Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of a self-complementary DNA heptacosamer with a 20-base-pair duplex flanked by seven-nucleotide overhangs at the 3′-terminus

Hyun Koo Yeo and Jae Young Lee*

Department of Life Science, Dongguk University-Seoul, 26 Pil-dong 3-ga, Jung-gu, Seoul 100-715, Republic of Korea

Correspondence e-mail: jylee001@dongguk.edu

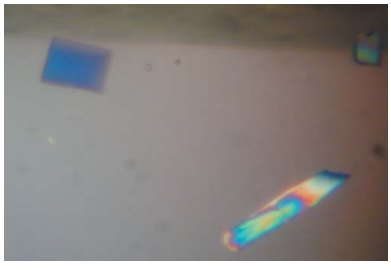
Received 18 January 2010
Accepted 22 March 2010

The self-complementary DNA heptacosamer (a 27-mer oligonucleotide) with sequence d(CGAGCACTGCGCAGTGCTCGTTGTTAT) forms a 20-base-pair duplex flanked by seven-nucleotide overhangs at the 3′-terminus. Crystals of the oligonucleotide were obtained by sitting-drop vapour diffusion and diffracted to 2.8 Å resolution. The oligonucleotide was crystallized at 277 K using polyethylene glycol as a precipitant in the presence of magnesium chloride. The crystals belonged to the triclinic space group, with unit-cell parameters $a = 48.74$, $b = 64.23$, $c = 79.34$ Å, $\alpha = 91.37$, $\beta = 93.21$, $\gamma = 92.35^\circ$.

1. Introduction

Our knowledge of the architectural principles of nucleic acid structure has advanced significantly over the past decade with the determination to high resolution of larger structures of great biological significance. The conformation of DNA structures is fundamental to diverse DNA transactions such as replication, repair, nucleosome assembly and regulation of gene expression (Rhodes *et al.*, 1996; Travers, 2004). Many proteins bind to DNA in a sequence- or structure-specific manner appropriate to their biological role. Comparative structural analyses of DNA oligonucleotides and protein–DNA complexes have provided insight into the physical principles of DNA conformation, reorganization and recognition. A self-complementary oligonucleotide flanked by several extra single-strand nucleotides is a suitable DNA substrate for proteins which have an affinity for both duplex and single-strand regions, such as DNA helicases. A self-complementary nucleotide sequence can potentially adopt a unimolecular hairpin as well as a bimolecular duplex. The hairpin structure is favoured at a lower cationic concentration, while the bimolecular duplex is formed at higher cationic strength (Nakano *et al.*, 2007). The difference in the preference for secondary-structure formation depending on the cationic strength can lead to the structural transition of self-complementary nucleotides from a hairpin to a bimolecular duplex (Nakano *et al.*, 2007; Hald *et al.*, 1995; Wemmer *et al.*, 1985; Marky *et al.*, 1983).

The crystal structure of *Escherichia coli* UvrD helicase has been determined in complex with multiple self-complementary oligonucleotides that form 18–28-base-pair duplexes flanked by seven-nucleotide overhangs at the 3′-terminus (Lee & Yang, 2006). Because the UvrD–DNA complex structures were determined in the absence of the structures of the free DNA oligonucleotides, it is worthwhile determining the free DNA oligonucleotide structures, which will provide insight into DNA recognition by UvrD and induced fit by the DNA-binding motifs of UvrD. In addition, the overhanging nucleotides may project into the groove of a neighbouring duplex in a semi-continuous stack of duplexes, forming a base triplet (Vlieghe *et al.*, 1996). On searching the NDB (Berman *et al.*, 1992), the longest duplex structures found were a 21-base-pair DNA determined by NMR (Masliah *et al.*, 2008) and a 17-mer oligonucleotide forming a 16-base-pair duplex with a 5′-terminal overhang determined by X-ray crystallography (Huang *et al.*, 2005). The heptacosameric self-complementary oligonucleotide forming a 20-base-pair duplex flanked by seven-nucleotide single-strand overhangs at the



3′-terminus is the longest DNA to be crystallized in the absence of protein. As part of our crystallographic studies on *E. coli* UvrD helicase in complex with oligonucleotides containing single-stranded and double-stranded regions of DNA (Lee & Yang, 2006), here we report the crystallization and preliminary X-ray data of the heptacosamer with sequence d(CGAGCACTGCGCAGTGCTCGTTGTTAT).

2. Materials and methods

2.1. Crystallization

The self-complementary oligonucleotide d(CGAGCACTGCGACGTGCTCGTTGTTAT) was obtained from the Keck Oligonucleotide Synthesis Facility (Yale University). The oligonucleotide was further purified by reverse-phase HPLC using a Poros Oligo R3 column (Applied Biosystems), detritylation and a final desalting step using an NAP-25 column (GE Healthcare). The oligonucleotide was lyophilized prior to solubilization in TE buffer to obtain the final stock concentration. The concentration of the oligonucleotide was measured at 260 nm using an absorption coefficient of 247 700 l cm^{-1} mol^{-1}. The oligonucleotide was annealed by heating to 363 K for 10 min and gradual cooling to room temperature.

The purified oligonucleotide was resuspended in ice-cold buffer (20 m*M* Tris–HCl pH 8.0, 0.15 *M* KCl, 5 m*M* $MgCl_2$, 1 m*M* DTT and 5% glycerol). The final concentration of the oligonucleotide used in crystallization trials was about 0.05 m*M*. Crystallization was performed by the sitting-drop vapour-diffusion method using 96-well CrystalQuick plates (Greiner Bio-One) at 277 K. Each sitting drop, which was prepared by mixing 1 µl each of oligonucleotide solution and reservoir solution, was placed over 0.1 ml reservoir solution. Results from the initial searches for crystallization conditions performed using Crystal Screen I (Jancarik & Kim, 1991), Crystal Screen II, Natrix, PEG/Ion and Index Screens (Hampton Research) and Wizard I and II (DeCODE Genetics) were optimized.

2.2. X-ray data collection

Crystals were transferred into a cryoprotectant solution containing 20%(*v*/*v*) glycerol in reservoir solution and then flash-cooled in liquid nitrogen. X-ray diffraction data were collected from crystals at 100 K in the nitrogen-gas stream from an Oxford Cryojet system with a MAR 300 CCD detector using synchrotron radiation on SER-CAT beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory. Data were collected using a 1° oscillation per image with a crystal-to-detector distance of 300 mm and a total of 360 frames were recorded. A data set was collected to 2.8 Å resolution from a single crystal. Data were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL* program suite (Otwinowski & Minor, 1997).

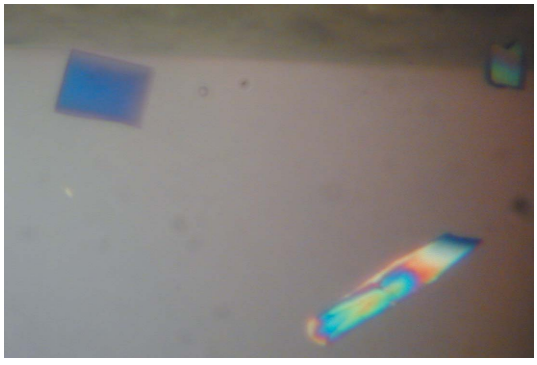


Figure 1
Triclinic crystals of a self-complementary DNA heptacosamer. Their approximate dimensions are 0.2 × 0.15 × 0.05 mm.

Table 1
Data-collection statistics for the DNA heptacosamer crystal.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.000
Space group	*P*1
Unit-cell parameters (Å, °)	*a* = 48.74, *b* = 64.23, *c* = 79.34, α = 91.37, β = 93.21, γ = 92.35
Temperature (K)	100
Resolution range (Å)	50.0–2.80 (2.90–2.80)
No. of observations	64279
Unique reflections	21117
Data completeness (%)	90.3 (57.4)
Redundancy	3.1 (2.5)
Average $I/\sigma(I)$	29.2 (2.4)
R_{merge}† (%)	6.0 (26.7)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the intensity of reflection *hkl*, $\sum_{hkl}$ is the sum over all reflections and $\sum_i$ is the sum over *i* measurements of reflection *hkl*.

3. Results

The self-complementary DNA heptacosamer d(CGAGCACTGCGCAGTGCTCGTTGTTAT) was synthesized and purified for crystallization experiments using reverse-phase HPLC. Crystals of the oligonucleotide were obtained by sitting-drop vapour diffusion at 277 K. The optimized reservoir solution for crystallization was 20%(*v*/*v*) polyethylene glycol (PEG) 3350, 0.1 *M* bis-tris pH 5.5 and 0.25 *M* magnesium chloride. Crystals grew reproducibly to maximum dimensions of approximately 0.2 × 0.15 × 0.05 mm within two weeks (Fig. 1). Crystals were exchanged into a cryoprotectant consisting of 25%(*v*/*v*) PEG 3350, 0.25 *M* $MgCl_2$, 0.1 *M* bis-tris pH 5.5 and 20%(*v*/*v*) glycerol. A crystal was flash-cooled in liquid nitrogen and diffracted to 2.8 Å resolution with a slightly anisotropic diffraction pattern. The crystal belonged to the triclinic space group *P*1, with unit-cell parameters *a* = 48.74, *b* = 64.23, *c* = 79.34 Å, α = 91.37, β = 93.21, γ = 92.35°. A total of 21 117 unique reflections were measured with 90.3% completeness at 2.8 Å resolution. Solvent-content analysis using the program *MATTHEWS_COEF* (Kantardjieff & Rupp, 2003) indicated that the crystal could contain between five and eight molecules of the bimolecular duplex in the asymmetric unit. The Matthews coefficient ranged from 1.85 to 2.96 Å^3 Da^{-1} and the solvent content from 55.1% to 72.0%, based on a partial specific volume of 0.5 cm^3 g^{-1} (Matthews, 1968; Cohen & Eisenberg, 1968). Table 1 summarizes the statistics of data collection. Structure determination is currently ongoing using molecular replacement with the program *Phaser* (Storoni *et al.*, 2004) with a model based on either ideal B-DNA or the 20-base-pair duplex DNA from the UvrD–DNA complexes (Lee & Yang, 2006). Six molecules of duplex DNA were found in the asymmetric unit and were refined as a rigid body with a correlation coefficient of 0.771 and an *R* factor of 43.4% with *REFMAC*5 (Murshudov *et al.*, 1999) using data in the resolution range 30.0–2.8 Å. Model improvement and refinement of the structure are in progress.

We thank the staff members at Advanced Photon Source beamline 22-ID (SER-CAT) for their help with data collection. This research was supported by the Basic Science Research Program through the

National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0073810).

References

Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.
Cohen, G. & Eisenberg, H. (1968). *Biopolymers*, **6**, 1077–1100.
Hald, M., Pedersen, J. B., Stein, P. C., Kirpekar, F. & Jacobsen, J. P. (1995). *Nucleic Acids Res.* **23**, 4576–4582.
Huang, D. B., Phelps, C. B., Fusco, A. J. & Ghosh, G. (2005). *J. Mol. Biol.* **346**, 147–160.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Lee, J. Y. & Yang, W. (2006). *Cell*, **127**, 1349–1360.
Marky, L. A., Blumenfeld, K. S., Kozlowski, S. & Breslauer, K. J. (1983). *Biopolymers*, **22**, 1247–1257.
Masliah, G., Rene, B., Zargarian, L., Fermandjian, S. & Mauffret, O. (2008). *J. Mol. Biol.* **381**, 692–706.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Nakano, S., Kirihata, T., Fujii, S., Sakai, H., Kuwahara, M., Sawai, H. & Sugimoto, N. (2007). *Nucleic Acids Res.* **35**, 486–494.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–325.
Rhodes, D., Schwabe, J. W., Chapman, L. & Fairall, L. (1996). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **351**, 501–509.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Travers, A. A. (2004). *Philos. Transact. A Math. Phys. Eng. Sci.* **362**, 1423–1438.
Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1996). *Acta Cryst.* D**52**, 766–775.
Wemmer, D. E., Chou, S. H., Hare, D. R. & Reid, B. R. (1985). *Nucleic Acids Res.* **13**, 3755–3772.