Lipsett, M. N. & Davies, D. H. (1962). *Proc. Natl Acad. Sci. USA*, **48**, 2013–2018.

Griffin, L. C., Tidmarsh, G. F., Bock, L. C., Toole, J. J. & Leung, L. L. K. (1993). *Blood*, **81**, 3271–3276.

Guschlbauer, W., Chantot, J. F. & Thiele, D. (1990). *J. Biomol. Struct. Dynam.* **8**, 491–511.

Henderson, E. R., Hardin, C. C., Walk, S. K., Tinoco, I. Jr. & Blackburn, E. H. (1987). *Cell*, **51**, 899–908.

Higashi, T. (1990). *J. Appl. Cryst.* **23**, 253–257.

Jakubowski, J. A. & Maraganore, J. M. (1990). *Blood*, **75**, 399–406.

Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992). *Nature (London)*, **356**, 126–131.

Karshikov, A., Bode, W., Tulinsky, A. & Stone, S. R. (1992). *Protein Sci.* **1**, 727–735.

Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. & Feigon, J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 3745–3749.

Maraganore, J. M., Chao, B., Joseph, M. L., Jablonski, J. & Ramachandran, K. L. (1989). *J. Biol. Chem.* **264**, 8692–8698.

Mathews, I. I., Padmanabhan, K. P., Ganesh, V., Tulinsky, A., Ishii, M., Chen, J., Turck, C. W., Coughlin, S. R. & Fenton, J. W. II (1994). *Biochemistry*, *33*, 3266–3279.

Mathews, I. I., Padmanabhan, K. P., Tulinsky, A. & Sadler, J. E. (1994). *Biochemistry*, **33**, 13547–13552.

Nagashima, M., Lundh, E., Leonard, J. C., Morser, J. & Parkinson, J. F. (1993). *J. Biol. Chem.* **268**, 2888–2892.

Paborsky, L. R., McCurdy, S. N., Griffin, L. C., Toole, J. J. & Leung, L. L. K. (1993). *J. Biol. Chem.* **268**, 20808–20811.

Padmanabhan, K., Padmanabhan, K. P., Ferrara, J. D., Sadler, J. E. & Tulinsky, A. (1993). *J. Biol. Chem.* **268**, 17651–17654.

Patterson, A. L. (1939). *Nature (London)*, **143**, 939–941.

Patterson, A. L. (1944). *Phys. Rev.* **65**, 195–201.

Pauling, L. & Shappell, M. D. (1930). *Z. Kristallogr.* **75**, 128–135.

Pinnavaia, T. J., Marshall, C. L., Mettler, C. M., Fisk, C. L., Miles, H. T. & Becker, E. D. (1978). *J. Am. Chem. Soc.* **100**, 3625–3627.

Pinnavaia, T. J., Miles, H. T. & Becker, E. D. (1975). *J. Am. Chem. Soc.* **97**, 7198–7200.

Qiu, X., Yin, M., Padmanabhan, K. P., Krstenansky, J. L. & Tulinsky, A. (1993). *J. Biol. Chem.* **268**, 20318–20326.

Rydel, T. J., Yin, M., Padmanabhan, K. P., Blankenship, D. T., Cardin, A. D., Correa, P. E., Fenton, J. W. II & Tulinsky, A. (1994). *J. Biol. Chem.* **269**, 22000–22006.

Schultze, P., Macaya, R. F. & Feigon, J. (1994). *J. Mol. Biol.* **235**, 1532–1547.

Sen, D. & Gilbert, W. (1991). *Curr. Opin. Struct. Biol.* **1**, 435–438.

Sheehan, J. P., Tollefsen, D. M. & Sadler, J. E. (1994). *J. Biol. Chem.* **269**, 32747–32751.

Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M. & Maraganore, J. M. (1991). *J. Mol. Biol.* **221**, 1379–1393.

Smith, F. W. & Feigon, J. (1992). *Nature (London)*, **356**, 164–168.

Sundquist, W. I. (1991). *Nucleic Acids in Molecular Biology*, Vol. 5, edited by Lilly, D. M. & Eckstein, F., pp. 1–24. New York: Springer-Verlag.

Szostak, J. W. (1992). *Trends Biochem. Sci.* **17**, 89–93.

Tsiang, M., Jain, A. K., Dunn, K. E., Rojas, M. E., Leung, L. L. K. & Gibbs, C. S. (1995). *J. Biol. Chem.* **270**, 16854–16863.

Tulinsky, A. & Qiu, X. (1993). *Blood Coag. Fibrinolysis*, **4**, 305–312.

Vu, T. K. H., Wheaton, V. I., Huang, D. T., Charo, I. & Coughlin, S. R. (1991). *Nature (London)*, **353**, 674–677.

Wang, K. Y., Krawczyk, S. H., Bischofberger, N., Swaminathan, S. & Bolton, P. H. (1993). *Biochemistry*, **32**, 11285–11292.

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**, 8112–9119.

Wang, Y. & Patel, D. J. (1993). *Structure*, **1**, 263–282.

Wang, Y., de los Santos, C., Gao, X., Greene, K., Live, D. & Patel, D. J. (1991). *J. Mol. Biol.* **222**, 819–832.

Westhof, E., Dumas, P. & Moras, D. (1985). *J. Mol. Biol.* **184**, 119–145.

Williamson, J. R. (1993). *Curr. Opin. Struct. Biol.* **3**, 357–362.

Wu, Q., Sheehan, J. P., Tsiang, M., Lentz, S. R., Birktoft, J. J. & Sadler, J. E. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 6775–6779.

Wu, Q., Tsiang, M. & Sadler, J. E. (1992). *J. Biol. Chem.* **267**, 24408–24412.

Ye, J., Rezaie, A. R. & Esmon, C. T. (1994). *J. Biol. Chem.* **269**, 17965–17970.