Crystal structure of the DNA sequence d(CGTGAATTCACG)₂ with DAPI

Hristina I. Sbirikova-Dimitrova* and Boris Shivachev*

Institute of Mineralogy and Crystallography (IMC), Bulgarian Academy of Sciences, Acad. Georgi Bonchev Str., Bl. 107, 1113 Sofia, Bulgaria. *Correspondence e-mail: sbirikova@mail.bg, blshivachev@gmail.com

Received 30 May 2017

Accepted 2 August 2017

Edited by A. Nakagawa, Osaka University, Japan

Keywords: DNA; single crystal; DAPI; fluorescent intercalator displacement; ligand; palindrome.

PDB reference: d(CGTGAATTCACG) with DAPI, 5t4w

Supporting information: this article has supporting information at journals.iucr.org/f

The structure of 4',6-diamidine-2-phenylindole (DAPI) bound to the synthetic B-DNA oligonucleotide d(CGTGAATTCACG) has been solved in space group $P2_12_12_1$ by single-crystal X-ray diffraction at a resolution of 2.2 Å. The structure is nearly isomorphous to that of the previously reported crystal structure of the oligonucleotide d(CGTGAATTCACG) alone. The adjustments in crystal packing between the native DNA molecule and the DNA–DAPI complex are described. DAPI lies in the narrow minor groove near the centre of the B-DNA fragment, positioned over the A–T base pairs. It is bound to the DNA by hydrogen-bonding and van der Waals interactions. Comparison of the two structures (with and without ligand) shows that DAPI inserts into the minor groove, displacing the ordered spine waters. Indeed, as DAPI is hydrophobic it confers this behaviour on the DNA and thus restricts the presence of water molecules.

1. Introduction

The d(CGTGAATTCACG) DNA duplex is interesting because it features an EcoRI restriction site (Kapuscinski, 1995; Kim *et al.*, 1990). DAPI (4',6-diamidino-2-phenylindole) shows strong fluorescence when bound to DNA, which led to the rapid adoption of DAPI for fluorescent staining in fluorescence microscopy (Russell *et al.*, 1975). In addition, DAPI is a DNA-specific binding molecule which forms a fluorescent complex by attaching to the minor groove of A/T-rich sequences of DNA. When bound to double-stranded DNA (dsDNA), DAPI has an absorption maximum at a wavelength of 358 nm (ultraviolet) and its emission maximum is at 461 nm (blue). Therefore, in fluorescence microscopy DAPI is excited with ultraviolet light and is detected using a blue/cyan filter. Apart from analytical fluorescence light microscopy, DAPI is also popular for the labelling of cell cultures to detect DNA from contaminating mycoplasma or viruses (Biancardi *et al.*, 2013). We accomplished our aim of cocrystallizing DNA with DAPI, as described in previous work (Sbirikova & Shivachev, 2016). The mechanisms of DAPI–DNA complex formation, including minor-groove binding, hydrogen bonding and hydrophobic activity, are discussed.

2. Materials and methods

2.1. Macromolecule production

The dry oligonucleotide sequence 5'-CGTGAATTCACG-3' was purchased from Eurofins MWG Genomics.

2.2. Crystallization

The DNA sequence 5'-CGTGAATTCACG-3' (single-stranded) was dissolved in buffer to 3 mM and was annealed

Table 1
Crystallization.

Method	Hanging drop
Plate type	24-well
Temperature (K)	290
Final DNA concentration (mM)	1
Buffer composition of DNA solution	60 mM sodium cacodylate pH 7.0, 17 mM MgCl ₂ , 2 mM spermine
Composition of reservoir solution	50% (v/v) MPD
Volume and ratio of drop	3 μ l (1:1:1 DNA:DAPI:reservoir)
Volume of reservoir (μ l)	550

for 1 min at 75°C before use in order to obtain double-stranded DNA (dsDNA). The buffer solution consisted of 60 mM sodium cacodylate pH 7.0, 17 mM MgCl₂, 2 mM spermine. DAPI was dissolved in the same buffer at 3 mM. Crystals were grown from hanging drops (3 μ l) equilibrated against 50% (v/v) 2-methyl-2,4-pentanediol (MPD) at room temperature. Large crystals (0.4 $\times$ 0.3 $\times$ 0.25 mm) suitable for single-crystal X-ray studies formed within a month (Fig. 1). Crystallization information is summarized in Table 1.

2.3. Data collection and processing

Crystals were mounted on loops and flash-cooled at 130 K directly under a Cobra nitrogen cryostream (Oxford Cryosystems). All data were collected at low temperature (130 K) on an Oxford Diffraction SuperNova diffractometer using Cu K α radiation (λ = 1.54056 Å) from a microfocus source. Determination of the unit-cell parameters, data integration, scaling and absorption corrections were carried out using *CrysAlis^{Pro}* (Rigaku).

2.4. Structure solution and refinement

The phases were obtained by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using PDB entry 1d29 (Larsen *et al.*, 1991) as a starting model. Refinement of the structure involved several cycles of refinement using *REFMAC5* (Murshudov *et al.*, 2011) and *Coot* (Emsley *et al.*, 2010). The water molecules and ligand (DAPI) were positioned from the $F_o - F_c$ difference map using the *Coot* interface (Emsley *et al.*,

Figure 1
Observed crystals of 5'-CGTGAATTCACG-3' complexed with DAPI.**Table 2**
Data collection and processing.

Diffraction source	SuperNova Dual, microfocus Cu (Nova)
Wavelength (Å)	1.54
Temperature (K)	130
Detector	Atlas CCD
Crystal-to-detector distance (mm)	56
Rotation range per image (°)	0.5
Exposure time per image (s)	80
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	24.50, 41.09, 65.18
α , β , γ (°)	90, 90, 90
Mosaicity (°)	1.3, 1.01, 2.28
Resolution range (Å)	34.76–2.30
<i>R</i> _{merge}	0.08 (0.21)
Total No. of reflections	43577
No. of unique reflections	3229
Completeness (%)	90.1 (77.9)
Multiplicity	8.3 (5.4)
$\langle I/\sigma(I) \rangle$	40.1 (7.9)
Overall <i>B</i> factor from Wilson plot (Å ²)	38.04

Table 3
Structure refinement.

Resolution range (Å)	34.76–2.30 (2.36–2.30)
Completeness (%)	90.1 (77.9)
σ Cutoff	None
No. of reflections, working set	2912 (161)
No. of reflections, test set	291 (15)
Final <i>R</i> _{cryst}	0.216 (0.337)
Final <i>R</i> _{free}	0.258 (0.454)
Cruickshank DPI	0.378
No. of non-H atoms	
DNA	486
Ligand	21
Solvent	3
Total	510
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.559
Average <i>B</i> factors (Å ²)	
Nucleic acid	39.5
Ligand	41.1
Water	31.6

2010). Visual analyses of the model and the electron-density maps were carried out using *Coot* (Emsley *et al.*, 2010). A summary of the most important crystal data and refinement indicators is provided in Tables 2 and 3. *X3DNA* (Zheng *et al.*, 2009) was used to carry out structural analysis and geometrical calculations of DNA parameters. *USFC Chimera* (Pettersen *et al.*, 2004) was used to prepare the figures. The coordinates and structure factors have been deposited in the PDB as entry 5t4w.

3. Results and discussion

Single-crystal data collection was attempted from several different crystals. It should be noted that the data set collected to a resolution of 2.3 Å was from a crystal that was harvested from the drop 5 d after it was spotted. Crystals with similar or even larger dimensions that were allowed to 'stabilize' for more than a week in the crystallization drop usually diffracted to resolutions of up to 2.5 Å. Attempts to collect data at room

temperature (19°C) were also performed on a few crystals; however, the observed quality of the diffraction was not comparable with that for experiments conducted at 130 K. The presence of DAPI in the crystallization conditions may have

Figure 2
(a) View of the asymmetric unit of PDB entry 5t4w including the DAPI molecule (red) and (b) view of the asymmetric unit of PDB entry 5ju4 (Mg^{2+} is in green).

played a role in destabilizing the crystal structure. It should be noted that there was no difference in the crystallization conditions used to obtain the structures of d(CGTTGAATTCACG)₂ without (PDB entry 5ju4; Sbirikova & Shivachev, 2016) and with (PDB entry 5t4w) bound DAPI. In both crystallization trials crystals formed after a month and if they were allowed to equilibrate the diffraction quality was worse. We cannot provide a plausible explanation for the presence and absence of DAPI in crystals obtained under the same crystallization conditions.

The asymmetric unit of PDB entry 5t4w consists of two chemically equivalent self-complementary strands (each of 12 bases in length) forming antiparallel right-handed DNA (Fig. 2). The B-type DNA duplex is formed by classical Watson–Crick (W–C) hydrogen-bonding base-pairing interactions between the two strands: bases C1 to G12 from the first strand interact with bases G13 to C24 from the opposite (second) strand (the numbering corresponds to the sequence in the crystal structure). The minor groove of the present double-stranded oligonucleotide features a central TpA step (AATT) surrounded by C/G-rich regions. Its overall secondary structure (Fig. 2) is comparable to previously reported low-temperature and room-temperature structures with the same sequence: PDB entries 5ju4, 1d28 and 1d29 (Sbirikova & Shivachev, 2016; Narayana *et al.*, 1991; Larsen *et al.*, 1991). The base-pair morphology values for shear, stretch, stagger, buckle, opening and propeller twist obtained using X3DNA (Zheng *et al.*, 2009) are shown in Supplementary Table S1. It is obvious from the values that the interaction of DAPI changes the conformation of the DNA when they form a complex. The variation in the shear and stretch decreases,

Figure 3
Representation of canonical Watson–Crick interactions within the duplex and noncanonical G...G interactions between the adjacent DNA duplexes. (a) G...G interaction I between bases Dg2 and Dg24. (b) G...G interaction II between bases Dg12 and Dg14.

while that in the stagger, buckle, opening and propeller twist values increases. The most pronounced differences are in the core TpA region; for example, the values of buckle are 6.05 and 8.06 for A–T pairs 5–20 in 5t4w and 5ju4, respectively, while the values of propeller twist are -22.16 and -15.28 and those of opening are 9.28 and 1.8, respectively, for T–A pairs 7–18. The interaction of DAPI destabilizes the minor-groove conformation, while in contrast the major groove and the CpG regions at the end of the sequence show significant stabilization. Consequently, owing to this compensatory behaviour, although the DNA sequence is altered the hydrogen bonds are kept unchanged and the intrastrand interactions in PDB entries 5ju4 and 5t4w produce a motif that is in agreement with the Dickerson–Drew dodecamer and has classical right-handed B-DNA duplex structural features (Wing *et al.*, 1984; Wei *et al.*, 2013).

Both the 5t4w and 5ju4 (Sbirkova & Shivachev, 2016) crystal structures confirm the nonclassical interstrand hydrogen-bonding interactions involving G bases, as previously reported by Hossain *et al.* (2002). The base pairs C1–G24, G2–C23 and G12–C13, C11–G14 located at the ends of the duplexes form G···G bonds to the adjacent DNA duplexes. The observed G···G hydrogen bonding does not correspond to that of Hoogsteen (1959). Representative electron densities and hydrogen-bonding interactions are shown in Fig. 3. Based on the donor–acceptor distances ($D\cdots A$), the observed G···G hydrogen bonds are slightly weaker than classical W–C hydrogen bonds (Hossain *et al.*,

2002; Desiraju & Steiner, 1999). (The $D\cdots A$ distance for $G\cdots G$ is around $3.0\text{ \AA}$, while that for $C\cdots G$ is around $2.85\text{ \AA}$.) The $G\cdots G$ interactions are located at the ends of the DNA strands, while the $DNA\cdots DAPI\cdots DNA$ interactions (Fig. 4) involve the AATT region. DAPI requires more space and thus occupies a ‘water-rich’ region in the minor groove, conferring it with a hydrophobic behaviour, and compensating the negative charge of the DNA. On the other hand, the upper and lower surfaces of the purine and pyrimidine rings are also hydrophobic and the $G\cdots G$ interactions exploit the interaction of the edges of the bases (which are hydrophilic), thus eliminating the need for water molecules. The Mg^{2+} ions compensate for the DNA negative charge in the 5ju4 structure (Sbirkova & Shivachev, 2016). Charge compensation is achieved through $Mg^{2+}\cdots Da6A$, $Dt7A$ (Chiu & Dickerson, 2000; Sbirkova & Shivachev, 2016). The $2F_o - F_c$ difference map of PDB entry 5t4w suggests the absence of heavier atoms (than water), for example Mg^{2+} ions. Indeed, in the present structure the charge is compensated by the $DAPI^{2+}$ molecule and thus no Mg^{2+} ion is required, and it cannot be detected in the structure. The DAPI molecule is additionally hydrogen-bonded to $Dt19B$, $Da6A$, $Dt7A$, $Dt8A$ and $Dc9A$, strengthening the DNA–DAPI interactions (Fig. 4). The asymmetric unit of PDB entry 5t4w contains only three water molecules (while 74 are located in PDB entry 5ju4). The observed reduction in the ordered waters in the hydration shell may be owing to the hydrophobic property conferred on the DNA–DAPI complex by DAPI. A similar effect of DAPI is observed when it is present in complex with the Dickerson sequence in PDB entry 1d30 (Larsen *et al.*, 1989). Only four DNA–DAPI crystal structures are available in the PDB and NDB: PDB entries 1d30 (Larsen *et al.*, 1989), 432d (Vlieghe *et al.*, 1999), 3ajl (Tsunoda *et al.*, 2010) and 3gjh (Juan *et al.*, 2010). $-C(NH_2)_2^+$ groups with identical hydrogen-bonding capabilities are present at both ends of the DAPI molecule. Thus, the orientation of the DAPI molecule in the DNA minor groove is governed by the indole NH proton (Granzhan *et al.*, 2014). Comparison of the insertion of DAPI into the DNA shows that the DAPI molecule tends to favour the formation of a hydrogen bond between the indole NH and a T base (Larsen *et al.*, 1989; Vlieghe *et al.*, 1999; Tsunoda *et al.*, 2010; Juan *et al.*, 2010). The orientation of DAPI is similar to that observed in PDB entry 1d30 (Larsen *et al.*, 1989), while the positioning in the minor groove is shifted in the direction of the indole moiety (or towards the water molecule near to DAPI in PDB entry 1d30). Even though the position of DAPI varies slightly, most of the hydrogen-bonding interactions are retained.

Funding information

This study was partially supported by the Bulgarian Fund for Research Investigations (FNI), grant Nos. T02/14 and DRNF 02/1.

References

- Biancardi, A., Biver, T., Secco, F. & Mennucci, B. (2013). *Phys. Chem. Chem. Phys.* **15**, 4596–4603.

Figure 4

Observed interactions of DAPI with the DNA bases; hydrogen bonds are shown as dashed lines and distances are shown in Å.

- Chiu, T. K. & Dickerson, R. E. (2000). *J. Mol. Biol.* **301**, 915–945.
- Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*. Oxford University Press.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Granzhan, A., Kotera, N. & Teulade-Fichou, M.-P. (2014). *Chem. Soc. Rev.* **43**, 3630–3665.
- Hoogsteen, K. (1959). *Acta Cryst.* **12**, 822–823.
- Hossain, M. T., Kondo, J., Ueno, Y., Matsuda, A. & Takénaka, A. (2002). *Biophys. Chem.* **95**, 69–77.
- Juan, E. C. M., Shimizu, S., Ma, X., Kurose, T., Haraguchi, T., Zhang, F., Tsunoda, M., Ohkubo, A., Sekine, M., Shibata, T., Millington, C. L., Williams, D. M. & Takénaka, A. (2010). *Nucleic Acids Res.* **38**, 6737–6745.
- Kapuscinski, J. (1995). *Biotech. Histochem.* **70**, 220–233.
- Kim, Y., Grable, J. C., Love, R., Greene, P. J. & Rosenberg, J. M. (1990). *Science*, **249**, 1307–1309.
- Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K. & Dickerson, R. E. (1989). *J. Biomol. Struct. Dyn.* **7**, 477–491.
- Larsen, T. A., Kopka, M. L. & Dickerson, R. E. (1991). *Biochemistry*, **30**, 4443–4449.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- 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.* **D67**, 355–367.
- Narayana, N., Ginell, S. L., Russu, I. M. & Berman, H. M. (1991). *Biochemistry*, **30**, 4449–4455.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Russell, W., Newman, C. & Williamson, D. (1975). *Nature (London)*, **253**, 461–462.
- Sbirkova, H. I. & Shivachev, B. L. (2016). *Bul. Chem. Commun.* **48**, 589–592.
- Tsunoda, M., Sakaue, T., Naito, S., Sunami, T., Abe, N., Ueno, Y., Matsuda, A. & Takénaka, A. (2010). *J. Nucleic Acids*, **2010**, 107289.
- Vlieghe, D., Sponer, J. & Van Meervelt, L. (1999). *Biochemistry*, **38**, 16443–16451.
- Wei, D., Wilson, W. D. & Neidle, S. (2013). *J. Am. Chem. Soc.* **135**, 1369–1377.
- Wing, R. M., Pjura, P., Drew, H. R. & Dickerson, R. E. (1984). *EMBO J.* **3**, 1201–1206.
- Zheng, G., Lu, X.-J. & Olson, W. K. (2009). *Nucleic Acids Res.* **37**, W240–W246.