Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a DNA duplex with all-AT base pairs

Núria Valls, Michael Richter and Juan A. Subirana*

Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Avinguda Diagonal 647, E-08028 Barclelona, Spain

Correspondence e-mail: juan.a.subirana@upc.edu

Received 7 July 2005
Accepted 19 September 2005

PDB Reference: $d(AAATATTT)_2$, 2a2t, r2a2tsf.

NDB Reference: BD0088.

The structure of the oligonucleotide duplex $d(AAATATTT)_2$ has been obtained in two crystal forms. In both cases the duplexes show Watson–Crick base pairs and are organized as a helical arrangement of stacked oligonucleotides. The helices contain either 11 duplexes in seven turns (form I) or 14 duplexes in nine turns (form II). As a result, the crystals have rather large unit cells. Such an organization of DNA duplexes has not been previously described in oligonucleotide crystal structures. The columns of stacked duplexes are organized in a pseudohexagonal arrangement. They do not show any direct lateral interactions; instead, these are indirectly mediated by shared and disordered hydrated ions. Such interactions are similar to those found in DNA fibres. The structural parameters of the individual base steps coincide with those found in other structures which contain CG terminal base pairs.

1. Introduction

At present, over 600 oligonucleotides have been crystallized (Berman *et al.*, 1992). However, as noted elsewhere (Subirana, 2003), most of them have sequences which start and end with CG base pairs, which have a strong influence on the crystal structure as they establish interhelical contacts (Dickerson *et al.*, 1987). Therefore, it appears of interest to study the crystal structure of DNA fragments which only contain AT base pairs. Our first studies in this direction, with oligonucleotides with an alternating sequence, yielded unexpected results: d(ATATAT) shows Hoogsteen base pairs (Abrescia *et al.*, 2002, 2004), whereas d(ATATATATATAT) forms coiled coils (Campos *et al.*, 2005).

It is also well known that noncoding regions of the genome, including introns, are rich in AT base pairs (Lander *et al.*, 2001). Scaffold-associated regions are also AT-rich (Liebich *et al.*, 2002). The study of single sequence repeats in some genomes (Karaogly *et al.*, 2005) also shows a predominance of AT-rich repeats. It is worth noting that in the case of fungi (Karaogly *et al.*, 2005) some sequences are common, whereas other sequences are rare. For example, of the tetranucleotide repeats AAAA, ATAT and AAAT are abundant, while AATT is practically absent. Since the biological role and structure of such AT-rich regions is unknown, the study of oligonucleotides with such sequences is of particular interest.

In this paper, we present the structure of the oligonucleotide d(AAATATTT) with Watson–Crick base pairs. It crystallizes in a new way which to our knowledge has not been previously described. We have found two crystal forms (I and II) and in both of them the duplexes are organized as a helical arrangement of stacked oligonucleotides with several duplexes (either 5.5 or 7) in the asymmetric unit. No obvious

lateral interactions between duplexes in the crystal appear to be present, but the oligonucleotides are organized as columns with precise interhelical arrangements similar to those found in DNA fibres (Langridge *et al.*, 1960). In fact, our results provide information on the way DNA molecules interact sidewise through their electrostatic field.

2. Methods and materials

2.1. Synthesis and crystallization

The deoxyoligonucleotide d(AAATATTT) was synthesized as the ammonium salt on an automatic synthesizer by the phosphoramidite method and was purified by gel filtration and reverse-phase HPLC. Two different kinds of crystals were grown under very similar conditions using the hanging-drop method by vapour diffusion at 286 K. Crystals of form I grew in a drop containing 0.5 m*M* DNA duplex, 25 m*M* sodium cacodylate buffer pH 6.5, 5 m*M* $MnCl_2$, 1 m*M* spermine and 10% MPD equilibrated against a 25% MPD reservoir. Thin laminar crystals of approximately 0.15 × 0.75 mm in size appeared in about two weeks. Form II crystals were grown in a drop containing 1 m*M* DNA duplex, 25 m*M* sodium cacodylate buffer pH 6.5, 8 m*M* $MnCl_2$, 1 m*M* spermine and 10% MPD equilibrated against 30% MPD reservoir. In this case long crystals of ∼1 × 0.05 × 0.05 mm in size appeared within a few days.

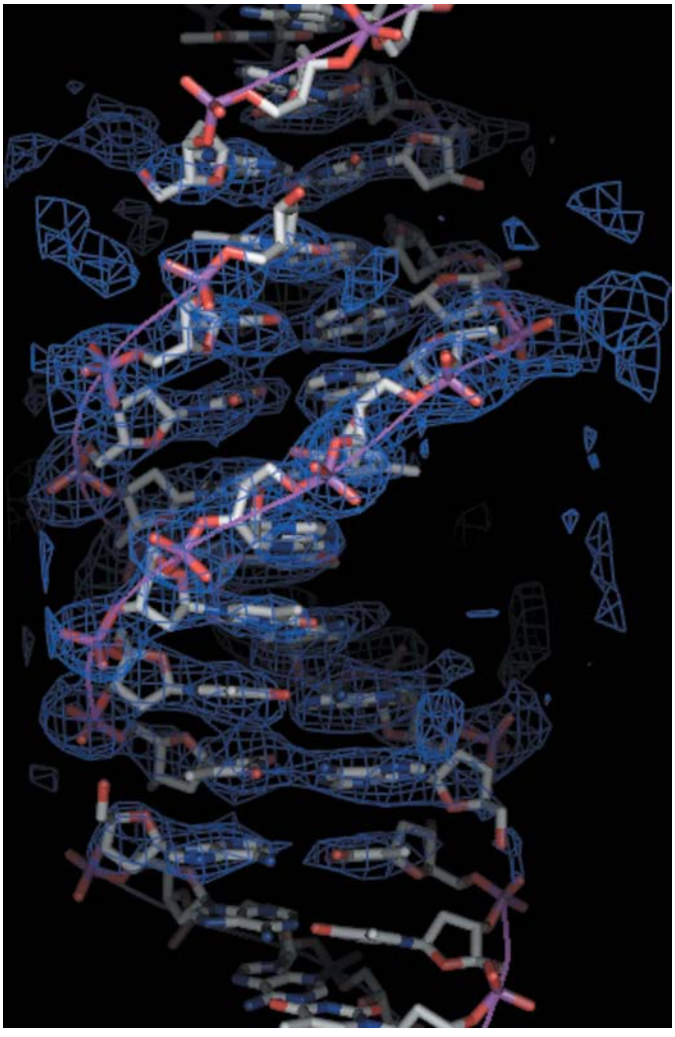

Figure 1
$2mF_o - DF_c$ electron-density map of duplex *B* at the 1σ level. The program *PyMOL* (DeLano, 2002) was used for the preparation of this figure and Fig. 2.

Table 1
Crystallographic and refinement statistics.

Values in parentheses are for the last shell.

	Form I (model *A*)	Form II
Wavelength (Å)	0.9794	0.9794
Temperature (K)	120	120
Space group	*C*2	*P*321
Resolution range (Å)	15–3.1 (3.21–3.1)	50–6.0
Unit-cell parameters (Å, °)	$a = 147.62$, $b = 24.98$, $c = 82.14$, $\beta = 90.51$	$a = b = 43.0$, $c = 367$, $\gamma = 120$
Unique reflections	5710	—
Completeness (%)	99.3 (97.3)	—
R_{sym}†	0.063 (0.184)	—
Contents of ASU	5.5 DNA duplexes, 6 Mn^{2+}, 6 cacodylates, 16 H_2O	7 DNA duplexes
R_{work}‡/R_{free}§	0.272/0.357	—

† $R_{sym}(I) = \sum_{hkl}\sum_j |I_j(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_j I_j(hkl)$ calculated for the whole data set. ‡ $R_{work} = \sum_{hkl} |F_o(hkl) - kF_c(hkl)| / \sum_{hkl} F_o(hkl)$. § R_{free} is the R factor of reflections used for cross-validation in the refinement.

2.2. Data collection and structure determination

Crystals were flash-cooled at 120 K and diffraction data were collected at the ESRF synchrotron beamline BM16 on a MAR CCD detector at λ = 0.9794 Å. Data for form I were integrated and scaled with *HKL*2000 (Otwinowski & Minor, 1997). Crystallographic data and refinement parameters are given in Table 1. A complete set of oscillation patterns was also collected for form II crystals, from which the unit cell and space group were determined. No attempt to process the data was carried out as the resolution was only ∼6 Å.

As a first approach to structure determination of form I crystals the program *MOLREP* (Vagin & Teplyakov, 1997) was used. The starting model was an ideal duplex structure of d(AAATATTT) with constant twist ($\omega = 36°$) and rise (3.25 Å) created with the program *TURBO-FRODO* (Roussel *et al.*, 1998). We could find all 5.5 duplexes in the asymmetric unit, but refinement did not progress using this model. Apparently, the *MOLREP* solution did not place the duplexes with the correct relative twist. In order to improve the positions of individual duplexes in a column, we assumed that the value of twist between the ends of two neighbouring duplexes would be the same as that found in other cases where the terminal base pairs are AT. There are only four structures available of this type, which are from dodecamers with related sequences crystallized in different space groups (Rozenberg *et al.*, 1998; Hizver *et al.*, 2001). Interestingly, in all of them the twist value ω_T between the ends of consecutive duplexes is negative, about −20°; they do not form a pseudo-continuous helix, which would require a positive value of ω_T of ∼36°. Thus, we assumed that in our case ω_T would also be close to

Table 2
Twist values ω (°) of the six independent duplexes in form I.

Values in parentheses indicate the standard deviations.

Base step	*A*	*B*	*C*	*D*	*E*	*F*	Average
AA/TT	36.0	36.2	35.4	41.9	28.5	34.1	35.4 (4.3)
AA/TT	36.0	37.3	43.3	37.0	37.5	38.8	38.3 (2.6)
AT/AT	34.4	35.5	29.3	32.8	34.0	33.9	33.3 (2.2)
TA/TA	39.5	35.7	41.9	37.6	38.8	37.9	38.6 (2.1)
AT/AT	33.6	34.1	38.9	31.8	35.2	34.1	34.6 (2.4)
TT/AA	37.5	37.3	39.1	40.9	37.8	39.0	38.6 (1.4)
TT/AA	34.0	37.2	31.4	38.6	37.3	34.8	35.5 (2.7)
$\langle\omega\rangle$	35.9 (2.1)	36.2 (1.2)	37.0 (5.2)	37.2 (3.8)	35.6 (3.5)	36.0 (2.4)	36.3 (0.7)
TA(ω_T)†	−23.0	−24.1	−28.0	−29.3	−23.7		−25.6
Ω‡		228.0	229.2	231.4	231.5	225.4	229.1

† ω_T is the twist value between the terminal base steps of consecutive oligonucleotides. It corresponds to a TA step without binding phosphates. ‡ Ω is the overall twist of one oligonucleotide with respect to the previous one in a column. It has been measured taking the first base pair in each oligonucleotide as a reference. The average value coincides with that calculated from the overall helical structure: $\Omega = 7 \times 360°/11 = 229.1°$.

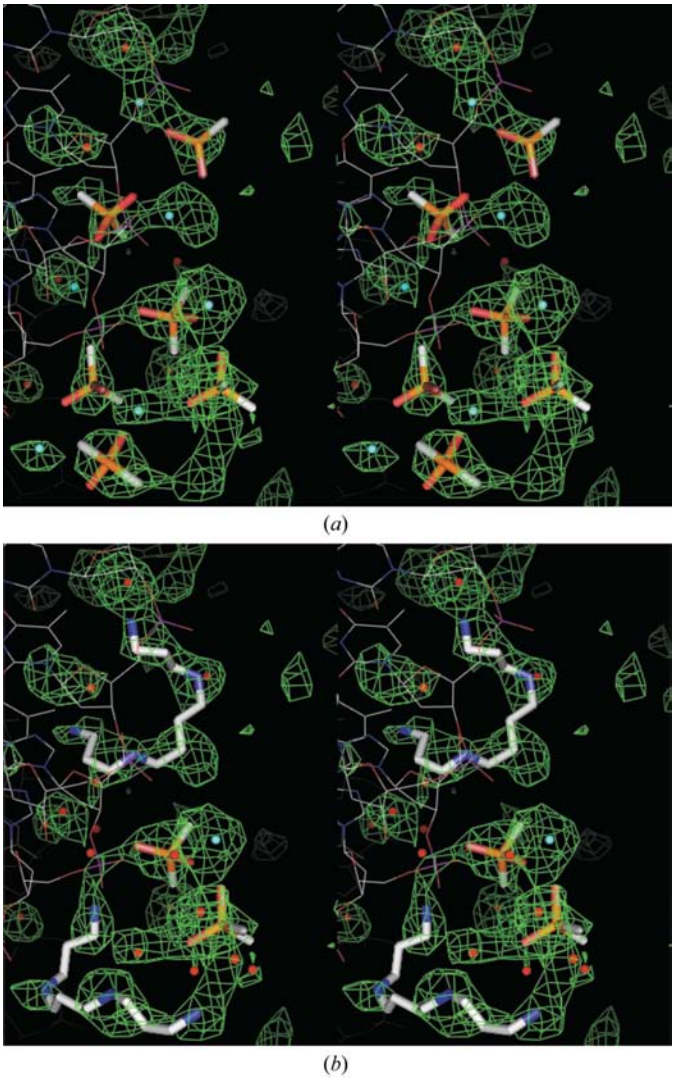

Figure 2
Stereoview of the $mF_o - DF_c$ electron-density map at the 2.5σ level in the neighbourhood of duplex *E* for (*a*) model *A* and (*b*) model *B*. The point of view is approximately perpendicular to the helical axes. The map was calculated prior to the introduction of ions and spermine. Cacodylates are shown as tetrahedra and Mn^{2+} ions as small cyan spheres.

−20°. With this assumption and using the duplex positions found with *MOLREP* we built a column of consecutive octamers which could successfully be refined.

Refinement was carried out with the program *REFMAC*5 (Murshudov *et al.*, 1997) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) and 5% of the unique reflections were set apart in order to calculate a free *R* factor as an independent cross-validation indicator of the progress of refinement (Brünger, 1992). The initial model was subjected to rigid-body refinement treating each duplex as a rigid body, followed by several cycles of maximum-likelihood isotropic restrained refinement which yielded an R_{work} of 0.355 and an R_{free} of 0.424. At this point, electron-density maps for DNA fitted very well into the model, as shown in Fig. 1, but significant extra density was visible in the $mF_o - DF_c$ map next to duplex *E* (Fig. 2). A search for ions was carried out in this region. The anomalous map obtained with the native data (λ = 0.9794 Å) did not show any clear cacodylate ions (As has a *K* edge at 1.0448 Å). For the Mn atoms, data were collected from a different crystal at a wavelength of 1.8929 Å where the Mn signal was maximal in a fluorescence spectrum measured from the crystal. The anomalous map of these data was calculated and only one significant peak was found, which we assigned to an Mn atom (Mn number 101 in the PDB file). Two cacodylate molecules were modelled next to this Mn, lowering R_{work} to 0.343 and R_{free} to 0.417. At this point two different models (*A* and *B*) were assumed in order to assign the rest of the extra density. In model *A* all density was assigned to Mn and cacodylate ions (Fig. 2*a*), while in model *B* spermine molecules were also introduced into the model, replacing some of the Mn and cacodylate ions (Fig. 2*b*).

Both refinements gave very similar R_{work} and R_{free} values after adding water molecules. We applied TLS refinement (Winn *et al.*, 2001). Final statistics are shown in Table 1. Data beyond 3.1 Å were scarce. When these were introduced into the final refinement no improvement in the map quality was detected. Model building was performed using the *TURBO-FRODO* program (Roussel *et al.*, 1998).

2.3. Calculations

The program 3*DNA* (Lu & Olson, 2003) was used in order to calculate the structural parameters of the individual oligonucleotide duplexes in form I. An overall axis for the six different duplexes which form this structure was also determined from the arrangement of the C_1' residues. Twist angles were calculated with respect to the overall axis from the C_1'–C_1' vectors. In this way twist angles are additive and the overall twist of one duplex with respect to its neighbour can easily be calculated. Interestingly, when the helical parameters are

calculated for each duplex with respect to its own helical axis, greater oscillations in value are found. For example, in the central ATAT alternating region the oscillation is greater than that reported below (Table 2), as is usually found in such sequences (Abrescia *et al.*, 1999). It appears that the calculations based on the C_1'–C_1' atoms smooth out the changes in twist.

The program *Cerius*2 (Accelrys Inc.) was used to prepare Figs. 3–6.

3. Results

3.1. Structure of form I

Form I presents an apparently complex organization of duplexes, with 22 oligonucleotides in the unit cell, as shown in Fig. 3. The individual molecules are organized into columns, which correspond to a helical arrangement of 11 duplexes in seven turns. As a result, the average rotation Ω (twist) of one duplex with respect to its neighbours in a column is $\Omega = 360^\circ \times 7/11 = 229.1^\circ$. One such column is labelled in Fig. 3, where it can easily be appreciated that each column has two dyad axes, one centred in oligonucleotide *F*, which is symmetric, and another one between two *A* duplexes which run in opposite directions. As a result of the presence of the dyad axes, the duplexes have different orientations and may be considered to be polar. Only duplex *F* is symmetric. The individual helical axes shown in Fig. 3 deviate by 3–4° from the overall axis of the whole column (*A*–*F*).

A partial view of the crystal structure is given in Fig. 4. The duplex columns are organized with approximate hexagonal symmetry. Each duplex is related to its neighbour in the next column by a vertical displacement of about 3.6 Å and an average rotation which can be calculated from the average angular twist Ω of oligonucleotides in a column. It turns out to be 163.3°, equivalent to 360° × 5/11. As a result of the combination of these two movements (rotation and vertical displacement), the relative orientation of equivalent structural features (for example, the minor groove) is 120° in neighbouring duplexes. The global consequence is that a given duplex (for example, blue in Fig. 4) will be surrounded by six duplexes: two in the same orientation (blue), two at +120° (green) and two at −120° (grey).

As indicated in §2, the interaction between the terminal base pairs of consecutive oligonucleotides in a column shows an unexpected negative twist ω_T (Table 2). A view of this

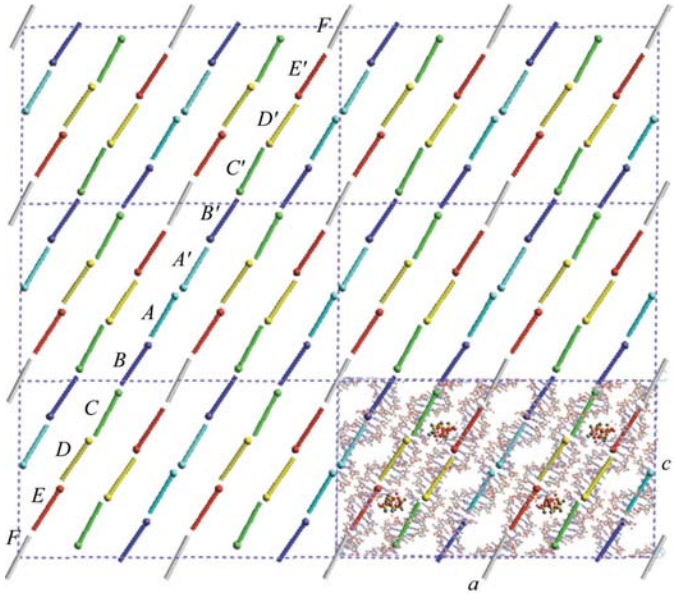

Figure 3
A view of six unit cells of form I projected onto the *xz* plane. In the bottom right corner the whole content of a unit cell has been represented. The letters correspond to the six different duplexes in the asymmetric unit, each in a different colour. The duplexes are represented by their own axis calculated using 3*DNA* (Lu & Olson, 2003). The polarity is indicated by a small sphere at one end. Each column of duplexes is formed by 11 different units, with a dyad between *A* and *A*′ (cyan) and another one centred on the *F* duplex. Neighbouring columns do not lie on the same plane, as is apparent in Fig. 4. The external region of electron density found next to duplex *E* (red) is represented by small spheres in the unit cell at the bottom right.

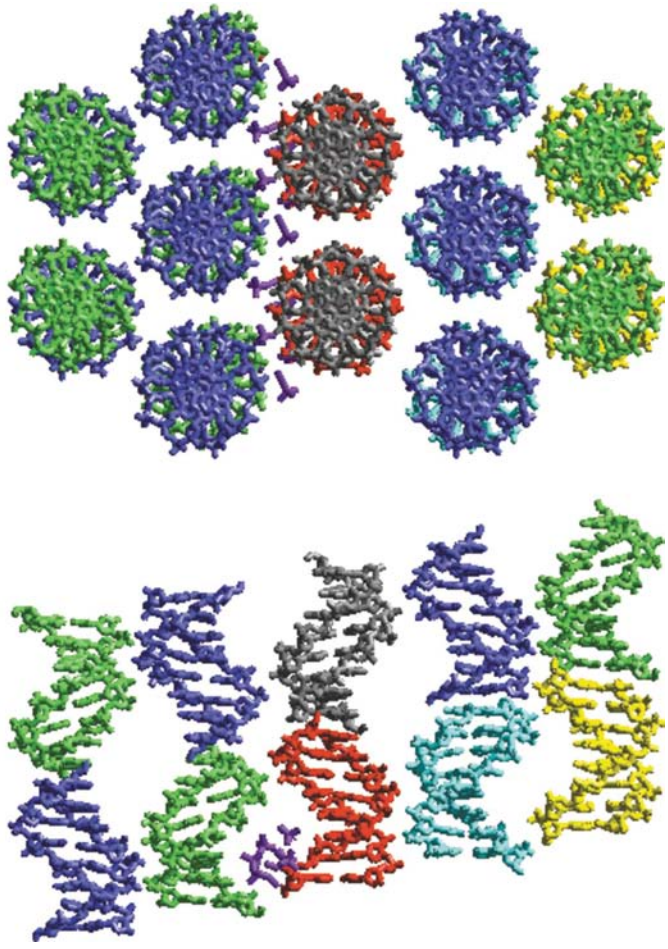

Figure 4
Representation of a segment of the crystal structure of form I centred on one corner of the unit cell. In the upper frame the duplexes are projected onto a plane perpendicular to the overall helical axis. The approximate hexagonal arrangement of neighbouring columns is clearly apparent. In the lower frame a perpendicular view is presented. Each duplex is displaced with respect to its neighbours by about 3.6 Å in the direction of the helical axis.

Table 3
Average conformational parameters of form I.

Step	ω (°)	Standard twist† (°)	Rise (Å)
AA/TT (end)	35.4	35.8	3.18
AA/TT	38.5	35.8	3.20
AT	34.0	33.4	3.36
TA	38.6	40.0	3.22
Average	36.3	35.7	3.24
TA (ω_T)‡	−25.6	—	3.50

† According to Gorin *et al.* (1995). ‡ Values for the step between the ends of two consecutive oligonucleotides.

interaction is shown in Fig. 5. It is mainly stabilized by stacking of the two terminal base pairs involved. Some sugar–sugar van der Waals contacts are also present. There are also hydrogen bonds between O5′ of adenine 1 and one of the O atoms of the phosphate of thymine 8 in the next duplex.

3.1.1. Conformational parameters. The helical parameters for each duplex are given in Table 2 and their average values are given in Table 3. Given the comparatively low resolution of our structure, we concentrate our analysis on the value of the twist, which appears to be less sensitive to refinement approximations. Nevertheless, we should point out that AT base pairs in general show the expected high propeller twist and in several cases form bifurcated hydrogen bonds with a base in the next base pair, as found in other structures (Nelson *et al.*, 1987; Coll *et al.*, 1987). There are several features which stand out from Tables 2 and 3.

(i) Each individual duplex has different values of twist for the various steps present, but the pattern is the same in each duplex, with an alternation in the magnitude of the twist. An exception is the twist for the first AA/TT step in duplex D, which is significantly larger than in the other cases.

(ii) The central alternating region shows the expected values of the twist, with ω_{TA} greater than ω_{AT}, as found in previously studied oligonucleotides (Fig. 5 in Abrescia *et al.*, 1999). However, in our case the alternation is less pronounced than in other cases. What appears to be unusual is that ω_{AA} is consistently larger in the second step than in the first step. As a result of this variation, all duplexes (except one end of *D*) present an alternation of twist. A search of the NDB for the AAAT sequence present in other structures shows that this is a general feature of this sequence. In two previously studied oligonucleotides (Edwards *et al.*, 1992; Huang *et al.*, 2005), the three twist values for this sequence practically coincide with those found by us (Table 3). What appears to be typical for the AAAT sequence is the alternation of twist values found in all cases.

(iii) The individual parameters of twist in Table 2 may contain errors owing to the comparatively low resolution of our study. On the other hand, the consistent values found in all duplexes indicate that the refinement process has been adequate, since the starting model had a twist of 36° for all steps. The values in Tables 2 and 3 agree with those previously reported (Edwards *et al.*, 1992; Huang *et al.*, 2005; Gorin *et al.*, 1995; Subirana & Faria, 1997).

(iv) The average value of twist for the whole octamers is 36.3° (Table 3), which corresponds to 9.9 base pairs per turn. This value deviates from mixed-sequence DNA, which tends to be underwound, but agrees with the report of Rhodes & Klug (1981), who found that the average twist of poly(dA)·poly(dT) is close to ten base pairs per turn.

3.1.2. The interhelical space. In both forms of d(AAATATTT) the crystal structure is highly hydrated. The oligonucleotide columns are organized at a comparatively large distance, in the range 24–25 Å. This fact is related to the large volume occupied by each base pair (1720–1750 Å^3 in both forms). In most oligonucleotide structures this volume is usually close to 1300 Å^3. The high hydration explains the rather low resolution of the diffraction patterns.

No direct interactions among the oligonucleotide columns are apparent in the crystal structure of form I. No ionic or water bridges were detected. All distances between phosphate O atoms of different duplexes are greater than 5 Å. Only in the neighbourhood of duplex *E* (red in Figs. 3 and 4) was a clear region of additional electron density detected. It is shown in detail in Fig. 2. It forms a continuous tube along the *b* axis of the crystal (Fig. 4), but does not show any contacts with other duplexes. The extra density could be approximated by either a catenate structure of cacodylate and Mn^{2+} ions or by spermine molecules. We have chosen the first possibility, since the electron density appears to be rather discontinuous and formed by approximately spherical regions. Catenate structures of cacodylate and divalent cations have been described (Ciunik & Glowiak, 1981). In any case, the description of this electron-density region is only tentative. The location of some Mn^{2+} and cacodylate ions could even be exchanged without any significant influence in the results of refinement. It should be also noted that this electron density only shows possible interactions with two phosphate groups in duplex *E*. We placed either spermine N atoms or Mn^{2+} ions in such sites.

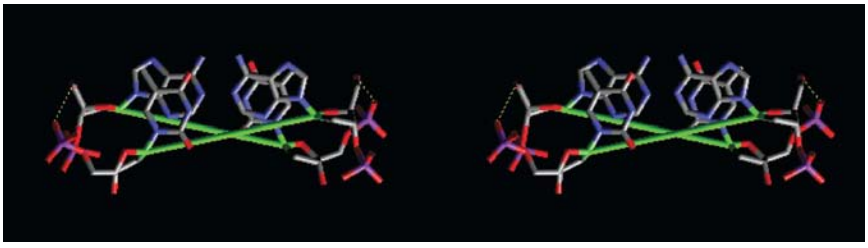

Figure 5
Stereoview of the stacking interaction between the terminal base pairs of oligonucleotides *B* and *C*. The interaction is practically identical between the other oligonucleotides that are in contact. It has also been found in other cases (Rozenberg *et al.*, 1998; Hizver *et al.*, 2001) where a pseudo-TA step is present. The green lines connect the C_1' atoms so that the negative twist is clearly apparent. Hydrogen bonds are shown as thin dashed lines in yellow.

3.2. Structure of form II

With rather small changes in the crystallization conditions, we detected another structure. In spite of the rather low resolution (about 6 Å) the pattern could be

indexed as a trigonal cell with $a = b = 43.0$, $c = 367$ Å. The value of c indicates that the crystal contains stacks of 14 duplexes. We assigned space group $P321$, which corresponds to an asymmetric unit of seven oligonucleotides, as shown in Fig. 6. The unit cell will thus contain 42 duplexes. Its structure can be interpreted as a helix of 14 duplexes in nine turns. The average turn per duplex will be $\Omega = 9 \times 360^\circ/14 = 231.4^\circ$ and the twist value between terminal steps -22.7°, values similar to those found in form I. A more detailed analysis of form II and comparison with other all-AT duplexes will be reported elsewhere.

4. Discussion

The structures we present in this paper are the first structures reported for a DNA duplex with all-AT Watson–Crick base pairs. Also, only one octamer structure of B-form DNA has previously been reported (Tereshko *et al.*, 1996). The two structures we have found form a similar helical arrangement of stacked oligonucleotides. It is not clear why form I is stabilized as the rather complex unit cell shown in Figs. 3 and 4.

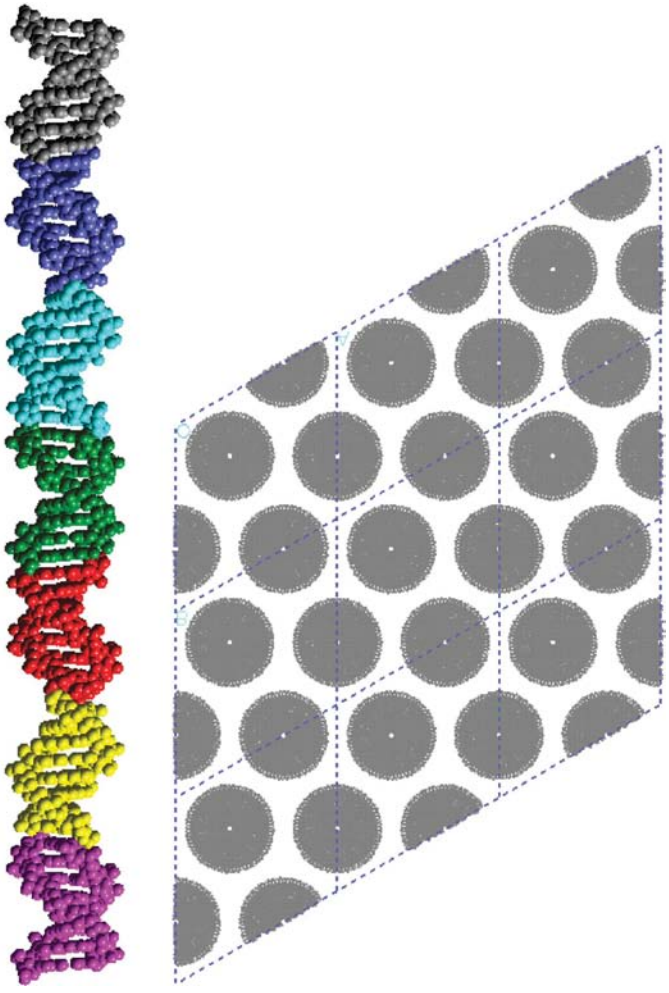

Figure 6
A model of the asymmetric unit (seven duplexes) of form II of d(AAATATTT). Each duplex is rotated 231.4° with respect to its neighbours in the column (see text). The projection on the *xy* plane of nine unit cells in the *P*321 space group is also shown.

The duplexes interact among themselves by end-to-end stacking through a pseudo-TA step (Fig. 5). The absence of phosphates apparently determines a negative twist of about -25° in this step. A similar value has been found (Rozenberg *et al.*, 1998; Hizver *et al.*, 2001) in several dodecamers which also have a pseudo-TA step between their terminal bases. It appears that this step does not like the standard 36° twist for the B form; rather, it turns to the left in the absence of phosphates.

On the other hand, no direct lateral interactions between the oligonucleotide columns are apparent. However, their relative azimuthal orientation appears to be strongly determined by electrostatic interactions between neighbouring DNA columns. Their orientation is similar in the two forms we have presented. In form I each column is surrounded by two columns with the same orientation (0° rotation), two columns with -120° rotation and another two with $+120^\circ$ rotation. In form II each column is surrounded by six columns which have alternate rotations of -120° and $+120^\circ$. A rotation of 120° is equivalent to a vertical displacement of one third the pitch of the DNA. It is interesting to note that in crystalline DNA fibres (Langridge *et al.*, 1960) the same relative molecular orientation is found: 0 and 120°. It appears that this selective orientation is determined by the electric field created by the DNA phosphates and their ionic atmosphere, since no apparent direct interactions between them have been seen. With these relative orientations (0 and 120°) the minor grooves of neighbouring duplexes never face each other, a fact which might be the reason for stabilizing such orientations.

The individual twist values of the different base steps in the sequence (Table 3) fall within the pattern of variation (Gorin *et al.*, 1995; Subirana & Faria, 1997) in oligonucleotides with related sequences. In particular they confirm the differences previously reported (Edwards *et al.*, 1992; Huang *et al.*, 2005). Thus, the difference in twist of the two AA/TT steps found by us (Table 3) appears to be an intrinsic feature of the AAAT sequence, rather than an end effect. Since the oligonucleotide structures described in this paper have no apparent lateral interactions and are also highly hydrated, the coincidence in twist values is of particular interest, since it confirms that the values previously reported for structures studied at higher resolution (Gorin *et al.*, 1995; Subirana & Faria, 1997) are basically a consequence of the base sequence and not of the influence of packing interactions in the crystal.

In summary, our work has shown a novel packing arrangement in crystals of B-form DNA duplexes, a helical array of stacked duplexes. We have found a similar organization in other all-AT oligonucleotides, as we will report elsewhere.

We thank the BM16 Spanish beamline staff of ESRF (Grenoble) for assistance in data collection. We are also thankful to Drs Isabel Usón and Lourdes Campos for help throughout this work. This work has been supported by grants

BIO2002-00317 from the Ministerio de Ciencia y Tecnología and 2001 SGR 00250 from the Generalitat de Catalunya. NV acknowledges a fellowship from the Generalitat de Catalunya.

References

Abrescia, N. G. A., González, C., Gouyette, C. & Subirana, J. A. (2004). *Biochemistry*, **43**, 4092–4100.

Abrescia, N. G. A., Malinina, L., Fernandez, L. G., Huynh-Dinh, T., Neidle, S. & Subirana, J. A. (1999). *Nucleic Acids Res.* **27**, 1593–1599.

Abrescia, N. G. A., Thompson, A., Huynh-Dinh, T. & Subirana, J. A. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 2806–2811.

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.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Campos, J. L., Urpí, L., Sanmartín, T., Gouyette, C. & Subirana, J. A. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 3663–3666.

Ciunik, Z. & Glowiak, T. (1981). *Acta Cryst.* B**37**, 693–695.

Coll, M., Frederick, C. A., Wang, A. & Rich, A. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 8385–8389.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.

Dickerson, R. E., Goodsell, D. S., Kopka, M. L. & Pjura, P. E. (1987). *J. Biomol. Struct. Dyn.* **5**, 557–579.

Edwards, K. J., Brown, D. G., Spink, N., Skelly, J. V. & Neidle, S. (1992). *J. Mol. Biol.* **226**, 1161–1173.

Gorin, A. A., Zhurkin, V. B. & Olson, W. K. (1995). *J. Mol. Biol.* **247**, 34–48.

Hizver, J., Rozenberg, H., Frolow, F., Rabinovich, D. & Shakked, Z. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 8490–8495.

Huang, D.-B., Phelps, C. B., Fusco, A. J. & Ghosh, G. (2005). *J. Mol. Biol.* **346**, 147–160.

Karaogly, H., Lee, C. M. Y. & Meyer, W. (2005). *Mol. Biol. Evol.* **22**, 639–649.

Lander, E. S. *et al.* (2001). *Nature (London)*, **409**, 860–921.

Langridge, R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960). *J. Mol. Biol.* **2**, 19–37.

Liebich, I., Bode, J., Reuter, I. & Wingender, E. (2002). *Nucleic Acids Res.* **30**, 3433–3442.

Lu, X.-J. & Olson, W. K. (2003). *Nucleic Acid Res.* **31**, 5108–5121.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Nelson, H., Finch, J., Luisi, B. & Klug, A. (1987). *Nature (London)*, **330**, 221–226.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rhodes, D. & Klug, A. (1981). *Nature (London)*, **292**, 378–380.

Roussel, A., Inisan, A. G., Knoops-Mouthuy, E. & Cambillau, E. (1998). *TURBO-FRODO* version OpenGL.1, University of Marseille.

Rozenberg, H., Rabinovich, D., Frolow, F., Hegde, R. S. & Shakked, Z. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 15194–15199.

Subirana, J. A. (2003). *Nature (London)*, **423**, 683.

Subirana, J. A. & Faria, T. (1997). *Biophys. J.* **73**, 333–338.

Tereshko, V., Urpí. L., Malinina, L., Huynh-Dinh, T. & Subirana, J. A. (1996). *Biochemistry*, **35**, 11589–11595.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Winn, M., Isupov, M. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.