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

Crystallization and preliminary X-ray study of the D-altritol oligonucleotide GTGTACAC

Margriet Ovaere,[a] Arthur Van Aerschot,[b] Mikhail Abramov,[b] Piet Herdewijn[b] and Luc Van Meervelt[a]*

[a]Department of Chemistry, Biomolecular Architecture and BioMacS, Katholieke Universiteit Leuven, Belgium, and [b]Laboratory of Medicinal Chemistry, Rega Institute for Medical Research and BioMacS, Katholieke Universiteit Leuven, Belgium

Correspondence e-mail: luc.vanmeervelt@chem.kuleuven.be

Received 21 January 2010
Accepted 24 February 2010

In altritol nucleic acids (ANAs), the natural five-membered ribose ring of RNA is replaced by the six-membered D-altritol ring. ANAs are good candidates to act as siRNAs in the RNA-interference pathway. Crystals of the fully modified altritol self-complementary octamer GTGTACAC were grown by the hanging-drop vapour-diffusion technique at 289 K. Diffraction data were recorded on SLS beamline X06DA and processed to 3.0 Å resolution. The crystals belonged to the hexagonal space group $P6_122$ or $P6_522$, with unit-cell parameters $a = 25.05$, $c = 117.58$ Å.

1. Introduction

The method of choice nowadays for control of gene expression involves the use of the natural RNAi mechanism, a process in which post-transcriptional gene silencing is obtained (Meister & Tuschl, 2004). Unlike the antisense strategy, in which a single-stranded oligonucleotide with either DNA-like or RNA-like conformational preferences can be used to hybridize with the target mRNA, dsRNA molecules of defined length (or dsRNA mimics, both called small interfering RNAs or siRNAs) are mandatory for loading the RNA-induced silencing complex (RISC). The latter is an enzymatic ensemble that is responsible for breakdown of the target mRNA, following unwinding of the dsRNA and selection and positioning of the correctly remaining antisense strand, allowing hybridization with its target. Hence, exogeneously applied siRNA duplexes can be used for gene silencing, but their enzymatic stability should be improved by chemical modifications in order to increase their lifetime and bioavailability (Dorsett & Tuschl, 2004).

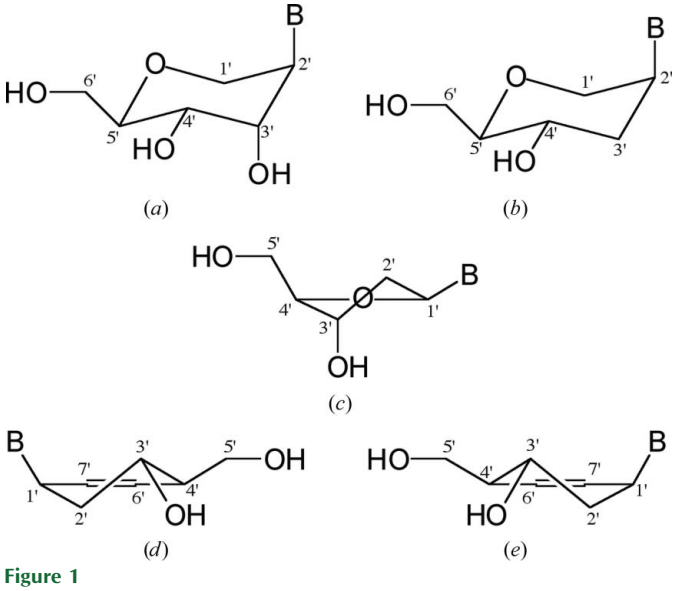

Figure 1
Comparison of the different monomeric structures of (*a*) ANA and (*b*) HNA, both with their base moiety typically oriented axial, (*c*) DNA in its preferred C2′-*endo* conformation and (*d*) left-handed and (*e*) right-handed CeNA residues, both in their preferred $^3H_2$ half-chair conformation mimicking the C3′-*endo* conformation of ribose nucleosides.

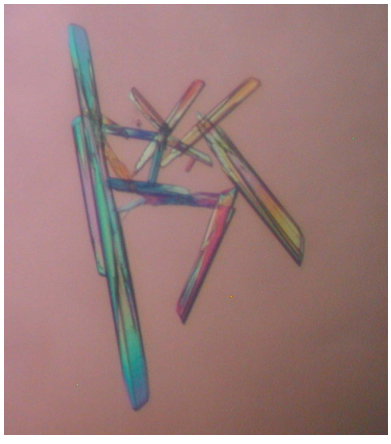

Table 1
Overview of the crystal structures with sequence GTGTACAC available in the Nucleic Acid Database (NDB; Berman *et al.*, 1992) with their corresponding space groups.

NDB code	Nucleic acid type	Space group
ADH034†, ADH038‡, ADH039‡	DNA	$P6_122$
ADH014§	DNA	$P4_32_12$
HD0001¶	HNA	$P3_212$
HD0002¶	HNA	$P6_522$
UD0068††	CeNA	$R32$

† Jain *et al.* (1991). ‡ Thota *et al.* (1993). § Jain *et al.* (1989). ¶ Declercq *et al.* (2002). †† Robeyns *et al.* (2008).

Many different chemical modifications have therefore been studied and much attention has been devoted to locked nucleic acids (LNAs; Elmén *et al.*, 2005; Puri *et al.*, 2008). However, the incorporation of altritol nucleic acid (ANA) monomers (Fig. 1*a*) at specific positions has not only been able to augment the intrinsic activity of siRNA molecules, but more important also dramatically increased the persistence of the expected gene-silencing effects (Fisher *et al.*, 2007, 2009). A large-scale screen with chemically modified siRNAs confirmed the high activity, high stability and low toxicity of the ANAs (Bramsen *et al.*, 2009). Likewise, some other nucleoside analogues with a six-membered ring substituting for the ribofuranose sugar [as in anhydrohexitol (HNA; Fig. 1*b*) and cyclohexenyl nucleic acids (CeNAs; Figs. 1*d* and 1*e*)] have been reported to be endowed with strong siRNA properties (Fisher *et al.*, 2009; Nauwelaerts *et al.*, 2007).

While the RNA-like conformation of ANA modifications has been documented previously (Allart, Khan *et al.*, 1999), the structure of a fully modified duplex has not yet been determined. However, crystal structures of GTGTACAC sequences of HNA (Declercq *et al.*, 2002), CeNA (left-handed structure; Robeyns *et al.*, 2008) and DNA (Jain *et al.*, 1989, 1991; Thota *et al.*, 1993) are already available (Table 1). Solution structures of a CeNA–RNA (Nauwelaerts *et al.*, 2007) and an ANA–RNA duplex (Froeyen *et al.*, 2000) are also available. A complete comparison between the structures of the different modifications therefore becomes possible.

2. Methods

2.1. Synthesis and purification

The altritol nucleic acid octamer GTGTACAC was assembled on a solid support using traditional phosphoramidite chemistry as reported previously (Allart, Khan *et al.*, 1999). The phosphoramidite building blocks made use of the benzoyl protecting group for the hydroxyl moiety at the 3′-position of the altritol modifications (Allart, Busson *et al.*, 1999). To isolate the required material, several 1 µmol scale runs were deprotected and purified by ion-exchange chromatography on a Mono Q column using a denaturing $NaClO_4$ gradient in the presence of 15% CH_3CN pH 7.4, followed by desalting on a Biogel P2 column. In view of the lower yielding true universal supports and to avoid special synthesis of altritol-functionalized solid supports, a 'universal' 3′-end phosphate-generating support was used (Guzaev *et al.*, 1995). Homemade DMTr-sulfonyl-diethanol-LCAA-CPG (80 µmol g^{-1}) was thus prepared, in this case generating a 4′-phosphate terminus on the fully modified altritol analogue.

2.2. Crystallization

Screening for crystallization conditions was carried out by the hanging-drop vapour-diffusion method at 289 K using the 24-matrix screen for nucleic acid fragments (Berger *et al.*, 1996). 2 µl drops were prepared by mixing 1 µl screen with 1 µl 1.5 m*M* oligonucleotide and were equilibrated against 500 µl 35% MPD solution. After 2 d, crystals were formed in the condition 40 m*M* cacodylate pH 7, 12 m*M* spermine tetrachloride, 40 m*M* LiCl, 20 m*M* $MgCl_2$, 80 m*M* $SrCl_2$ with 10%(*v*/*v*) MPD as precipitant. The crystals exhibited a thin plate-like shape with a length of up to 0.9 mm (Fig. 2). The crystals were flash-cooled in liquid nitrogen at 100 K for storage and shipping to the synchrotron facility. No additional cryoprotectant was applied.

Table 2
Data-collection statistics for the ANA sequence GTGTACAC.

Values in parentheses are for the outermost shell.

X-ray source	X06DA (PXIII), SLS
Wavelength (Å)	1.0000
Resolution range (Å)	39.19–3.00 (3.16–3.00)
Space group	$P6_122$ or $P6_522$
Unit-cell parameters (Å)	$a = b = 25.05$, $c = 117.58$
No. of reflections used	9196
No. of unique reflections	579 (88)
R_{merge} (%)	3.7 (35.1)
Mean $I/\sigma(I)$	39.1 (8.4)
Completeness (%)	99.1 (99.7)
Multiplicity	15.9 (16.3)

2.3. Data collection and processing

X-ray diffraction data were recorded at 100 K from a crystal with dimensions 0.22 × 0.06 × 0.01 mm on the Swiss Light Source (SLS, Paul Scherrer Institut, Villigen, Switzerland) X06DA (PXIII) beamline with a sub-100 × 100 µm focused beam at the sample position. Using a MAR225 mosaic CCD detector, a total of 180 frames with a φ increment of 1.0° were collected to 2.7 Å resolution at a wavelength of 1.0 Å. The diffraction data were processed using *MOSFLM* (Leslie, 1992) and scaled using the program *SCALA* (Evans, 2006) to a 3.0 Å cutoff in order to obtain a reasonable R_{merge} in the outer resolution shell. The latter was used as part of the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). At a slightly higher resolution of 2.9 Å the R_{merge} for the outer shell increases to 60% and the $I/\sigma(I)$ drops to 1.3.

Data-collection statistics are listed in Table 2.

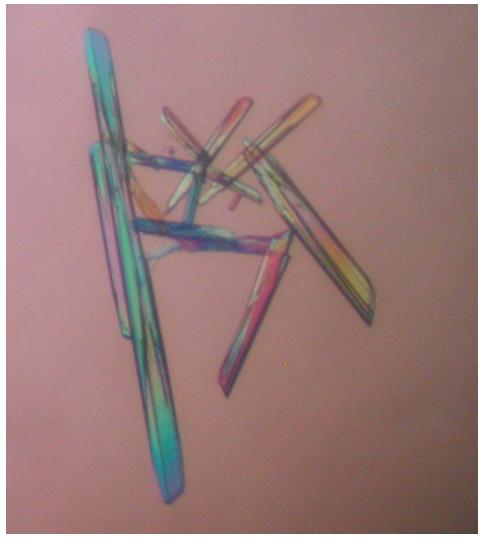

Figure 2
Typical crystals of the ANA self-complementary octamer GTGTACAC.

3. Results and discussion

Crystals of an altritol-modified oligonucleotide were grown for the first time. The crystals of the ANA sequence GTGTACAC have a hexagonal unit cell with parameters $a = b = 25.05$, $c = 117.58$ Å. The diffraction-pattern symmetry belongs to the 622 crystal class. Systematic absence analysis indicated the presence of a 6_1 or 6_5 screw axis, resulting in space group $P6_122$ or $P6_522$, respectively. Assignment of the final space group will be made during structure solution.

Based on calculation of the Matthews coefficient, the asymmetric unit cannot contain a full duplex. Estimation of the Matthews coefficient for half a duplex amounts to 1.97 Å^3 Da^{-1}, giving a solvent fraction of 57.8%. The full helix is then generated by a crystallographic twofold axis.

The volume per base pair (bp), assuming 4 bp in the asymmetric unit, is about 1331 Å^3. While this value lies within the observed volume range of both A-type DNA (1299–1836 Å^3) and B-type DNA (1175–1462 Å^3) (Heinemann, 1991), the ANA sequence GTGTACAC is likely to crystallize in the A-type family, since both DNA, HNA and CeNA octamers of the same sequence are known to form A-type helices (Table 1). Nevertheless, the potential formation of a B-type helix cannot be excluded.

With the availability of crystallographic models, structure determination was performed by molecular replacement with the program *Phaser* (McCoy *et al.*, 2007) using all of the models listed in Table 1. Since the asymmetric unit contains only half of the biological unit, the input models were truncated to half a duplex rather than a single strand to avoid the influence of a different helical curvature between the structures. An adequate solution was found in space group $P6_522$ with a four-base-pair model of the HNA structure HD0001. In this solution the other four base pairs are generated by a twofold crystallographic axis. However, refinement using the *REFMAC* program (Murshudov *et al.*, 1997) did not converge during successive cycles. Plausible causes are flipped-out or disordered bases, the presence of spermine in the crystal structure, which could influence the packing (Jain *et al.*, 1989), or other unresolved features. Structure solution by molecular replacement with A-type and B-type models is still ongoing. In a further attempt to solve the structure, SAD phasing on P atoms as well as on the possibly abundant Mg and Sr atoms will be explored.

We are indebted to Guy Schepers for the synthesis and purification of the altritol octamer and thank the staff of Swiss Light Source beamline X06DA in Villigen for help with the synchrotron experiments.

References

Allart, B., Busson, R., Rozenski, J., Van Aerschot, A. & Herdewijn, P. (1999). *Tetrahedron*, **55**, 6527–6546.
Allart, B., Khan, K., Rosemeyer, H., Schepers, G., Hendrix, C., Rothenbacher, K., Seela, F., Van Aerschot, A. & Herdewijn, P. (1999). *Chem. Eur. J.* **5**, 2424–2431.
Berger, I., Kang, C., Sinha, N., Wolters, M. & Rich, A. (1996). *Acta Cryst.* D**52**, 465–468.
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.
Bramsen, J. B. *et al.* (2009). *Nucleic Acids Res.* **37**, 2867–2881.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Declercq, R., Van Aerschot, A., Read, R. J., Herdewijn, P. & Van Meervelt, L. (2002). *J. Am. Chem. Soc.* **124**, 928–933.
Dorsett, Y. & Tuschl, T. (2004). *Nature Rev. Drug Discov.* **3**, 318–329.
Elmén, J., Thonberg, H., Ljungberg, K., Frieden, M., Westergaard, M., Xu, Y., Wahren, B., Liang, Z., Ørum, H., Koch, T. & Wahlestedt, C. (2005). *Nucleic Acids Res.* **33**, 439–447.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fisher, M., Abramov, M., Van Aerschot, A., Rozenski, J., Dixit, V., Juliano, R. L. & Herdewijn, P. (2009). *Eur. J. Pharmacol.* **606**, 38–44.
Fisher, M., Abramov, M., Van Aerschot, A., Xu, D., Juliano, R. L. & Herdewijn, P. (2007). *Nucleic Acids Res.* **35**, 1064–1074.
Froeyen, M., Wroblowski, B., Esnouf, R., De Winter, H., Allart, B., Lescrinier, E. & Herdewijn, P. (2000). *Helv. Chim. Acta*, **83**, 2153–2182.
Guzaev, A., Salo, H., Azhayev, A. & Lonnberg, H. (1995). *Tetrahedron*, **51**, 9375–9384.
Heinemann, U. (1991). *J. Biomol. Struct. Dyn.* **8**, 801–811.
Jain, S., Zon, G. & Sundaralingam, M. (1989). *Biochemistry*, **28**, 2360–2364.
Jain, S., Zon, G. & Sundaralingam, M. (1991). *Biochemistry*, **30**, 3567–3576.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
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.
Meister, G. & Tuschl, T. (2004). *Nature (London)*, **431**, 343–349.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nauwelaerts, K., Fisher, M., Froeyen, M., Lescrinier, E., Van Aerschot, A., Xu, D., Kang, H., DeLong, R. L., Juliano, R. & Herdewijn, P. (2007). *J. Am. Chem. Soc.* **129**, 9340–9348.
Puri, N., Wang, X., Varma, R., Burnett, C., Beauchamp, L., Batten, D. M., Young, M., Sule, V., Latham, K., Sendera, T., Echeverri, C., Sachse, C. & Magdaleno, S. (2008). *Nucleic Acids Symp. Ser.* **52**, 25–26.
Robeyns, K., Herdewijn, P. & Van Meervelt, L. (2008). *J. Am. Chem. Soc.* **130**, 1979–1984.
Thota, N., Li, X. H., Bingman, C. & Sundaralingam, M. (1993). *Acta Cryst.* D**49**, 282–291.