Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Duplex–quadruplex motifs in a peculiar structural organization cooperatively contribute to thrombin binding of a DNA aptamer

Irene Russo Krauss,[a] Andrea Pica,[a] Antonello Merlino,[a,b] Lelio Mazzarella[a,b] and Filomena Sica[a,b,c]*

[a]Department of Chemical Sciences, University of Naples 'Federico II', Complesso Universitario di Monte Sant'Angelo, I-80126 Naples, Italy, [b]Institute of Biostructure and Bioimaging, CNR, Via Mezzocannone 16, I-80134 Naples, Italy, and [c]National Institute Biostructures and Biosystems, Inter-University Consortium, Viale Medaglie d'Oro 305, I-00136 Rome, Italy

Correspondence e-mail: filosica@unina.it

Received 24 June 2013
Accepted 8 August 2013

PDB Reference: thrombin–HD22-27mer, 4i7y

Potent second-generation thrombin aptamers adopt a duplex–quadruplex bimodular folding and recognize thrombin exosite II with very high affinity and specificity. A sound model of these oligonucleotides, either free or in complex with thrombin, is not yet available. Here, a structural study of one of these aptamers, HD22-27mer, is presented. The crystal structure of this aptamer in complex with thrombin displays a novel architecture in which the helical stem is enchained to a pseudo-G-quadruplex. The results also underline the role of the residues that join the duplex and quadruplex motifs and control their recruitment in thrombin binding.

1. Introduction

Aptamers are single-stranded DNA or RNA sequences generated through *in vitro* selection which inherently adopt stable three-dimensional structures that have a good shape complementarity with many different molecular targets (small molecules, natural and synthetic amino acids and peptides, large biomacromolecules and also cells). This intrinsic property makes aptamers efficient binding molecules capable of recognizing their targets with high affinity and specificity (Keefe *et al.*, 2010; Tan *et al.*, 2011). For these reasons, aptamers are considered a kind of nucleic acid version of antibodies with improved properties, such as greater thermal stability, tolerance to wide ranges of pH and salt concentration, and simplicity of their synthetic routes (Jayasena, 1999). Among the large number of structural arrangements adopted by aptamers, the four-stranded topology known as the G-quadruplex has become of great interest, since it has been implicated in key biological processes (Patel *et al.*, 2007; Lipps & Rhodes, 2009; Fernando *et al.*, 2009; Kumari *et al.*, 2007) and represents an attractive molecular scaffold for use in drug design (Balasubramanian & Neidle, 2009). The main component of the G-quadruplex is the G-tetrad or G-quartet, a roughly planar arrangement of four guanine bases associated through a cyclic array of Hoogsteen-like hydrogen bonds, in which each guanine base both accepts and donates two hydrogen bonds. The G-quadruplex core is usually stabilized by the stacking of two or more G-tetrads and by the interaction of cations with the O6 atoms of guanines belonging to the cyclic array (Williamson, 1994). Within the rather rigid structure of the G-quadruplex core, great topological variation is allowed. The overall G-quadruplex arrangement is critically dependent on the nature of the counter-ions, the number and the orientation of the strands, the conformation of the guanine glycosidic bond angles, and the conformation of the loops connecting the guanines of the tetrad.

Several examples of aptamers that are supposed to adopt a mixed duplex–quadruplex structure have been reported in the literature (Tasset *et al.*, 1997; Chinnapen & Sen, 2002; Pileur *et al.*, 2003; Mori *et al.*, 2004; Chou *et al.*, 2005; Li *et al.*, 2009). In all of these cases the presence of both structural motifs is required for strong and specific binding to the target. The two folded elements are characterized by a large difference in shape and electrostatic potential. Indeed, with respect to a double helix, a G-quadruplex exhibits a substantially higher negative charge density that could strengthen the specificity in target recognition. To date, no crystallographic data exist on bimodular aptamers and their complexes with targets, although they are particularly important because of their ability to recognize proteins involved in several pathologies, such as AIDS, thrombosis, tumours *etc.* (Gatto *et al.*, 2009). Only recently has the structure of an RNA molecule with a mixed duplex–quadruplex fold in complex with a peptide from the human fragile X mental retardation protein been solved by NMR (Phan *et al.*, 2011).

One of the most interesting results of SELEX protocols directed towards human α-thrombin (thrombin) was the identification of potent second-generation aptamers that are supposed to adopt a mixed duplex–quadruplex conformation (Tasset *et al.*, 1997). In particular, this family of aptamers includes a 29-mer, named HD22-29mer (5′-AGTCCGT**GG-TAGGGCAGGTTGG**GGTGACT-3′), and a 27-mer (HD22-27mer), which lacks the first and the last residue with respect to HD22-29mer. A schematic model of these bimodular oligonucleotides is shown in Fig. 1(*a*) (Tasset *et al.*, 1997). Both aptamers embody a 15-residue sequence (indicated in bold) very similar to that of the best characterized thrombin aptamer TBA (5′-GGTTGGTGTGGTTGG-3′) and bind the protein with higher affinity [$K_d \simeq 0.5$ and 0.7 n*M* for the 29-mer and 27-mer aptamers, respectively (Macaya *et al.*, 1995; Tasset *et al.*, 1997)] compared with TBA [$K_d \simeq 100$ n*M* (Macaya *et al.*, 1995; Nagatoishi *et al.*, 2011)]. Structural studies have shown that TBA and its derivatives (TBAs) adopt an antiparallel chair-like G-quadruplex structure (Macaya *et al.*, 1993; Wang *et al.*, 1993). More recently, high-resolution X-ray data have shown that these aptamers bind to thrombin by embracing the protruding region of exosite I through their TT loops (Russo Krauss *et al.*, 2011, 2012). On the other hand, several results indicate that HD22 aptamers recognize exosite II rather than exosite I (Tasset *et al.*, 1997). However, high-resolution structural data on these aptamers and on their complexes with thrombin are still lacking.

Here we present the crystallographic analysis of the thrombin–HD22-27mer complex and the results of a spectroscopic study of the aptamer, both free and in complex with thrombin. The X-ray model of the thrombin–HD22-27mer complex at 2.4 Å resolution revealed interesting new features regarding the overall fold of the aptamer and how it interacts with its target. In particular, the aptamer shows a spatial organization that differs from the proposed models (Marson *et al.*, 2012; Mayer *et al.*, 2009; Müller *et al.*, 2008; Tasset *et al.*, 1997; Zhou *et al.*, 2010; Spiridonova *et al.*, 2003; Hasegawa *et al.*, 2008), such as that sketched in Fig. 1(*a*), that are characterized by a definite separation between the duplex and the quadruplex regions. In the present structure the duplex is directly enchained to a pseudo-G-quadruplex topology that has never been observed. This structural arrangement optimizes the adhesion of HD22-27mer on thrombin exosite II. Circular-dichroism (CD) measurements also suggest that the crystal structure is representative of the solution structure of the aptamer.

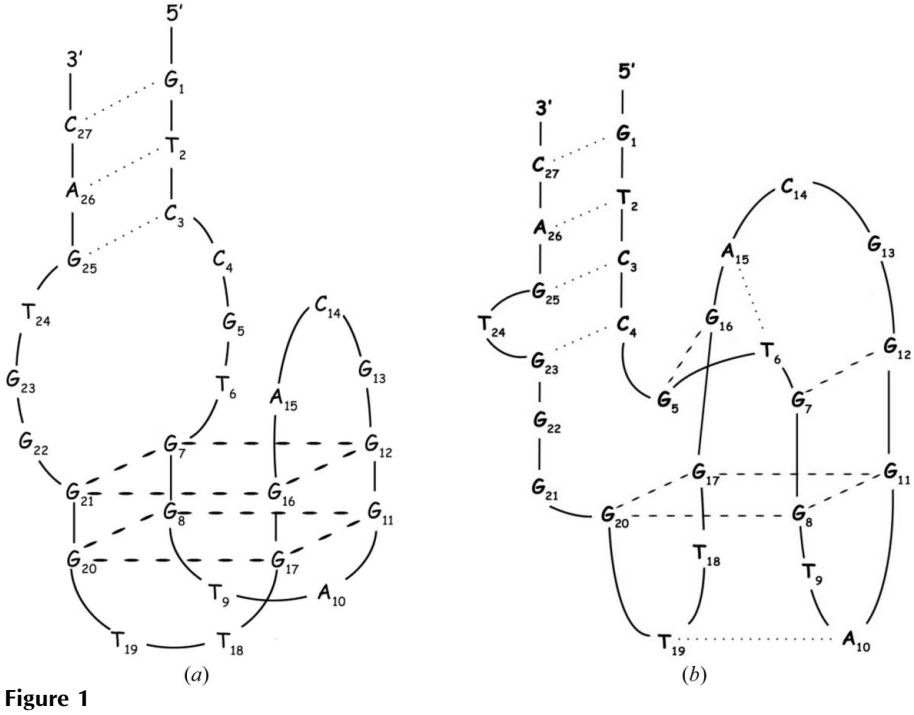

Figure 1
Schematic representations of the HD22-27mer organization (*a*) as proposed by Tasset *et al.* (1997) and (*b*) as found in the crystal structure of the complex with thrombin. Hydrogen bonds between bases are represented by dotted lines (dashed lines represent G–G interactions).

2. Material and methods

2.1. Sample preparation

The HD22-27mer aptamer was purchased from Genosys. Stock solutions were prepared by dissolving the lyophilized oligonucleotide in 10 m*M* sodium phosphate buffer pH 7.1 to a concentration of 4 m*M*. Nucleotide samples were heated at 358 K for 10 min and then cooled and stored for one night at 277 K.

Human D-Phe-Pro-Arg-chloromethyl ketone (PPACK)-inhibited α-thrombin was purchased from Haemtech and the initial buffer was changed to 0.75 *M* NaCl using a Centricon mini-concentrator and a refrigerated centrifuge. The protocol described by Russo Krauss *et al.* (2010) was used to

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell. One crystal was used for data collection.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 48.14, b = 81.90, c = 53.72, β = 99.7
Resolution (Å)	50.0–2.40 (2.49–2.40)
$R_{\rm merge}$ (%)†	9.9 (57.8)
$\langle I/\sigma(I)\rangle$	13.8 (2.0)
No. of observations	77242
No. of unique reflections	16175
Completeness (%)	99.8 (98.9)
Average multiplicity	4.8 (4.1)
$V_{\rm M}$ (Å^3 Da^{-1})	2.4
Solvent content (%)	53.4

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith intensity measurement of the reflection hkl, including symmetry-related reflections, and $\langle I(hkl)\rangle$ is its average.

prepare the thrombin–HD22-27mer complex. Briefly, a twofold molar excess of the annealed aptamer solution was placed on a frozen sample of inhibited thrombin and the sample was left for 3 h at 277 K. It was then diluted and the buffer was changed to 25 m*M* sodium phosphate buffer pH 7.1, 0.1 *M* NaCl. The solution was extensively washed to remove excess aptamer and was finally concentrated to about 0.2 m*M* using a Centricon mini-concentrator and a refrigerated centrifuge.

2.2. Crystallization and data collection

Crystallization conditions for the thrombin–HD22-27mer complex were identified after extensive screening using commercial kits and an automated crystallization system. The starting crystallization solution consisted of 20%(*w*/*v*) PEG 3350, 0.2 *M* sodium citrate. Optimization of the crystallization conditions produced well diffracting crystals by the hanging-drop vapour-diffusion method in 1 µl drops with a protein: reservoir solution ratio of 1:1 using a protein concentration of 8 mg ml^{-1}. The best crystals grew from solutions containing a lower concentration of PEG 3350 [14–16%(*w*/*v*)] with respect to the starting concentration. After the addition of 35% glycerol to the harvesting solution, crystals were flash-cooled at 100 K in supercooled N_2 gas produced by an Oxford Cryosystem and were maintained at 100 K during data collection. Diffraction data were collected using synchrotron radiation (wavelength 1.2651 Å) at Elettra, Trieste, Italy and were indexed, processed and scaled with *HKL*-2000 (Otwinowski & Minor, 1997). The crystals belonged to the monoclinic space group $P2_1$ and diffracted to 2.4 Å resolution. Matthews coefficient calculations suggested the presence of a 1:1 complex in the asymmetric unit. Detailed statistics of data collection are reported in Table 1.

2.3. Structure determination and refinement

The structure of the complex between HD22-27mer and thrombin was solved by the molecular replacement method using *Phaser* (McCoy *et al.*, 2005) with the coordinates of both inhibited thrombin (derived from the structure of its complex with TBA; PDB entry 4dih; Russo Krauss *et al.*, 2012) and a DNA duplex (derived from the structure of a DNA dodecamer duplex; PDB entry 1lu5; Silverman *et al.*, 2002) as search models. To avoid bias, PPACK, ions and water molecules were removed from the model. A clear solution was obtained with a log-likelihood gain (LLG) of 1511. The starting model was subjected to several cycles of rigid-body refinement followed by several cycles of coordinate minimization and *B*-factor refinement using *CNS* and *REFMAC*5 (Brünger *et al.*, 1998; Murshudov *et al.*, 2011). Each run was alternated with manual model building using *Coot* (Emsley *et al.*, 2010). Fourier difference maps calculated with ($F_{\rm o} - F_{\rm c}$) and ($2F_{\rm o} - F_{\rm c}$) coefficients showed continuous electron density in the proximity of the DNA duplex segment. The

Table 2
Summary of refinement statistics.

Resolution (Å)	50.00–2.40 (2.49–2.40)
No. of reflections	14526
$R_{\rm work}/R_{\rm free}$ (%)	18.1/24.9
No. of atoms	
Total	2951
Protein	2288
Aptamer	564
Ions	1
Water	98
Average B factor (Å^2)	
Overall	50.1
Protein	47.2
Aptamer	62.2
Ion	46.8
Water	48.3
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.687
Ramachandran plot, residues in (%)	
Most favoured region	91.9
Additionally allowed region	8.1
Generously allowed region	0

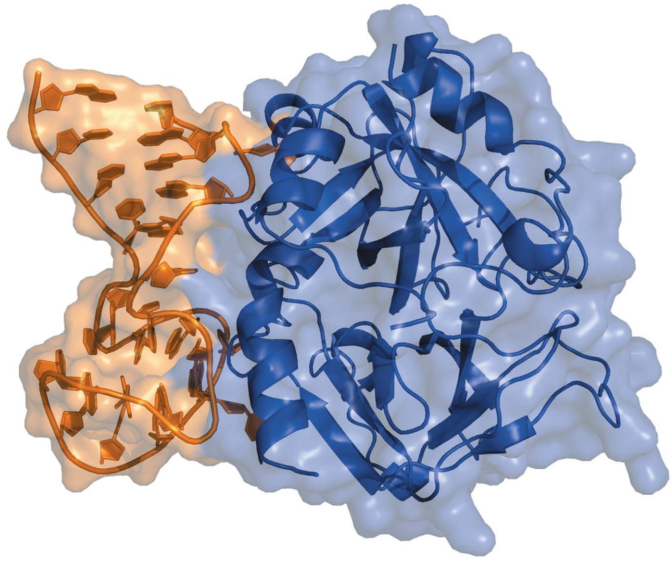

Figure 2
Overall structure of the complex between thrombin and HD22-27mer. Surface representation of the complex, with thrombin coloured blue and HD22-27mer coloured orange.

analysis of these maps calculated at various stages of refinement allowed the building of the whole aptamer molecule, as well as the fitting of PPACK into the active site, the identification of a glycosylation site and the positioning of several water molecules. After the inclusion of low-resolution data and bulk-solvent correction, the crystallographic R factor and $R_{\rm free}$ for the final model of the complex between thrombin and HD22-27mer in the resolution range 50.0–2.40 Å were 18.1 and 24.9%, respectively. At the end of refinement the geometry of the protein structure was checked using *PROCHECK* (Laskowski *et al.*, 1993) and *WHATCHECK* (Hooft *et al.*, 1996). A full list of refinement statistics is reported in Table 2. The figures were prepared with *PyMOL* (http://pymol.org). The coordinates of the structure have been deposited in the Protein Data Bank (PDB entry 4i7y).

2.4. CD measurements

Circular dichroism (CD) spectra were recorded at 283 K using a Jasco J-710 spectropolarimeter equipped with a Peltier thermostatic cell holder (Model PTC-348WI). CD measurements in the 200–350 nm range were carried out using a 0.1 cm path-length cell and a 40 μ*M* solution of the free aptamer. We also recorded CD spectra of the aptamer after the addition of an equimolar amount of thrombin and the addition of NaCl and phosphate buffer to final concentrations of 0.1 and 0.025 *M*, respectively. Before measurements, the samples were pre-equilibrated at 283 K for 5 min and the instrument was calibrated with an aqueous solution of D-10-(+)-camphorsulfonic acid at 290 nm.

Thermal unfolding curves were recorded in the 283–363 K range in 1 K steps with 0.5 min equilibration time between readings at a heating rate of 1 K min^{-1}. Transition temperatures were calculated from the second derivative of the ellipticity change *versus* temperature.

3. Results

3.1. Overall crystal structure

The choice of a proper protocol (Russo Krauss *et al.*, 2013) in the preparation of the thrombin–HD22-27mer complex was crucial for the success of the crystallization experiments. The

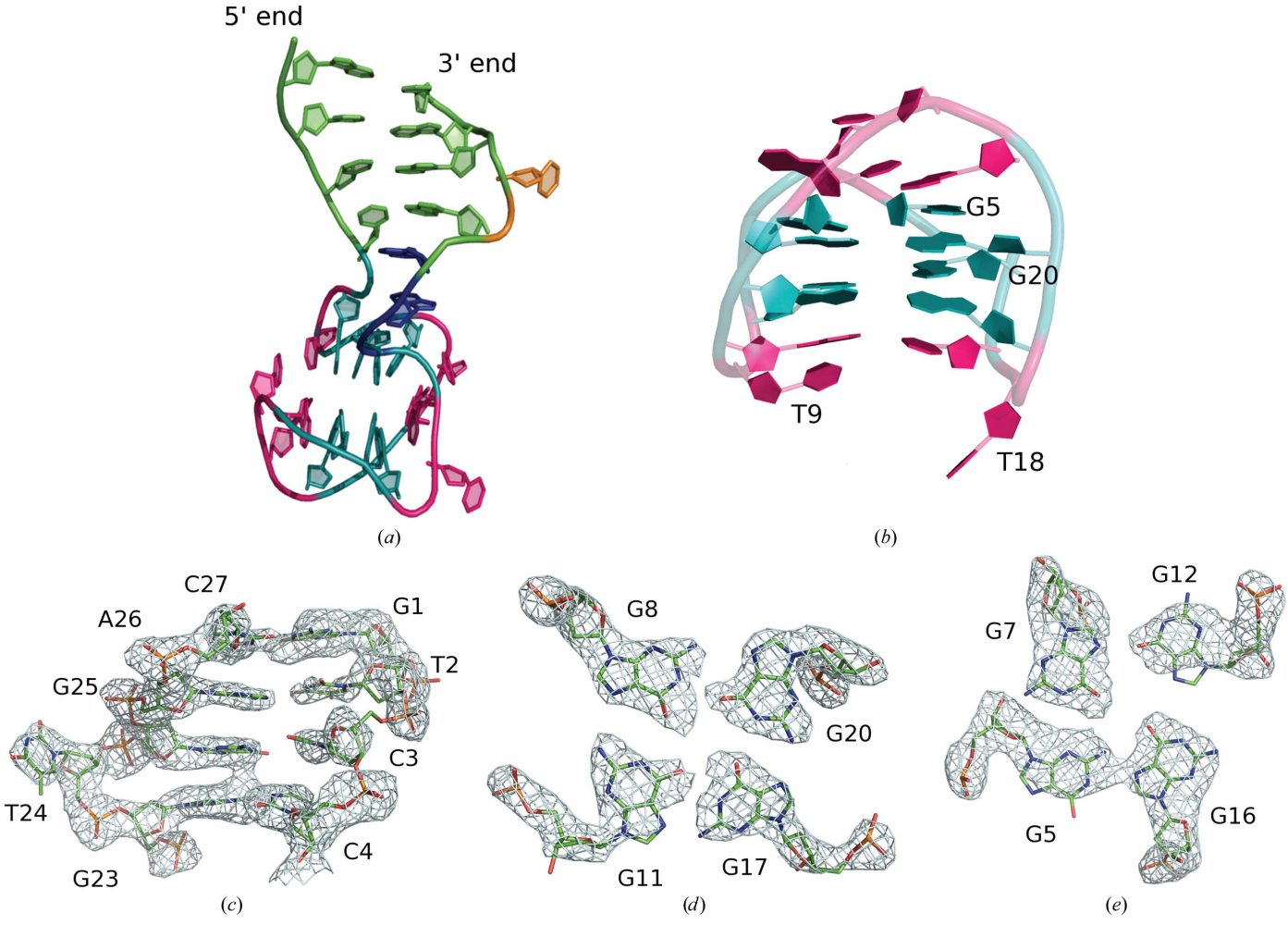

Figure 3
Crystal structure of the HD22-27mer aptamer. (*a*) Cartoon representation: the duplex region is coloured green, the G-core cyan, the loops pink, the connecting residues blue and the bulged-out thymine orange. (*b*) A different view of the pseudo-G-quadruplex architecture. $F_{\rm o} - F_{\rm c}$ unbiased OMIT simulated-annealing maps contoured at the 3.0σ level are shown for the duplex motif (*c*), for the G-tetrad (*d*) and for the pseudo-G-tetrad (*e*).

best crystals of the complex between thrombin and HD22-27mer diffracted X-rays to 2.4 Å resolution and belonged to the monoclinic space group $P2_1$. Electron-density maps calculated from the initial phases obtained by molecular replacement allowed the unambiguous model building of the whole DNA molecule. For thrombin, the heavy chain (residues 16–245) and the light chain (residues 1B–14K) were well defined in the electron-density map, with the exception of the γ-autolysis loop (residues 146–150) of the heavy chain. Crystallographic analysis shows that the HD22-27mer aptamer binds the thrombin molecule in the exosite II region, forming a 1:1 complex (Fig. 2). Interestingly, the ligand does not adopt a simple duplex or quadruplex fold, but presents a mixed structure (Fig. 1*b*). The crystal structure was refined to an R factor and an $R_{\rm free}$ of 18.1 and 24.9%, respectively. Detailed statistics of the refinement are reported in Table 2.

3.2. An aptamer with new structural features

X-ray analysis of HD22-27mer in complex with thrombin reveals that the oligonucleotide fold (Figs. 1*b* and 3*a*) is considerably different from the model sketched in Fig. 1(*a*). Residues 1–3 and 25–27 form the duplex motif; however, with respect to the proposed model, the crystal structure contains an additional Cyt4–Gua23 Watson–Crick base pair (Fig. 3*a*). This is favoured by the bulging out of Thy24, which is inserted into a pocket on the thrombin surface (see below). The duplex structure is followed by 16 residues (sequence number 5–20) organized into a G-tetrad core capped by the two-residue loops Thy9-Ade10 and Thy18-Thy19 on one side and the three-residue loop Gua13-Cyt14-Ade15 on the opposite side (Fig. 3*b*). Electron-density maps corresponding to the two structural motifs, duplex and quadruplex, are shown in Figs. 3(*c*), 3(*d*) and 3(*e*). It has to be noted that, on the basis of the sequence alignment between HD22-27mer and TBA, the G-quadruplex arrangement was expected to involve the 15 residues 7–21. The shape of the HD22-27mer G-tetrad core is roughly similar to that of TBA, with the extra residue Thy6 forming a one-residue loop adjacent to the three-residue loop (Figs. 1*b* and 3*b*). However, the chain topology is markedly different. In the core region the four guanines Gua8, Gua11, Gua17 and Gua20 define a well formed G-tetrad (Figs. 3*d* and 4*a*): it presents the classical alternation of *syn* and *anti* conformations of the guanine residues (with Gua8 and Gua17 adopting the *anti* conformation and Gua11 and Gua20 adopting the *syn* conformation) and the nucleobases are only slightly tilted out from the average plane of the tetrad. The remaining tetrad (Figs. 3*e* and 4*b*) begins with Gua5, which is located two residues upstream with respect to the putative TBA segment. This tetrad does not possess the canonical *anti–syn–anti–syn* but an *anti–syn–anti–anti* pattern of glycosidic angles (only Gua12 adopts a *syn* conformation). Moreover, the presence of a short one-residue loop slightly pushes apart the backbone atoms of Gua5 and Gua7, causing a

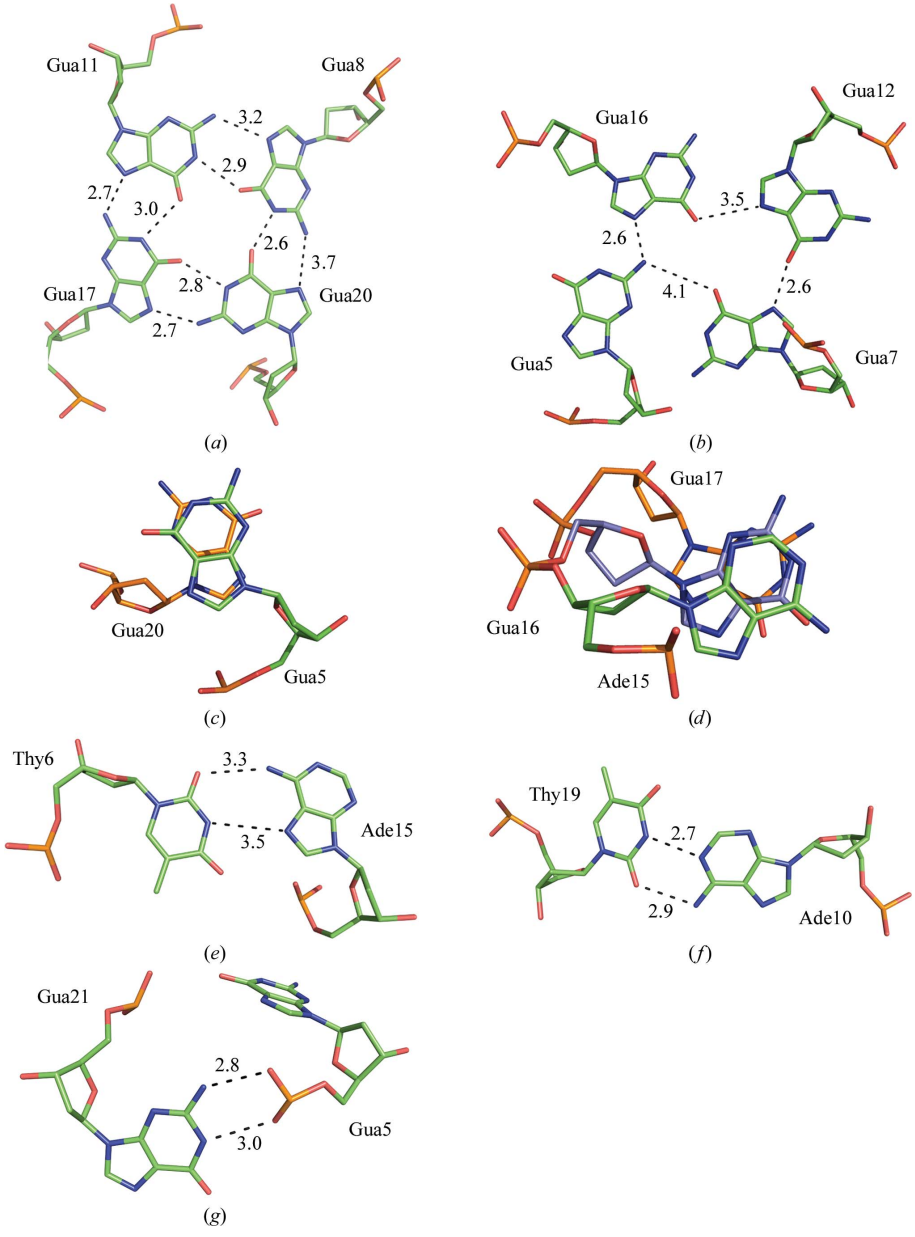

Figure 4
Snapshots of aptamer structural motifs. (*a*) G-tetrad, (*b*) pseudo-G-tetrad, (*c*) and (*d*) stacked bases, (*e*) *trans* Watson–Crick/Hoogsteen pair, (*f*) *trans* Watson–Crick/Watson–Crick pair and (*g*) G-fork. The distances reported are in Å.

Table 3
Interactions between thrombin and HD22-27mer.

Aptamer residue	Protein residue
Polar contacts	
Thy9 OP1	Asn95 ND2
Thy9 OP2	Arg97 NH1
Thy9 N1	Tyr94 O
Thy9 O2	Tyr94 O
Thy9 N3	Pro92 O
Thy18 O2	Trp237 NE1
Gua20 OP1	Arg93 NH2
Gua20 O4′	Arg93 NH1
Gua20 O5′	Arg93 NH1
Gua23 O4′	Arg233 NH1
Thy24 OP1	His230 NE2
Thy24 OP1	Arg233 NH1
Thy24 O2	Leu130 O
Thy24 O4′	Leu130 O
Gua25 OP2	Arg126 NH2
Cyt27 OP1	Lys169 NZ
Hydrophobic contacts	
Thy9	Arg93, Asn95, Trp96, Arg97
Thy18	Tyr89, Pro92, Trp237, Val241, Phe245
Thy19	His91, Pro92, Arg93, Trp237
Thy24	His230
Ade26	Arg165

Table 4
Thrombin–aptamer interface features.

Structure	HD22-27mer	TBA	Toggle-25
PDB code	4i7y	4dih	3dd2
Interface area (Å^2)	1118.5	563.2	754.5
Polar interface area (Å^2)	515.4	287.6	279.7
Nonpolar interface area (Å^2)	603.1	275.6	474.8
Residues at the interface			
Complex	55	22	32
Aptamer	15	8	10
Thrombin	40	14	22
Surface-complementarity index†	0.64	0.76	0.72

† See Russo Krauss *et al.* (2011).

break in the Hoogsteen hydrogen-bond cyclic pattern. Thus, the four guanine residues Gua5, Gua7, Gua12 and Gua16 can be better described as forming two pairs: in one pair Gua5 and Gua16 are linked through a reversed Hoogsteen N2–N7 hydrogen bond, whereas in the second pair only a loose polar contact between O6 (Gua12) and N7 (Gua7) is present (Fig. 4*b*). Despite these modifications, an almost perfect stacking of Gua5 and Gua20 (Fig. 4*c*) and of Gua16 and Gua17 (Fig. 4*d*) is observed. The latter pair also makes a good stacking interaction with Ade15 belonging to the three-residue loop. We shall call this unusual four-guanine set a pseudo-G-tetrad and the overall organization of residues 5–20 a pseudo-G-quadruplex. Interestingly, base pairing is also observed for the residues of the loops. In particular, Thy6–Ade15 form a *trans* Watson–Crick/Hoogsteen pair (Fig. 4*e*) and Ade10–Thy19 a *trans* Watson–Crick/Watson–Crick pair (Fig. 4*f*). The sharp transition from the duplex motif to the pseudo-G-quadruplex motif produces a somewhat strained kink in the molecule (Fig. 3*a*). Indeed, Gua5, which is the first residue of the pseudo-G-quadruplex segment, is also linked to Cyt4, the last residue of the duplex. Thus, Gua5 can be seen as the pivot residue that mediates the conversion from a duplex to a quadruplex arrangement. In the facing strand, the hinge segment between the two motifs includes two residues: Gua22, which does not interact with other residues of the aptamer, and Gua21, which forms a G-fork with Gua5 (Fig. 4*g*). As a result of the different numbers of residues involved in the loops connecting the duplex–quadruplex structures, the helical axes of the two motifs form an angle of approximately 90°. Interestingly, no electron density assignable to a cation is observed in this pseudo-G-quadruplex region.

3.3. Interaction with thrombin

The unusual mixed structure of HD22-27mer extensively adheres to the thrombin surface (Fig. 2). In particular, the aptamer binds to exosite II, burying a total of 1118 Å^2 of the protein accessible surface. Interaction between HD22-27mer and thrombin involves numerous residues of both molecules: Thy9, Thy18, Thy19, Gua20, Gua23, Thy24, Cyt27 and marginally Gua21, Gua22 and Ade26 of the aptamer and segments 89–101, 230–245 of thrombin, with a further contribution from residues Leu130, Arg165 and Lys169. A full list of contacts between thrombin and HD22-27mer is reported in Table 3, while a few examples of these contacts are presented in Fig. 5. As expected on the basis of the strong electropositive potential of exosite II, many polar contacts are observed at the interface. In particular, ion pairs are formed between the phosphate backbone of the nucleotide and positively charged residues of thrombin. Hydrophobic contacts, mainly involving loop residues Thy9, Thy18 and Thy19 and Ade26, also contribute to the stability of the complex. A further anchorage is represented by Thy24, which bulges out from the duplex region of the nucleotide into a protein pocket (Fig. 5*d*), where it is mainly involved in polar contacts.

The thrombin–HD22-27mer complex (Supplementary Figs. S1*a* and S1*d*[1]) was compared with two other thrombin–aptamer complexes: that with TBA (PDB entry 4dih; Russo Krauss *et al.*, 2012), which binds to exosite I (Supplementary Figs. S1*b* and S1*d*), and that with Toggle-25t (PDB entry 3dd2; Long *et al.*, 2008), an RNA aptamer that embodies two short duplex regions and binds to exosite II (Supplementary Figs. S1*c* and S1*d*). In particular, using the *COCOMAPS* tool (https://www.molnac.unisa.it/BioTools/cocomaps/), the protein–aptamer interfaces have been compared (Table 4). The much higher affinity of HD22-27mer (Tasset *et al.*, 1997) with respect to TBA (Macaya *et al.*, 1995; Nagatoishi *et al.*, 2011), which binds to exosite I, correlates very well with the almost doubled value of the interface area and of the number of interacting residues displayed by the HD22-27mer complex. A similar good correlation is also found when the comparison is performed with Toggle-25t, which recognizes exosite II with a K_d of about 3 n*M* (Long *et al.*, 2008).

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5101). Services for accessing this material are described at the back of the journal.

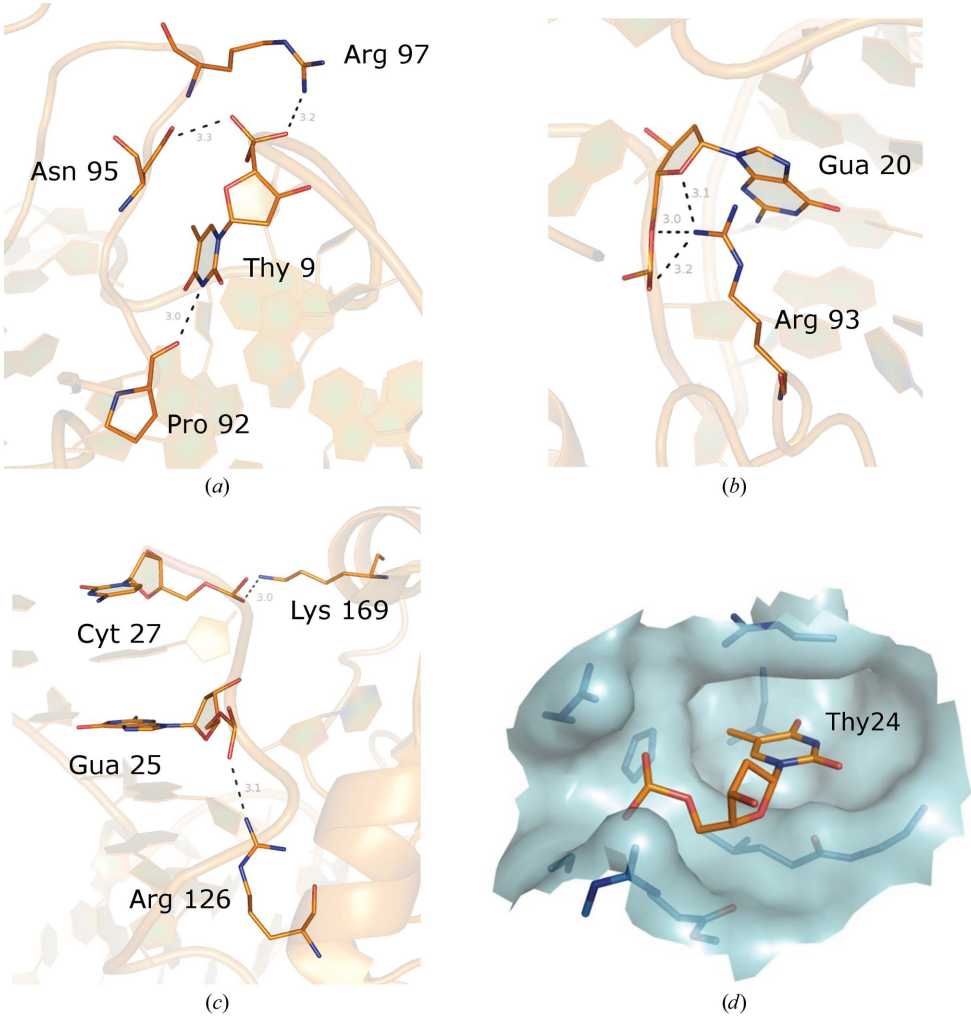

Figure 5
Some examples of the interactions between thrombin and HD22-27mer focusing on Thy9 (*a*), Gua20 (*b*), Gua25 and Cyt27 (*c*). (*d*) Insertion of HD22-27mer Thy24 into a pocket on the thrombin surface. The aptamer residue is represented as sticks and the thrombin molecule as a surface.

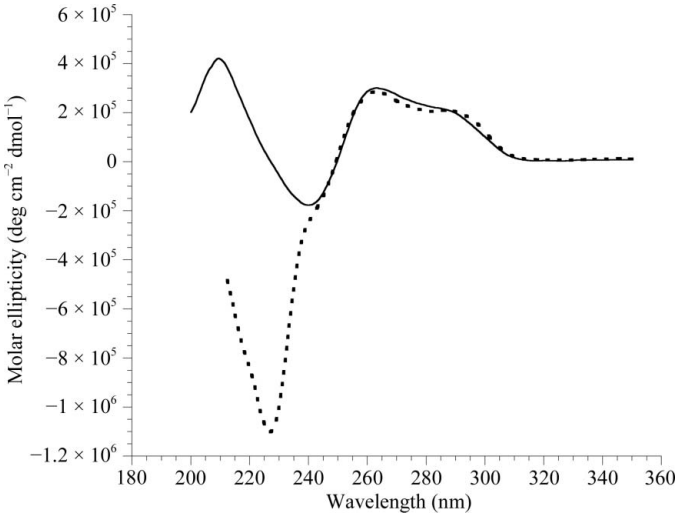

Figure 6
CD spectra of HD22-27mer before (bold line) and after (dashed line) the addition of thrombin.

3.4. Circular-dichroism study

The conformation and thermal stability of the HD22-27mer aptamer were investigated by means of circular dichroism. The CD spectra show two overlapping positive signals at ∼288 and ∼260 nm and a negative band at ∼240 nm (Fig. 6, bold line). Thrombin binding induces small variations in the spectral features of the oligonucleotide (Fig. 6, dashed line) but significantly affects the stability of the aptamer conformation, the transition temperature of which increases from 309 to 321 K (Supplementary Figs. S2 and S3). A similar thrombin-induced stabilization was previously observed in the case of TBA (Nagatoishi *et al.*, 2011; Russo Krauss *et al.*, 2012).

4. Discussion

Crystallographic analysis of the complex with thrombin shows that HD22-27mer adopts a novel sharply kinked conformation in which the helical axis of the regular duplex segment and that of the pseudo-G-quadruplex motif are approximately at right angles. The resulting overall shape of the molecule allows both motifs to interact with the protein (see below). The pseudo-G-quadruplex segment lacks the stabilizing effects produced by cyclic Hoogsteen hydrogen bonds in one of the two tetrads and by the cation binding (Williamson, 1994), but preserves the base stacking among the guanines of the core (Figs. 4*c* and 4*d*). Indeed, although the pivot residue Gua5 is displaced from the position required for the formation of a canonical G-tetrad, the *anti* conformation of this residue efficiently places its base in a position to build up the close packing of the eight guanine bases of the core. Moreover, hydrogen bonds between the bases in loops (Figs. 4*e* and 4*f*) also contribute to the intramolecular stabilization of this fold. Thus, it may be surmised that the aptamer conformation found in the crystal structure of the complex with thrombin represents one of the low-energy conformations adopted by the free nucleotide. Incidentally, the observed structure of HD22-27mer can also be representative of the longer HD22-29mer conformation, where the addition of the Ade and Thy residues at the 5′ and 3′ ends, respectively, should not significantly alter the aptamer structure and should only be expected to further extend the

duplex segment. The protein–aptamer interaction shows interesting new features. In the known crystallographic structures of TBA complexes (Russo Krauss *et al.*, 2011, 2012) the formation of the complex is essentially driven by the TT loops of the quadruplex. By contrast, in the present structure the protein recognition involves an extended region of the aptamer that includes the duplex segment (Gua23, Ade26 and Cyt27), the bulged-out residue Thy24 and the pseudo-G-quadruplex core (Gua20), as well as the quadruplex-connecting loops (Thy9, Thy18 and Thy19). Despite the lack of a well defined binding structural motif (such as the TT loops of TBAs), this large contact area (1118 Å^2) displays high complementarity with the thrombin surface. Indeed, the Sc complementarity index (Lawrence & Colman, 1993) is 0.65, only slightly lower than the values in the range 0.72–0.76 calculated for TBA complexes, in which the contact area is much smaller (563 Å^2).

Circular dichroism studies indicate that oligonucleotides embodying duplex–quadruplex domains display large conformational variability (Marson *et al.*, 2012; Dolinnaya *et al.*, 2012; Lin *et al.*, 2011; Zavyalova *et al.*, 2011; Zhou *et al.*, 2010) depending on the experimental conditions, such as the nature and concentration of salts, the temperature and the pH. In order to establish a correlation between solution and solid-state structure, we have complemented the crystallographic analysis with a CD study. Comparison of the CD spectra of the free HD22-27mer and its complex with thrombin (see §2 for details) indicates that the binding does not produce significant modifications of the aptamer spectral features, but enhances its thermal stability. Similar results were obtained in the case of TBA (Nagatoishi *et al.*, 2011; Russo Krauss *et al.*, 2012). This finding supports the hypothesis that the crystal structure is representative of the aptamer conformation in solution.

5. Conclusions

In recent years, several attempts to improve the biological activity of thrombin-binding aptamers *via* chemical and structural modifications have been performed (Pasternak *et al.*, 2011, and references therein; Musumeci & Montesarchio, 2012; Müller *et al.*, 2007, 2008; Rangnekar *et al.*, 2012). In this context, interesting results have been achieved with bimodular nucleotides adopting mixed duplex–quadruplex structures. Here, we have presented the structural characterization of the complex between thrombin and one of these aptamers, HD22-27mer. Our study provides new and interesting data on the structural and conformational behaviour of bimodular aptamers that will serve as a platform for the rational design of new molecules for use in anticoagulant therapies and diagnostic applications. In particular, the results emphasize the role of the residues controlling the spatial transition from the duplex to the quadruplex architecture and the recruitment of both motifs in thrombin binding. On more general grounds, the behaviour of HD22-27mer indicates that bimodular oligonucleotides should not be thought as fixed structured building blocks linked together. The two-residue upstream shift of the pseudo-G-quadruplex motif with respect to the TBA-like sequence, the bulging out of Thy24 in the duplex region and the intriguing connection between the two helical motifs, which controls their relative orientation, are examples of the variety of possible factors that play a role in determining the unexpected shape of the aptamer bound to the protein in this specific case. This intrinsic structural plasticity, which is confirmed by the various solution studies reported in the literature (Marson *et al.*, 2012; Zhou *et al.*, 2010; Lin *et al.*, 2011), makes these aptamers multifaceted ligands that are able to assume different folds depending on even small variations in environmental parameters.

We acknowledge the assistance of the staff at the crystallographic beamline of the Elettra synchrotron facility, Trieste, Italy. Giosuè Sorrentino and Maurizio Amendola (Institute of Biostructures and Bioimaging, Naples, Italy) are also gratefully acknowledged for technical assistance.

References

Balasubramanian, S. & Neidle, S. (2009). *Curr. Opin. Chem. Biol.* **13**, 345–353.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chinnapen, D. J. & Sen, D. (2002). *Biochemistry*, **41**, 5202–5212.
Chou, S.-H., Chin, K.-H. & Wang, A. H.-J. (2005). *Trends Biochem. Sci.* **30**, 231–234.
Dolinnaya, N. G., Yuminova, A. V., Spiridonova, V. A., Arutyunyan, A. M. & Kopylov, A. M. (2012). *J. Biomol. Struct. Dyn.* **30**, 524–531.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fernando, H., Sewitz, S., Darot, J., Tavaré, S., Huppert, J. L. & Balasubramanian, S. (2009). *Nucleic Acids Res.* **37**, 6716–6722.
Gatto, B., Palumbo, M. & Sissi, C. (2009). *Curr. Med. Chem.* **16**, 1248–1265.
Hasegawa, H., Taira, K., Sode, K. & Ikebukuro, K. (2008). *Sensors*, **8**, 1090–1098.
Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996). *Nature (London)*, **381**, 272.
Jayasena, S. D. (1999). *Clin. Chem.* **45**, 1628–1650.
Keefe, A. D., Pai, S. & Ellington, A. (2010). *Nature Rev. Drug Discov.* **9**, 537–550.
Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. (2007). *Nature Chem. Biol.* **3**, 218–221.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.
Li, T., Wang, E. & Dong, S. (2009). *PLoS One*, **4**, e5126.
Lin, P.-H., Chen, R.-H., Lee, C.-H., Chang, Y., Chen, C.-S. & Chen, W.-Y. (2011). *Colloids Surf. B*, **88**, 552–558.
Lipps, H. J. & Rhodes, D. (2009). *Trends Cell Biol.* **19**, 414–422.
Long, S. B., Long, M. B., White, R. R. & Sullenger, B. A. (2008). *RNA*, **14**, 2504–2512.
Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. & Feigon, J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 3745–3749.
Macaya, R. F., Waldron, J. A., Beutel, B. A., Gao, H., Joesten, M. E., Yang, M., Patel, R., Bertelsen, A. H. & Cook, A. F. (1995). *Biochemistry*, **34**, 4478–4492.
Marson, G., Palumbo, M. & Sissi, C. (2012). *Curr. Pharm. Des.* **18**, 2027–2035.
Mayer, G., Müller, J., Mack, T., Freitag, D. F., Hover, T., Pötzsch, B. & Heckel, A. (2009). *ChemBioChem*, **10**, 654–657.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Mori, T., Oguro, A., Ohtsu, T. & Nakamura, Y. (2004). *Nucleic Acids Res.* **32**, 6120–6128.
Müller, J., Freitag, D., Mayer, G. & Pötzsch, B. (2008). *J. Thromb. Haemost.* **6**, 2105–2112.
Müller, J., Wulffen, B., Pötzsch, B. & Mayer, G. (2007). *ChemBioChem*, **8**, 2223–2226.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Musumeci, D. & Montesarchio, D. (2012). *Pharmacol. Ther.* **136**, 202–215.
Nagatoishi, S., Isono, N., Tsumoto, K. & Sugimoto, N. (2011). *Biochimie*, **93**, 1231–1238.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pasternak, A., Hernandez, F. J., Rasmussen, L. M., Vester, B. & Wengel, J. (2011). *Nucleic Acids Res.* **39**, 1155–1164.
Patel, D. J., Phan, A. T. & Kuryavyi, V. (2007). *Nucleic Acids Res.* **35**, 7429–7455.
Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R. B. & Patel, D. J. (2011). *Nature Struct. Mol. Biol.* **18**, 796–804.
Pileur, F., Andreola, M. L., Dausse, E., Michel, J., Moreau, S., Yamada, H., Gaidamakov, S. A., Crouch, R. J., Toulmé, J. J. & Cazenave, C. (2003). *Nucleic Acids Res.* **31**, 5776–5788.
Rangnekar, A., Zhang, A. M., Li, S. S., Bompiani, K. M., Hansen, M. N., Gothelf, K. V., Sullenger, B. A. & LaBean, T. H. (2012). *Nanomedicine*, **8**, 673–681.
Russo Krauss, I., Merlino, A., Giancola, C., Randazzo, A., Mazzarella, L. & Sica, F. (2011). *Nucleic Acids Res.* **39**, 7858–7867.
Russo Krauss, I., Merlino, A., Randazzo, A., Mazzarella, L. & Sica, F. (2010). *Acta Cryst.* F**66**, 961–963.
Russo Krauss, I., Merlino, A., Randazzo, A., Novellino, E., Mazzarella, L. & Sica, F. (2012). *Nucleic Acids Res.* **40**, 8119–8128.
Russo Krauss, I., Merlino, A., Vergara, A. & Sica, F. (2013). *Int. J. Mol. Sci.* **14**, 11643–11691.
Silverman, A. P., Bu, W., Cohen, S. M. & Lippard, S. J. (2002). *J. Biol. Chem.* **277**, 49743–49749.
Spiridonova, V. A., Rog, E. V., Dugina, T. N., Strukova, S. M. & Kopylov, A. M. (2003). *Russ. J. Bioorg. Chem.* **29**, 450–453.
Tan, W., Wang, H., Chen, Y., Zhang, X., Zhu, H., Yang, C., Yang, R. & Liu, C. (2011). *Trends Biotechnol.* **29**, 634–640.
Tasset, D. M., Kubik, M. F. & Steiner, W. (1997). *J. Mol. Biol.* **272**, 688–698.
Wang, K. Y., Krawczyk, S. H., Bischofberger, N., Swaminathan, S. & Bolton, P. H. (1993). *Biochemistry*, **32**, 11285–11292.
Williamson, J. R. (1994). *Annu. Rev. Biophys. Biomol. Struct.* **23**, 703–730.
Zavyalova, E., Golovin, A., Reshetnikov, R., Mudrik, N., Panteleyev, D., Pavlova, G. & Kopylov, A. (2011). *Curr. Med. Chem.* **18**, 3343–3350.
Zhou, G., Huang, X. & Qu, Y. (2010). *Biochem. Eng. J.* **52**, 117–122.