

# Influence of the incorporation of a cyclohexenyl nucleic acid (CeNA) residue onto the sequence d(CGCGAATTCGCG)

Koen Robeyns<sup>1</sup>, Piet Herdewijn<sup>2</sup> and Luc Van Meervelt<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry, Katholieke Universiteit Leuven, Biomolecular Architecture and BioMacS, Celestijnenlaan 200F, B-3001 Leuven and <sup>2</sup>Laboratory of Medicinal Chemistry, Katholieke Universiteit Leuven, Rega Institute for Medical Research and BioMacS, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received October 8, 2007; Revised December 4, 2007; Accepted December 5, 2007

## ABSTRACT

Cyclohexene nucleic acids (CeNA), which are characterized by the presence of a cyclohexene moiety instead of a natural (deoxy)ribose sugar, are known to increase the thermal and enzymatic stability when incorporated in RNA oligonucleotides. As it has been demonstrated that even a single cyclohexenyl nucleoside, when incorporated in an oligonucleotide, can have a profound effect on the biological activity of the oligonucleotide, further research is warranted to study the complex of such oligonucleotides with target proteins. In order to analyse the influence of CeNA residues onto the helix conformation and hydration of natural nucleic acid structures, a cyclohexenyl-adenine building block (xAr) was incorporated into the Dickerson sequence CGCGA(xAr)TTCGCG. The crystal structure of this sequence determined to a resolution of 1.90 Å. The global helix belongs to the B-type family and shows a water spine, which is partially broken up by the apolar cyclohexene residue. The cyclohexene ring adopts the <sup>2</sup>E-conformation allowing a better incorporation of the residue in the dodecamer sequence. The crystal packing is stabilized by cobalt hexamine residues and belongs to space group P222<sub>1</sub>, never before reported for nucleic acids.

## INTRODUCTION

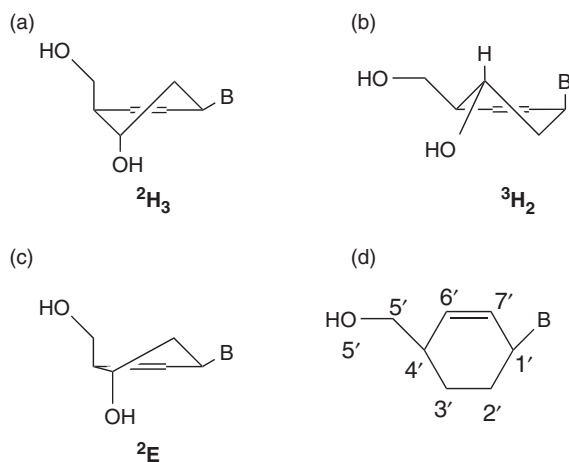
Cyclohexene nucleic acids (CeNAs) consist of a natural base and phosphate group, and are furthermore characterized by the replacement of the (deoxy)ribose sugar by a cyclohexene moiety. Recently cyclohexenyl nucleotide residues have been introduced in a double-stranded

siRNA, targeting the multi drug resistant gene (1). Interestingly incorporation of one cyclohexenyl residue in the sense and antisense strand of dsRNA leads to an enhanced biological activity. This has been demonstrated at the level of protein expression, at mRNA level and by functional assays (rhodamin accumulation and increased cytotoxicity of doxorubicin). Incorporation of cyclohexenyl nucleotides in RNA leads to an increase in its stability against nuclease degradation. The question is open if the excellent biological activity of the CeNA–RNA dimer is only due to its increased stability or that some structural reasons are involved. In perspective of projects in synthetic biology, we have evaluated the incorporation of cyclohexenyl nucleosides into DNA using DNA as template and DNA polymerase and reverse transcriptase as catalyst (2). The cyclohexenyl nucleoside are incorporated to a limited extend. The reason is currently under investigation. However, a better insight in the structure of a DNA duplex with an incorporated CeNA residue is the first step to understand the interaction between modified nucleic acid duplexes and polymerases. As a first study in the analysis of the influence of CeNA residues on the conformation of natural nucleic acids, we have solved the structure of the Dickerson sequence containing one CeNA unit, while a study of CeNA in complex with RNA is in progress.

The Dickerson sequence d(CGCGAATTCGCG) was the first DNA oligonucleotide sequence that was crystallized as a B-type double helix. Since it was first crystallized in 1980 (3,4), it was re-investigated numerous times and high-resolution structures have been published (5,6). Furthermore the Dickerson sequence is crystallized in different space groups [H3, Ca<sup>2+</sup>-form (7) H32, phosphoramidate backbone (8); P3<sub>2</sub>12, flipped out bases (9)], while the original Dickerson sequence crystallizes in the orthorhombic space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>.

Because of the existence of high-resolution structures and the intensity by which this sequence is studied,

\*To whom correspondence should be addressed. Tel: +32 16327609; Fax: +32 16327990; Email: luc.vanmeervelt@chem.kuleuven.be



**Figure 1.** CeNA in its two stable conformations: (a) the  ${}^2\text{H}_3$  half-chair conformation which mimics the C2'-*endo* conformation of the natural ribose sugar and (b) the  ${}^3\text{H}_2$  half-chair conformation mimicking the C3'-*endo* conformation. (c) The CeNA in the  ${}^2\text{E}$ -envelope conformation. (d) Numbering scheme used for the CeNA cyclohexene ring.

the Dickerson sequence d(CGCGAATTCGCG) is considered as ideal to study the influence of modified building blocks on the global helix conformation and hydration.

Here we present the structure of the modified Dickerson sequence CGCGA(xAr)TTCGCG, where (xAr) represents a cyclohexenyl-adenine nucleoside.

This modified Dickerson sequence with the incorporated cyclohexene residue has previously been studied using NMR and molecular modelling (10). The comparison between the chemical shifts from the modified and the native structure indicated that the presence of the cyclohexenyl-adenine residue has practically no effect on the global structure. The overall structure is comparable with the original B-type Dickerson structure. However, there were indications that the cyclohexenyl-adenine residue induces conformational mobility.

Circular dichroism (CD) spectra (10) suggested that the cyclohexene residue might adopt different conformations when incorporated in dsDNA and dsRNA helices, this mobility was confirmed by the NMR measurements. It was assumed that cyclohexene residues adopt the  ${}^3\text{H}_2$  conformation in A-type CeNA/RNA helices, and the  ${}^2\text{H}_3$  conformation in B-type CeNA/DNA helices (Figure 1a and b, respectively).

## MATERIALS AND METHODS

### Oligonucleotide synthesis

The CeNA nucleoside (11) with an adenine base, as well as the protected phosphoramidite nucleosides, that were used in the oligonucleotide synthesis, were synthesized by the Laboratory of Medicinal Chemistry at the Rega Institute (Leuven) (12). Assembly of the monomers into oligonucleotides was done as described by Maurinsh *et al.* (13).

### Crystallization conditions

The modified Dickerson dodecamer CGCGA(xAr)TTCGCG, with (xAr) the incorporated CeNA residue,

**Table 1.** Data collection statistics for the modified Dickerson sequence CGCGA(xAr)TTCGCG with (xAr) a cyclohexene residue

Space group	P222 <sub>1</sub>
Resolution range (Å)	20.00–1.90 (2.00–1.90) <sup>a</sup>
Measured reflections	24 668
Unique reflections	5476
Completeness (%)	99.4 (100.0) <sup>a</sup>
<i>R</i> <sub>merge</sub> (%)	5.0 (45.7) <sup>a</sup>
Multiplicity	4.5 (4.6) <sup>a</sup>
Mean <i>I</i> / $\sigma$ ( <i>I</i> )	15.3 (3.0) <sup>a</sup>
Average B-value (Å <sup>2</sup> )	38.0
Mosaicity (°)	1.1

<sup>a</sup>Values in parentheses are for the outermost shell.

was crystallized at 16°C by the ‘hanging-drop’ vapour diffusion method. The crystallization conditions were screened using a 24-matrix screen developed for oligonucleotides (14). The conditions that produced crystals were then further optimized by increasing the potassium concentration. This optimized crystallization condition contained 20 mM potassium cacodylate buffer (pH = 5.5), 10 mM cobalt hexamine, 44 mM potassium chloride, 6 mM sodium chloride, 5%(v/v) 2-methyl-2,4-pentanediol (MPD) and 0.75 mM ssDNA against a 35%(v/v) MPD stock solution. Block-like crystals were obtained after one day. No crystals could be obtained using the original crystallization conditions, Na<sup>+</sup>/Mg<sup>2+</sup>/spermine.

### Data collection

A dataset was collected at the EMBL synchrotron facility (Hamburg) on beamline BW7b ( $\lambda = 0.841$  Å) up to 1.85 Å resolution. The crystal used was about 0.1 × 0.1 × 0.2 mm. Data frames were taken with  $\Delta\phi$ -increments of 2° over a total  $\phi$ -range of 94° at a collection temperature of 100 K. The data were processed with ‘Mosflm’ (15) and scaled with ‘Scala’ (16) to 1.90 Å.

No signs of decay were observed. The data collection statistics up to 1.90 Å are summarized in Table 1.

The unit cell was determined as: *a* = 24.949 Å, *b* = 38.718 Å and *c* = 66.901 Å, which is similar to the unit cell found for the original Dickerson sequence and initially the space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> was assigned, following the same analogy.

### Structure determination

The space group could not be unambiguously determined based on systematic absences, although these indicated that space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> would be unlikely. The correct space group was assigned during the molecular replacement procedure.

Since the NMR studies (10) indicated that the incorporation of the cyclohexenyl-adenine residue has practically no effect on the global structure and also because there is no (NMR-) model of the modified Dickerson sequence available, the high-resolution structure (NDB code BDL084) was used as search model for the molecular replacement procedure. This was done in the assumption that the global crystal packing was similar to the original Dickerson structure. The molecular replacement program ‘Phaser’ (17) was used to find one

duplex in the space group  $P2_12_12_1$  and its enantiomers ( $P222$ ,  $P2_12_12$ ,  $P22_12_1$ ,  $P2_122_1$ ,  $P222_1$ ,  $P22_12$  and  $P2_122$ ). The automatic mode failed to give any results and a clear solution was only found by omitting the packing function step in the molecular replacement procedure in the space group  $P222_1$  where the helical dyad axis coincides with a crystallographic 2-fold rotation axis. A second helix was also found to have the dyad axis coinciding with a crystallographic 2-fold axis. Finally the asymmetric unit does not consist of one double helix, as observed for the native Dickerson sequence, but of two single stands, generating the double helices by applying the space group symmetry.

The structure was refined using 'Refmac' (18), using the standard dictionary files and a new dictionary entry was made for the cyclohexene residue. Target values for the modified sugar moiety were obtained from the Cambridge Structural Database (CSD) (19). The modified cyclohexene sugar moiety was fitted onto the sixth residue of the positioned MR model, to incorporate the modified residue into the molecular replacement solution. No ideal torsion angles were imposed for the CeNA residues, except for planarity restraints applied to the adenine bases and to the double bond in the cyclohexene ring. The chiral volumes of the sugar ring atoms were calculated and monitored during the refinement, as well as the bond distances and angles. The  $2|F_o| - |F_c|$  and  $|F_o| - |F_c|$  maps were also inspected to check the agreement between the model and the experimental electron density maps.

The terminal cytosine residue of the second strand (residue C1 of *duplex Two*) could not be fitted in the density maps and was found to be disordered, due to steric hindrance between the terminal cytosine residue and its neighbour generated by a crystallographic 2-fold axis.

Three cobalt hexamine residues were found when inspecting the density maps; a residue with ideal geometry was fitted into the  $2|F_o| - |F_c|$  density map using the program 'Coot' (20), and refined in Refmac in subsequent refinement cycles. One of the cobalt hexamine residues has an occupation factor of 50%. In total 75 water molecules were added using the 'arp\_waters' (ARP/wARP version 5.0) function in 'Refmac' (21). The diffuse solvent was modelled according to Babinet's principle (22). The convergence of the refinement was cross-validated using a 4.5%  $R_{free}$ -value test set. The final  $R$ - and  $R_{free}$ -values are 23.0% and 30.8%, respectively. Further refinement details are listed in Table 2.

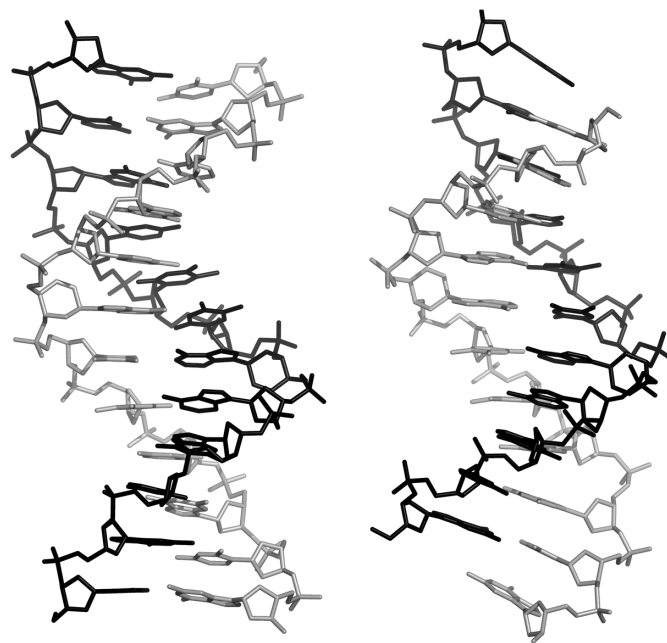
All molecular figures were created using the program PyMol (DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web <http://www.pymol.org>).

## RESULTS AND DISCUSSION

As the original Dickerson sequence, the modified sequence forms a right-handed, anti-parallel double helix in which all bases are engaged in standard Watson-Crick base pairs, with hydrogen bonding distances close to the standard values (Figure 2). On the contrary the modified sequence d(CGCGA(xAr)TTCGCG) crystallizes in the

**Table 2.** Refinement parameters for the modified Dickerson sequence CGCGA(xAr)TTCGCG with (xAr) a cyclohexene residue

Resolution range (Å)	17.76–1.90
Number of reflections	5209
Number of atoms	
Nucleic acid	468
Waters (treated as O)	75
Co(NH <sub>3</sub> ) <sub>6</sub> <sup>3+</sup>	21
Number of disordered atoms	19
Final $R$ -value (working set) (%)	23.0
Final $R$ -value (all data) (%)	23.3
Final $R_{free}$ -value (%)	30.8
RMS deviation from restraint target value	
Bond lengths (Å)	0.011
Angles (°)	1.803
Distances from restraint planes (Å)	0.009
Anti-bumping distance restraints (Å)	0.171
Mean $B$ -values (Å <sup>2</sup> )	44.2
DNA atoms	43.1
Solvent atoms	49.3



**Figure 2.** Stick representation showing the two duplexes present in the crystal structure, *duplex One* on the left, *duplex Two* on the right. The two strands are shown in silver and black, respectively and are connected through an internal 2-fold rotation axes. Notice the missing terminal cytosine residue in *duplex Two*.

orthorhombic space group  $P222_1$ , whereas the unmodified Dickerson sequence crystallizes in the space group  $P2_12_12_1$  with similar unit cell parameters. This is even more remarkably, knowing that this modified helix is the first oligonucleotide that crystallizes in this particular space group.

Since the asymmetric unit consists of two single-stranded oligonucleotides rather than one double helix, two distinct biological units (double helices) are present.

Each oligonucleotide strand generates its own duplex by applying the crystallographic 2-fold rotation axis present in the space group  $P222_1$ . Both oligonucleotide strands and the duplexes that they both generate are thus not equivalent. The two single-stranded oligonucleotides as well as their respective double helices will be named *One* and *Two* respectively, corresponding to chain A and chain B in the pdb file (NDB code BD0101). The numbering scheme for the CeNA residues is given in Figure 1d.

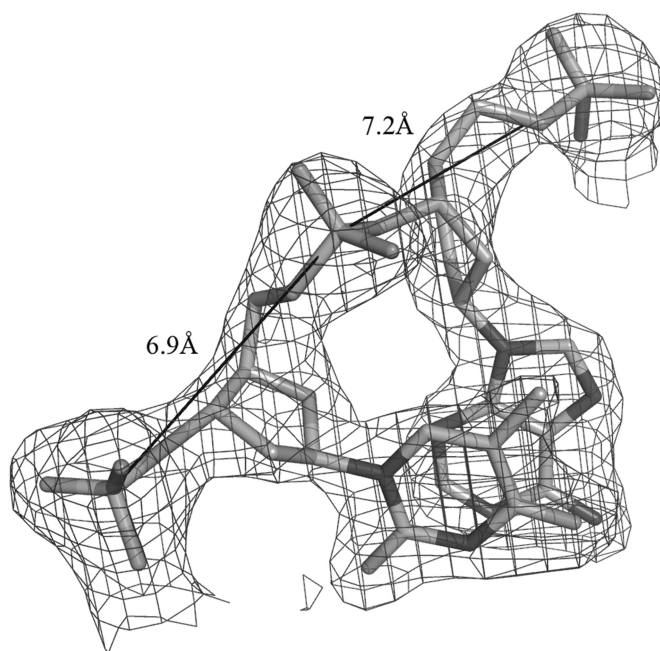
### Influence of the incorporated CeNA residue

The C1'–C1' distances from *duplex One* are all very similar to those found in *duplex Two* with an average distance of 10.7 and 10.6 Å, respectively. Also the  $\lambda$  angles (angle between the vectors C1'–C1' and C1'–N1/N9) are similar to one another, 53.6° and 53.8° for *duplex One* and *duplex Two*, respectively. Both the C1'–C1' distances and  $\lambda$  angles are comparable with the base-pair geometry found in high-resolution A-DNA and B-DNA crystal structures (23).

The overall structural features of both helices are similar and when superposing both duplexes onto each other, the root mean square (RMS) deviation is  $\sim 1.13$  Å, with the largest positional differences observed in the terminal residues. Similar results are found when superposing *duplex One* and *Two* onto the unmodified high-resolution Dickerson sequence (NDB code BDL084), a RMS deviation of 1.12 and 1.41 Å, with a maximum deviation of 4.85 and 5.38 Å is found, for *duplex One* and *duplex Two*, respectively. Comparing only the central GAATTC residues of the high-resolution and the modified dodecamer, the RMS deviation decreases considerably (0.58 and 0.56 Å for *duplex One* and *Two*, respectively), highlighting the rigidity of the central residues despite the modification.

The ring puckering of the five-membered deoxyribonucleotides ranges from C2'-endo over C1'-exo and C4'-exo to C3'-endo for both *duplex One* and *duplex Two*. In order to minimize the deviation from the global B-type helix, the cyclohexene ring is forced into the  $^2E$ -envelope conformation, rather than the suggested  $^2H_3$  half-chair conformation. The average intra-strand phosphate distance in B-type DNA is  $\sim 7.0$  Å (24). The C3' atom of the CeNA residue is pushed into the plane, formed by the double bond, to reduce the intra-strand phosphate distance to the average value found in B-type DNA and to preserve the global B-type helix. The intra-strand phosphate distance between the phosphate atom from the CeNA residue and phosphate atom attached to its 3'-end is 7.2 Å for both *duplexes One* and *Two* (Figure 3).

Helical parameters (Table 3) were calculated with the 3DNA package (25) and values similar to the ones found in unmodified B-type DNA are observed. The  $x$ -displacement is almost 0 Å (0.10 and  $-0.64$  Å for *duplex One* and *duplex Two*, respectively). The base pairs are inclined over 3.4° (*duplex One*) and 6.5° (*duplex Two*) and show on average a propeller twist of  $-11.0^\circ$  and a buckle of 0.0°. The average roll and slide for *duplex One* are 1.5° and 0.26 Å, respectively and 3.7° and  $-0.03$  Å, respectively



**Figure 3.** Stick representation and electron density of residues 6 (xAr) (CeNA) and T7. The CeNA sugar ring adopts the  $^2E$  envelope conformation rather than the  $^2H_3$  half-chair conformation. The C3' atom of the CeNA moiety is pushed into the plane formed by the double bond in order to prevent further stretching of the intra-strand phosphate distance, now 7.2 Å, and to preserve the global B-type conformation. Electron density maps are calculated with refined phases and contoured at  $1.1\sigma$  above the average background.

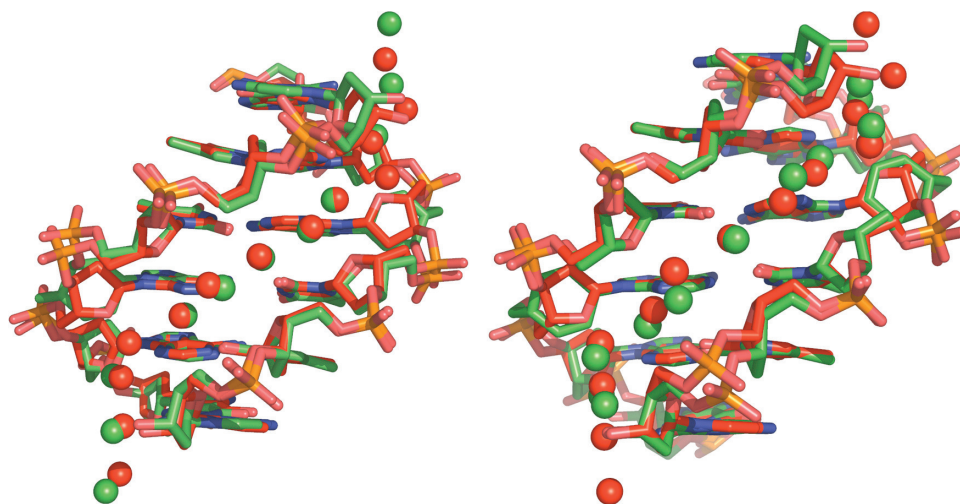
**Table 3.** Helical parameters for the modified Dickerson sequence CGCGA(xAr)TTCGCG with (xAr) a cyclohexene residue and comparison with classical B-type DNA

	Modified Dickerson sequence		
	B-DNA <sup>a</sup>	<i>Duplex One</i>	<i>Duplex Two</i>
$x$ -displacement (Å)	0.05	0.15	$-0.65$
Inclination (°)	2.1	3.4	6.5
Rise (Å)	3.32	3.3	3.2
Twist (°)	36.0	35.4	35.2
Slide (Å)	0.23	0.26	0.0
Roll (°)	0.6	1.5	3.66
Propeller twist (°)	$-11.4$	$-11.0$	$-11.0$
Buckle (°)	0.5	0.0	0.0

<sup>a</sup>Helical parameters for A and B type DNA are taken from (Olson *et al.* (2001)).

for *duplex Two*. *Duplex One* is characterized by a rise of 3.26 Å and a twist of 35.4°, similar to the rise and twist found in *duplex Two*, 3.22 Å and 35.2°, respectively. The twist found in both *duplex One* and *duplex Two* corresponds with about 10.2 base pairs in a full turn.

Since the helical parameters for both duplexes are very comparable, the minor and major groove widths are also very similar. The minor groove widths were calculated as the closest separation between the O4' atoms (or the C7' atoms for the CeNA residues) and were corrected for the



**Figure 4.** Stick representation of the central GAATTC residues with the hydration spine shown as spheres. The modified Dickerson sequence with the incorporated CeNA residue (shown in green) is superposed onto the high-resolution structure [Shui *et al.* (5)] (shown in red). *duplex One* is shown on the left, *duplex Two* on the right.

van der Waals radii of the corresponding atoms (oxygen atoms: 1.52 Å, carbon atoms: 1.70 Å). An average value of 5.38 and 5.14 Å is observed for *duplex One* and *duplex Two*, respectively (minor groove BDL084, 4.89 Å). The major groove was calculated as the distance between phosphate atom of residue P(*n*) and the phosphate atom P(*n* + 2) on the opposite strands, and the distance was corrected for the sum of the van der Waals radii of the two phosphate groups (5.80 Å). Average values of 10.21 and 9.72 Å are observed for *duplex One* and *duplex Two*, respectively (major groove BDL084, 9.78 Å). Both the minor and major groove broaden from the central region outwards.

The structure of the unmodified B-type helix CGCGAATTCGCG is characterized by a spine of water molecules that are located in the minor groove. This was already described by Drew and Dickerson (1981) (26) and confirmed by the high-resolution structure (5).

In Figure 5, the schematic hydrogen network in the minor groove is shown. This schematic representation of the water network is to be compared with a similar illustration for the high-resolution structure [figure 5 in Shui *et al.* (5); NDB code BDL084]. When comparing both figures it is clear that the water spine is partially preserved in the modified duplexes (missing water molecules in the water spine of the modified dodecamer are indicated with X). Moreover when superposing the modified duplexes onto the high-resolution structure it is unmistakable that the water molecules occupy similar sites, despite the introduced modification (Figure 4).

While the hydration spine extends asymmetrically in the unmodified duplex, with respect to the central AATT region, the water spine in the modified duplexes is symmetrical due to the internal 2-fold rotation axis. This is clearly seen in Figure 5 where the water molecules O3 and O6 are located onto a 2-fold axis (Symmetry generated residues and water atoms are indicated with an \*). This causes an extension of the water spine at one

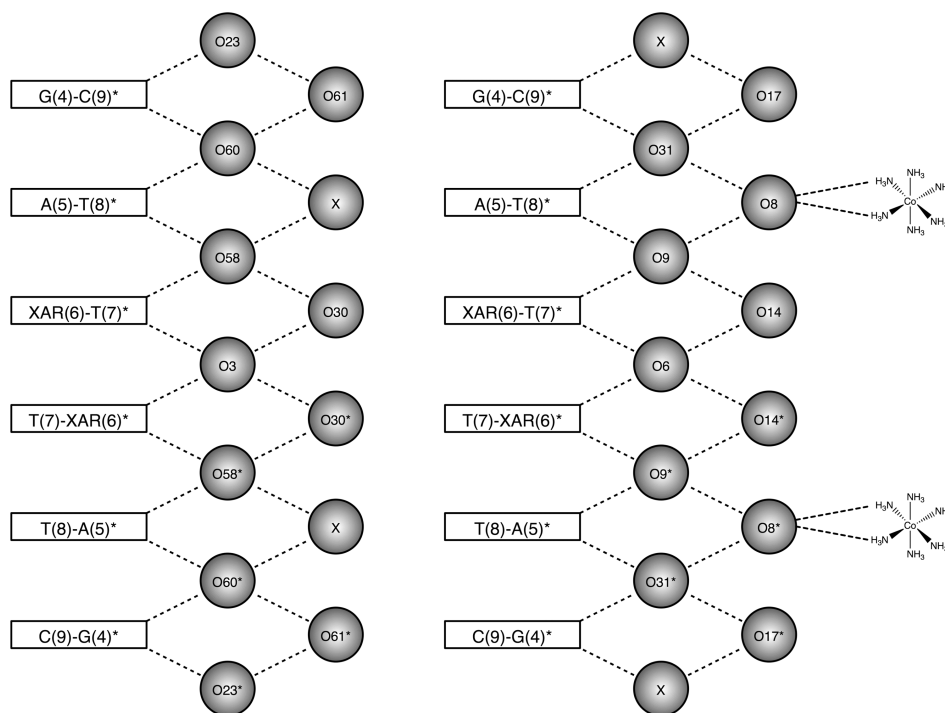
side of the helix and the terminal water molecules of the hydration spine deviate therefore more from the average sites of the high-resolution dodecamer.

In *duplex One*, one water molecule of the hydration spine could not be located in the density. This water molecule and its symmetry-related equivalent, which correspond to water molecule 55 and 67 in the high-resolution structure (BDL084), would lie next to the modified CeNA residue and it is believed that the introduction of the apolar cyclohexene ring results in the absence of the suitable O4'-acceptor site in the sugar ring. In *duplex Two*, however, this water molecule is present, mainly because a cobalt hexamine residue acts as a donor/acceptor for the water molecule (HOH 8) in the hydration spine (Figure 5), rather than the oligonucleotide.

The average temperature factor of the water molecules in the hydration spine of *duplex One* (49 Å<sup>2</sup>) is only slightly lower than the average temperature factor for all water molecules (50 Å<sup>2</sup>). In *duplex Two*, however the average temperature factor for the spine water molecules is much lower (41 Å<sup>2</sup>). This can be partially understood by the stabilizing effect of a cobalt hexamine residue on the hydration spine in *duplex Two*.

The base-stacking patterns in the modified dodecamer show a great resemblance with unmodified B-type DNA. No inter strand overlap between the base rings is observed and the intra-strand base stacking diminishes moving from the central base-pair steps outwards, with almost no overlap observed for the terminal base-pair steps.

It appears that the two Co(NH<sub>3</sub>)<sub>6</sub><sup>3+</sup> residues, which are located close to the modified CeNA residue, positively influence the base-stacking interaction of the central step in *duplex Two*. The intra-strand ring overlap for the central step in *duplex One* is ~1.37 Å<sup>2</sup> while for *duplex Two* an overlap of 2.41 Å<sup>2</sup> is observed. This increase in base-stacking interaction can be attributed solely to the larger negative slide found in *duplex Two* (-0.98 Å) as compared to the value of -0.38 Å in *duplex One*.



**Figure 5.** Schematic representation of the water network observed in the minor groove of the modified Dickerson sequence CGCGA(xAr)TTCGCG. This figure is to be compared with figure 5 in Shui *et al.* (5) where a similar illustration is shown for the high-resolution structure (NDB code BDL084). On the left *duplex One* is shown, on the right *duplex Two*. Missing water atoms, compared to the water-spine found in the high-resolution structure are indicated with X, symmetry generated residues and water atoms are indicated with \*. Notice in particular the presence of the  $\text{Co}(\text{NH}_3)_6$  residues, which stabilize the water spine in *duplex Two*. The water spine in *duplex One* and *duplex Two* stretches symmetrically to either side of the duplex, with the central water molecules (O3 and O6) being positioned on a crystallographic 2-fold axis.

The phosphate group in *duplex Two* adopts the  $B_I$  conformation ( $\varepsilon$ - $\zeta$  difference:  $-58.7^\circ$ ) whereas the phosphate group in *duplex One* is found to be in the  $B_{II}$  conformation ( $\varepsilon$ - $\zeta$  difference:  $14.7^\circ$ ). In general the  $B_{II}$  conformation appears to be associated with less favorable intra-strand base-stacking energies, as is observed in *duplex One*. The more positive slide found in *duplex One* can also be ascribed to the  $B_{II}$  conformation of the phosphate group (27). It appears therefore that the cobalt hexamine residues, attached to *duplex Two*, facilitate the conversion to the  $B_I$  configuration for the modified CeNA residue. This conversion increases the base pair overlap and results in a more stable configuration.

#### The CGCGAATTCGCG dodecamer in a new crystal form

The double helices stack on top of each other to form pseudo-continuous helices, similar to the pseudo-continuous helices found in the unmodified Dickerson structure. In the packing two such columns are observed, consisting of *duplexes One* and the other consisting of *duplexes Two*. Both columns are aligned along the *c*-axis. Unlike the natural dodecamer, which crystallizes in the space group  $P2_12_12_1$ , the CeNA modification crystallizes in space group  $P222_1$ . The introduction of 2-fold axes applies more restraints to the double helices and to their packing. These 2-fold axes are responsible for the disorder of the first cytosine residue of *duplex Two*. Rebuilding the disordered cytosine residue results in unfavorable contacts

between the phosphate group atoms of the second residue of *duplex Two* and its symmetry equivalent.

The Dickerson sequence has already been crystallized in different space groups. It was reported that adding calcium ions changed the packing and the structure (28). Here the terminal cytosine residue was also found to be disordered, due to the crystal packing.

The presence of other cations ( $\text{K}^+$ ,  $\text{Cs}^+$  or  $\text{Tl}^+$ ) does not change the crystal packing of the Dickerson sequence (29–31).

In the crystal packing of the Dickerson sequence with the incorporated CeNA residues, the dyad axes coincide with the crystallographic 2-fold axes of the  $P222_1$  space group. This is also the case with the Dickerson sequence crystallized in the space group  $P3_212$  (9). Here the terminal bases are flipped out, with the duplexes stacking on top of each other.

Modifying the sugar-phosphate backbone does not normally change the crystal packing (32–34). However, when replacing the sugar-phosphate backbone by a phosphoramidate backbone ( $\text{O}3' \rightarrow \text{N}3'$ ) the Dickerson sequence crystallizes in the  $H32$  space group (twinned structure with three helices in the asymmetric unit) (8). Here the helices stack in continuous columns contrary to the pseudo-continuous helices observed in the native Dickerson sequence.

The packing in our structure is greatly stabilized by the three cobalt hexamine residues present in the crystal structure. Two cobalt hexamine residues, which are located

~10 Å from each other, are highly hydrated and are situated in-between the pseudo-continuous columns. These two cobalt hexamine residues are connected with six (symmetry related) single-strand oligonucleotides through hydrogen bonds with bridging water molecules. The  $\text{Co}(\text{NH}_3)_6^{3+}$  residues not only stabilize the terminal residues of *duplex One* but the modified residue in *duplex Two* as well. The phosphate group attached to the 3'-end of the modified CeNA residue in *duplex Two* is involved in hydrogen bonding with both  $\text{Co}(\text{NH}_3)_6^{3+}$  residues, the cobalt hexamine residues hereby stabilizes the global B-type conformation of the duplex and the  ${}^2\text{E}$ -conformation of the cyclohexene residue in *duplex Two*.

The third  $\text{Co}(\text{NH}_3)_6^{3+}$  residue, which has an occupancy of 50%, does not stabilize the modified residue, but is located close to the disordered residue in *duplex Two*. It is believed that the third  $\text{Co}(\text{NH}_3)_6^{3+}$  residue replaces the disordered cytosine residue of *duplex Two* by stabilizing its Watson–Crick complement (Guanine 12 of *duplex Two*) through hydrogen bonds with the O6 and N7 atoms of the guanine residue. Furthermore this third  $\text{Co}(\text{NH}_3)_6^{3+}$  residue interacts with symmetry-related duplexes, stabilizing the crystal packing.

## CONCLUSIONS

The modified Dickerson sequence CGCGA(xAr)TT CGCG with one incorporated cyclohexene adenine residue crystallizes in the orthorhombic space group P222<sub>1</sub>, and data were collected up to 1.90 Å. The helix was found to be an anti-parallel right-handed duplex, with both strands connected by a crystallographic 2-fold rotation axis. The single cyclohexene residue causes almost no conformational changes; the global helix belongs to the B-type family. Standard Watson–Crick base pairs are observed and the backbone torsion angles are similar to those found in the original Dickerson sequence.

The conformation of the furanose sugar moieties varies from C3'-*endo* over O4'-*endo* to C3'-*exo*, and the modified sugar ring adopts the  ${}^2\text{E}$  conformation, rather than the expected  ${}^2\text{H}_3$  half-chair conformation.

The water spine found in the original Dickerson sequence is partially broken by the introduction of the apolar cyclohexene residues, especially in *duplex One*, in *duplex Two* the spine is better preserved.

The presence of two cobalt hexamine residues greatly stabilizes the crystal packing and the modified residues. These cobalt hexamine residues are highly hydrated and show hydrogen bonds with phosphate groups as well as with DNA bases.

The reported space group is quite unusual, not only because the Dickerson sequence has never before been crystallized in the space group P222<sub>1</sub>, but also because this space group has never been reported for nucleic acid sequences.

The odd crystal packing also causes the terminal cytosine residue of one of the two chains to be disordered. The crystal packing is however quite similar to the one observed in the native Dickerson sequence, with duplexes stacking in pseudo-continuous helices. There is also

resemblance with other types of crystal packing. As in the structure reported by Liu *et al.* (1999) (NDB code BD0014), the terminal cytosine residue is disordered and similar to the structure of Johansson *et al.* (9) (NDB code BD0032), the dyad axis coincides with a crystallographic 2-fold axis.

Although designed to be more rigid and to belong to the A-type family, the cyclohexene residue adopts its conformation to a B-type DNA when incorporated in the Dickerson sequence. Further studies will be established to investigate the influence of CeNA on a duplex involving RNA as the natural partner to obtain a better insight in the recognition of CeNA by the RISC complex.

The structural characterization of a DNA duplex with an incorporated CeNA residue is a first important step in the understanding of the interaction between CeNA and polymerases when synthesizing CeNA oligomers on a DNA template.

## Accession codes

Final coordinates and structure factor amplitudes have been deposited with the Protein Data Bank (code 2P8D) and Nucleic Acid Data Bank (code BD0101).

## ACKNOWLEDGEMENTS

We thank the staff of the EMBL Hamburg Outstation for help with the synchrotron experiments. BioMacS, the K.U. Leuven Interfaculty Centre for Biomacromolecular Structure, is supported by the Impulse Project of the K.U. Leuven. Funding to pay the Open Access publication charges for this article was provided by K.U. Leuven.

*Conflict of interest statement.* None declared.

## REFERENCES

- Nauwelaerts, K., Fisher, M., Froeyen, M., Lescrier, E., VanAerschot, A., Xu, D., DeLong, R., Kang, H., Juliano, R.L. *et al.* (2007) Structural characterization and biological evaluation of small interfering RNAs containing cyclohexenyl nucleosides. *J. Am. Chem. Soc.*, **129**, 9340–9348.
- Kempeneers, V., Renders, M., Froeyen, M. and Herdewijn, P. (2005) Investigation of the DNA-dependent cyclohexenyl nucleic acid polymerization and the cyclohexenyl nucleic acid-dependent DNA polymerization. *Nucleic Acids Res.*, **33**, 3828–3836.
- Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl Acad. Sci. USA*, **78**, 2179–2183.
- Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980) Crystal-structure analysis of a complete turn of B-DNA. *Nature*, **287**, 755–758.
- Shui, X.Q., Sines, C.C., McFail-Isom, L., VanDerveer, D. and Williams, L.D. (1998) Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry*, **37**, 16877–16887.
- Williams, L.D., Sines, C.C., McFail-Isom, L., Howerton, S.B. and VanDerveer, D. (2000) Cations mediate B-DNA conformational heterogeneity. *J. Am. Chem. Soc.*, **122**, 11048–11056.
- Liu, J. and Subirana, J.A. (1999) Structure of d(CGCGAATTCGCG) in the presence of Ca<sup>2+</sup> ions. *J. Biol. Chem.*, **274**, 24749–24752.
- Tereshko, V., Gryaznov, S. and Egli, M. (1998) Consequences of replacing the DNA 3'-oxygen by an amino group: High-resolution

- crystal structure of a fully modified N3' and P5' phosphoramidate DNA dodecamer duplex. *J. Am. Chem. Soc.*, **120**, 269–283.
9. Johansson,E., Parkinson,G. and Neidle,S. (2000) A new crystal form for the dodecamer C-G-C-G-A-A-T-T-C-G-C-G: symmetry effects on sequence-dependent DNA structure. *J. Mol. Biol.*, **300**, 551–561.
  10. Wang,J., Verbeure,B., Luyten,I., Lescrinier,E., Froeyen,M., Hendrix,C., Rosemeyer,H., Seela,F., Van Aerschot,A. *et al.* (2000) Cyclohexene nucleic acids (CeNA): serum stable oligonucleotides that activate RNase H and increase duplex stability with complementary RNA. *J. Am. Chem. Soc.*, **122**, 8595–8602.
  11. Wang,J. and Herdewijn,P. (1999) Enantioselective synthesis and conformational study of cyclohexene carbocyclic nucleosides. *J. Org. Chem.*, **64**, 7820–7827.
  12. Gu,P., Griebel,C., Van Aerschot,A., Rozenski,J., Busson,R., Gais,H.J. and Herdewijn,P. (2004) Synthesis of enantiomeric-pure cyclohexenyl nucleoside building blocks for oligonucleotide synthesis. *Tetrahedron*, **60**, 2111–2123.
  13. Maurinsh,Y., Rosemeyer,H., Esnouf,R., Medvedovici,A., Wang,J., Ceulemans,G., Lescrinier,E., Hendrix,C., Busson,R. *et al.* (1999) Synthesis and pairing properties of oligonucleotides containing 3-hydroxy-4-hydroxymethyl-1-cyclohexenyl nucleosides. *Chem. Eur. J.*, **5**, 2139–2150.
  14. Berger,I., Kang,C.H., Sinha,N., Wolters,M. and Rich,A. (1996) A highly efficient 24-condition matrix for the crystallization of nucleic acid fragments. *Acta Cryst. D*, **52**, 465–468.
  15. Leslie,A.G.W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, **26**.
  16. Evans,P.R. (1997) Scala. *Joint CCP4 + ESF-EACBM Newsletter on Protein Crystallography*, **33**.
  17. McCoy,A.J., Grosse-Kunstleve,R.W., Storoni,L.C. and Read,R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Cryst. D*, **61**, 458–464.
  18. Murshudov,G.N., Vagin, A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by maximum-likelihood method. *Acta Cryst. D*, **53**, 240–255.
  19. Allen,F.H. (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Cryst. B*, **58**, 380–388.
  20. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst. D*, **60**, 2126–2132.
  21. Lamzin,V.S. and Wilson,K.S. (1993) Automated refinement of protein models. *Acta Cryst. D*, **49**, 129–147.
  22. Moews,P.C. and Kretsinger,R.H. (1975) Refinement of structure of carp muscle calcium-binding parvalbumin by model-building and difference Fourier-analysis. *J. Mol. Biol.*, **91**, 201–225.
  23. Olson,W.K., Bansal,M., Burley,S.K., Dickerson,R.E., Gerstein,M., Harvey,S.C., Heinemann,U., Lu,X.J., Neidle,S. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.
  24. Saenger. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
  25. Lu,X.J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
  26. Drew,H.R. and Dickerson,R.E. (1981) Structure of a B-DNA dodecamer. III. Geometry of hydration. *J. Mol. Biol.*, **151**, 535–556.
  27. Madhumalar,A. and Bansal,M. (2005) Sequence preference for BI/II conformations in DNA: MD and crystal structure data analysis. *J. Biomol. Struct. Dyn.*, **23**, 13–27.
  28. Liu,J., Malinina,L., Huynh-Dinh,T. and Subirana,J.A. (1998) The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG). *FEBS Lett.*, **438**, 211–214.
  29. Howerton,S.B., Sines,C.C., VanDerveer,D. and Williams,L.D. (2001) Locating monovalent cations in the grooves of B-DNA. *Biochemistry*, **40**, 10023–10031.
  30. Sines,C.C., McFail-Isom,L., Howerton,S.B., VanDerveer,D. and Williams,L.D. (2000) Cations mediate B-DNA conformational heterogeneity. *J. Am. Chem. Soc.*, **122**, 11048–11056.
  31. Woods,K.K., McFail-Isom,L., Sines,C.C., Howerton,S.B., Stephens,R.K. and Williams,L.D. (2000) Monovalent cations sequester within the A-tract minor groove of [d(CGCGAATTCGCG)]<sub>2</sub>. *J. Am. Chem. Soc.*, **122**, 1546–1547.
  32. Berger,I., Tereshko,V., Ikeda,H., Marquez,V.E. and Egli,M. (1998) Crystal structures of B-DNA with incorporated 2'-deoxy-2'-fluoro-arabino-furanosyl thymines: implications of conformational pre-organization for duplex stability. *Nucleic Acids Res.*, **26**, 2473–2480.
  33. Boggon,T.J., Hancox,E.L., McAuleyHecht,K.E., Connolly,B.A., Hunter,W.N., Brown,T., Walker,R.T. and Leonard,G.A. (1996) The crystal structure analysis of d(CGCGAASSCGCG)(2), a synthetic DNA dodecamer duplex containing four 4'-thio-2'-deoxythymidine nucleotides. *Nucleic Acids Res.*, **24**, 951–961.
  34. Portmann,S., Altmann,K.H., Reynes,N. and Egli,M. (1997) Crystal structures of oligodeoxyribonucleotides containing 6'-alpha-methyl and 6'-alpha-hydroxy carbocyclic thymidines. *J. Am. Chem. Soc.*, **119**, 2396–2403.