Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Influence of packing interactions on the average conformation of B-DNA in crystalline structures

V. Tereshko† and J. A. Subirana*

Departament d'Enginyeria Química, ETSEIB, Av. Diagonal 647, 08028 Barcelona, Spain

† Present address: Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611-3008, USA.

Correspondence e-mail: subirana@eq.upc.es

Received 10 February 1998
Accepted 11 January 1999

The molecular interactions in crystals of oligonucleotides in the B form have been analysed and in particular the end-to-end interactions. Phosphate–phosphate interactions in dodecamers are also reviewed. A strong influence of packing constraints on the average conformation of the double helix is found. There is a strong relationship between the space group, the end-to-end interactions and the average conformation of DNA. Dodecamers must have a B-form average conformation with 10 ± 0.1 base pairs per turn in order to crystallize in the $P2_12_12_1$ and related space groups usually found. Decamers show a wider range of conformational variation, with 9.7–10.6 base pairs per turn, depending on the terminal sequence and the space group. The influence of the space group in decamers is quite striking and remains unexplained. Only small variations are allowed in each case. Thus, crystal packing is strongly related to the average DNA conformation in the crystals and deviations from the average are rather limited. The constraints imposed by the crystal lattice explain why the average twist of the DNA in solution (10.6 base pairs per turn) is seldom found in oligonucleotides crystallized in the B form.

1. Introduction

The first dodecamer d(CGCGAATTCGCG) to be crystallized in the B form (Drew *et al.*, 1981) has been thoroughly analyzed. In particular, its packing arrangement has been studied in detail (Dickerson *et al.*, 1987). Since then, many related dodecamers have been crystallized in the $P2_12_12_1$ and related space groups. Their structures have been stored in the Nucleic Acid Database (NDB; Berman *et al.*, 1992). It seems timely to review the packing arrangement of all these related dodecamers in order to determine how packing and oligonucleotide structural features are related. Selected examples in the different space groups will be analyzed in detail. Our study shows that the main features are maintained in all cases, as previously described for two structures by Dickerson *et al.* (1987), although differences in detail are found. In a few cases, another packing scheme in the $R3$ space group has been found for dodecamers (Timsit *et al.*, 1989), but no general conclusions can be drawn given the limited number of structures available. Packing in decamers will also be analyzed. Decamers crystallize with a different organization, stacking as continuous columns (Heinemann & Alings, 1989), although recently complex end-to-end interactions have been found in one decamer (Spink *et al.*, 1995).

The analysis of oligonucleotide structures by single-crystal X-ray diffraction has been the subject of controversy, since it is not clear to what extent sequence and 'packing forces' determine in each case the conformational parameters in a

Table 1
Crystallographic and conformational parameters of oligonucleotides.

Oligonucleotides are coded by a shortened version of their reference number in the Nucleic Acid Database (Berman *et al.*, 1992). Average parameters, helical axes and μ have been calculated with *NEWHELIX*93. μ is the angle between helical axes, as shown in Fig. 6. In the dodecamers it coincides with twice the angle between the helical axis and the longest side of the unit cell. The internal symmetry r.m.s. compares half an oligonucleotide with the other half, using all the atoms (whole) or only the central part (eight nucleotides in dodecamers; four nucleotides in octamer). The Dickerson interaction (DI) between the four terminal C·G base pairs is compared with the same structure in B1 by using all the atoms of the four base pairs or only the four guanines (in parentheses) directly involved in the interaction, as shown in Fig. 1. τ corresponds to the angle of rotation described in Fig. 6. It has been calculated with the program *CURVES* as described in the text. In B42 there is a different type of DI at either end of the dodecamer. References: B1, Drew *et al.* (1981); B4, Fratini *et al.* (1982); B42, Leonard & Hunter (1993); B47, DiGabriele & Steitz (1993); B70/71, Urpí *et al.* (1996).

Code	Sequence	Space group	Molecules in asymmetric unit	Average twist	Average rise (Å)	Volume per base pair (Å^3)	Axis angle (μ; °)	Internal symmetry r.m.s. (Å)		DI (r.m.s.; Å)	τ (°)
								Whole	Centre		
B1	CGCGAATTCGCG	$P2_12_12_1$	1	35.9	3.35	1385	33.5	1.77	0.76	—	182
B4	CGCGAATT^{Br}CGCG	$P2_12_12_1$	1	36.0	3.37	1293	42.6	0.80	0.47	0.64 (0.23) 0.99 (0.59)	179
B42	CGTAGATCTACG	$C2$	1	35.9	3.42	1210	24.9	1.65	1.23	0.71 (0.49)	185
B47	CGCGAAAAAACG	$P2_12_12$	(2)	36.1	3.32	1323	29.6	—	—	1.16 (0.67)	180
B70(1)	CGCTCTAGAGCG	$P2_1$	2	35.8	3.55	1316	20.2	1.38	0.77	1.20 (0.62)	180
B70(2)	CGCTCTAGAGCG	$P2_1$	2	36.0	3.52	1316	18.3	0.86	0.78	1.15 (0.64)	181
B71	CGCTAGCG	$P2_12_12_1$	3	36.6	3.49	1242	40.6	0.55	0.30	0.84 (0.17)	183

given crystal structure. Furthermore, the average structure found in solution differs from that found in crystals (Niederweis *et al.*, 1992). Statistical analysis of the conformational parameters has allowed the determination of the average local influence of sequence, although sufficient data are not available for a complete analysis (Gorin *et al.*, 1995; Subirana *et al.*, 1995; Subirana & Faria, 1997; El Hassan & Calladine, 1997). The general conclusion is that each base step has different conformational features, which only in some cases appear to be influenced by flanking sequences (Subirana & Faria, 1997). On the other hand, the influence of 'packing forces' remains elusive. In fact 'packing forces' are nothing other than intermolecular interactions (ionic, hydrogen bonds, van der Waals), which are of the same nature as those involved in DNA–protein interactions. Their influence appears most clearly in those base steps which are considered to be flexible by the various studies just mentioned. For example, in the case of the CA step, which is very variable, a low value of twist ($31.3 \pm 2.8^\circ$) is found in most dodecamers, while a high value ($45.2 \pm 6.1^\circ$) is found in most decamers. Unfortunately, the neighbouring sequences are also different (Subirana & Faria, 1997), and it is not possible to decide unambiguously whether this effect is either a consequence of the overall sequence or to the different packing environment (or both). In any case, DNA has a considerable range of conformational variability while preserving the standard Watson–Crick base-pairing scheme. This is most evident in protein–DNA complexes in which the DNA helix is considerably distorted (for example in the Pvu II endonuclease complex; Cheng *et al.*, 1994).

Packing forces in a B-form oligonucleotide crystal may be divided in two groups: end-to-end interactions and lateral interactions with neighbour molecules. The end-to-end interactions play a very important role in crystals. In all oligonucleotides crystallized in the B form, the molecules form infinitely long columns stabilized by end-to-end interactions. The crystal is formed by further interaction of these columns with neighbouring columns, most often in parallel arrays. The helical axis of the molecules is, in general, approximately oriented along a crystallographic axis. The combination of these different interactions determines the space group in which the molecules crystallize.

In this paper, we first analyze the packing in B-form dodecamers, taking into account end-to-end and electrostatic phosphate–phosphate interactions, both of which are influenced by the presence of counterions. These packing constraints demonstrate a strong relationship with the average parameters of the double helix. Only a few dodecamer structures are analyzed, since most are very similar. In a second part of the paper, we describe some features of packing in B-DNA decamer crystals. Packing in decamers shows a strong relation to the average parameters of the double helix. Our analysis complements previous detailed comparisons of crystalline structures (Heinemann *et al.*, 1994; Dickerson, 1998).

2. Packing of dodecamers in different space groups

Dodecamers show a peculiar end-to-end interaction between the guanines of the two terminal base pairs. This interaction requires that the oligonucleotides have terminal C·G base pairs. We will call it the Dickerson interaction (DI), since it was found in the first dodecamer which was crystallized in the laboratory of Professor Dickerson (Wing *et al.*, 1980). This interaction is essential for the organization of the dodecamers in the $P2_12_12_1$ space group. It is also present in other dodecamers which crystallize in related space groups, and in one octamer. The latter case is shown as an example in Fig. 1. There are four hydrogen bonds between N2 and N3 of neighbouring guanine bases which stabilize the DI. As a result, the geometry of the four bases is rather rigid, as indicated by the comparative r.m.s. values given in Table 1, with a maximum value of 0.67 Å for B47. Nevertheless, in some cases (BDL047, BDL070) one of the terminal cytosines is disordered and does not form a Watson–Crick pair with the

corresponding guanine. There is also a close contact between the O2 atoms of cytosines 3 and 15 and the C2′ atom of the terminal sugars. This feature of the DI may explain why no dodecamer with a starting sequence CG*R* (*R* = purine) has been yet crystallized, since the DI interaction may not be possible when a bulkier purine is present in the third position instead of the usual pyrimidine.

The DI has been analyzed in detail by Dickerson *et al.* (1987) in two dodecamers which crystallized in the $P2_12_12_1$ space group. These two dodecamers are included in Table 1 and we will take the first one (B1) as a reference for comparison. In this packing arrangement the asymmetric unit is a single dodecamer. The dodecamers pack in infinite columns, in which individual molecules are related by a screw axis, as shown in Fig. 2. These infinite columns are packed side by side and there are phosphate–phosphate interactions between them which will be analyzed in more detail below.

The B42 dodecamer studied by Leonard & Hunter (1993) is packed in a similar way, as shown in Fig. 2. However, in this case the DI has a dyad axis which relates the four interacting C·G base pairs. This dyad axis is approximately horizontal in the plane of the paper through the centre of Fig. 1. As a result, the DIs at both ends of the molecule are different and each one has a dyad axis.

In the dodecamer studied in our laboratory (Urpí *et al.*, 1996; Tereshko *et al.*, 1996), two molecules are present in the asymmetric unit: B70(1) and (2). Each of them again forms infinite columns similar to those found in B1 and B4, but in this case neighbouring columns are formed by dodecamers with a different conformation. Each column shows an additional interaction with its neighbours, in which the cytosines of terminal base pairs are stacked.

The B47 dodecamer studied by DiGabriele & Steitz (1993) also has two molecules in the asymmetric unit. One layer of molecules is ordered, but the next layer is partially disordered as described in the original article. Because of this, this structure has not been included in Fig. 2. The molecules also pack in infinite columns in which individual dodecamers are related by a screw axis. Cytosine–cytosine interactions between terminal base pairs are also present. In this case a dyad axis is found at the centre of the two stacked cytosines.

The molecules have a different inclination in the unit cell, as is clearly apparent in Fig. 1. The angle μ between the average helical axes of consecutive molecules which results from such an inclination is given in Table 1. The longest dimension of the unit cell is related to this angle of inclination, although for a quantitative relationship it is necessary to also take into account the differences in average rise (given in Table 1) and the eventual inclination of the DI with respect to the cell axis.

Projections of the structures we have described are presented in Fig. 3. Each dodecamer is surrounded by six neighbouring dodecamers, but only in the B1 structure is an approximate hexagonal packing found; in all other cases the distance between neighbouring molecules varies considerably depending on the direction. Furthermore, the rows of neighbouring molecules are oriented differently with respect to the unit cell, as is clearly apparent in Fig. 3.

The observations presented in this section demonstrate that even though all the molecules discussed have a similar packing motif based on the DI (shown in Figs. 1 and 4), the lateral interactions between neighbouring molecules in the crystal vary in each case.

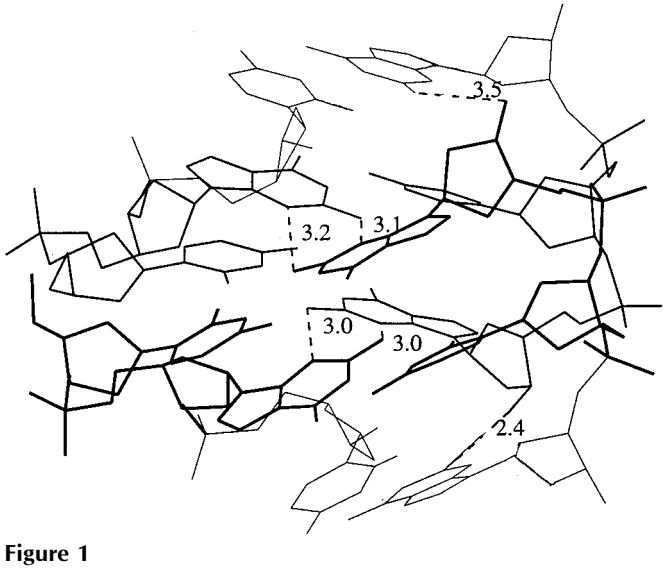


Figure 1
An example of the Dickerson interaction (DI) between terminal base pairs of two oligonucleotide molecules. In the figure, the ends of two octamer molecules (B71) are shown as an example. All the dodecamer structures analyzed in this paper have a similar geometry. The two terminal base pairs of one molecule are represented in heavy lines. They interact with the corresponding base pairs of a neighbouring molecule, shown as lines of medium thickness. The third base pairs of each molecule are shown as thin lines. Hydrogen bonds between N2 and N3 of guanines are shown as dashed lines with distances in Å. Hydrogen bonds between the O3′ terminal O atoms and N2 of the third guanine are also indicated. The N2–N3 distances are somewhat shorter in most of the other structures, being in the range 2.6–3.2 Å. An approximate local dyad axis runs horizontally through the centre of the figure.

3. General packing requirements

The dodecamer columns stabilized by the DI show different types of organization, which can be related to the variable distortion of the different DNA sequences and may also be influenced by the crystallization conditions and the counterions present. There is only one report of the same dodecamer being crystallized in two different space groups (Hunter *et al.*, 1988), but unfortunately only one of the crystal forms has been solved (Leonard & Hunter, 1993).

The variations in packing of the different structures can be summarized as follows.

(i) The angle μ between the helical axes of consecutive molecules in each column, given in Table 1 and obvious from Fig. 2, varies in different dodecamer structures; for example, B4 shows one of the largest angles, about twice that found between B70 molecules.

(ii) Neighbouring dodecamer columns have different relative displacements in the vertical direction, as is also obvious from Fig. 2. As a consequence, the phosphate–phosphate contacts vary in each case, as discussed below.

(iii) The distortions of molecules are not symmetric. The degree of internal symmetry may be estimated from the r.m.s. values given in Table 1. In general, the central part shows a higher symmetry than the ends of the molecules subjected to the DI. Some dodecamers, such as B1, have a low degree of internal symmetry, mainly because of different patterns of bending at either end of the molecule (Dickerson *et al.*, 1994). In the B70 structures the central parts are rather symmetric, but one of the molecules is much more bent than the other (Tereshko *et al.*, 1996).

The distortions of each structure are also apparent from the parameters given in Table 2. The slide value is a local feature of the two terminal base pairs and indicates how far the C·G base step at either end deviates from the average molecular helical axis. In the cases (B47, B70) where there is a cytosine–cytosine interaction between neighbouring dodecamers, the distortion is larger. The X and Y displacements from the overall helical axis are also quite variable.

4. Phosphate–phosphate interactions in dodecamer crystals

The oligonucleotide columns pack side by side with short phosphate–phosphate distances (given in Table 3) along crystallographic planes which coincide with a diagonal of the unit cell, as is apparent in Fig. 3. The phosphate–phosphate interactions occur in the plane of the drawing in Fig. 2. It is characteristic of the packing arrangement of dodecamers that short phosphate–phosphate distances are only present along such planes; phosphate–phosphate interactions are much weaker between planes.

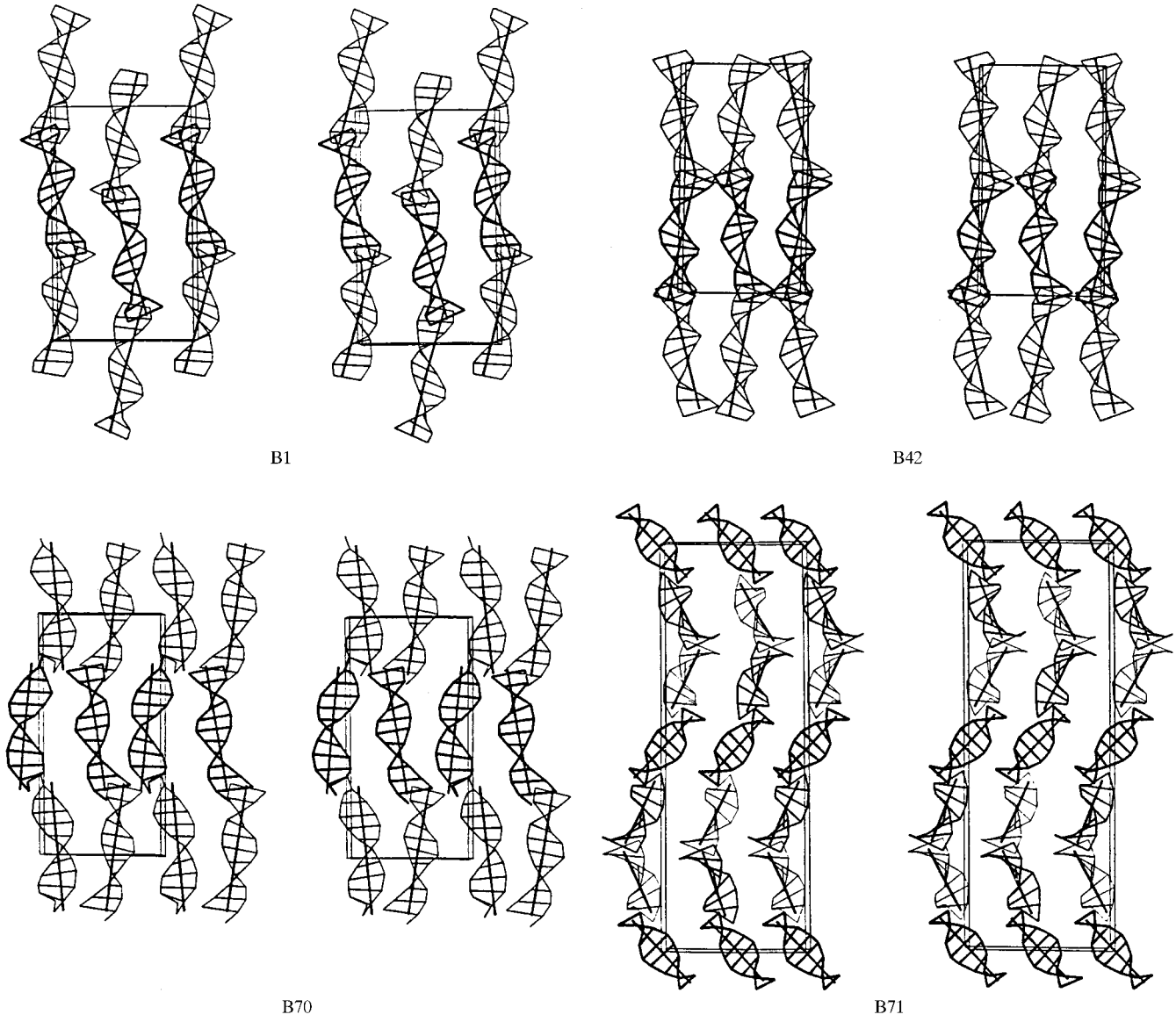


Figure 2
Stereopairs of simplified views of molecular packing in the different cases. Only the helical axes and the C1′ atoms are represented with virtual bonds between them. In the B70 case, a terminal cytosine is disordered and is not shown in the figure.

Phosphate–phosphate short distances are given in Table 3 and are schematically shown in Fig. 5. A short phosphate–phosphate distance is most likely to indicate that mobile positive counterions are present in the neighbourhood, but at the resolution available in most cases such ions have not been localized. In the high-resolution structure of CGCGAATTCGCG (Shui *et al.*, 1998) one magnesium ion is found in this region. Water 77 in the BDL001 coordinate file corresponds to this magnesium ion, which was not detected in the original study of Drew *et al.* (1981). Furthermore, it is likely that ions are very mobile in the crystal. Only in two cases, indicated in Table 3, have ions been assigned. However, in all the dodecamers which have been crystallized, divalent cations and spermine are present in the crystallization medium and are likely to contribute to stabilizing the crystal lattice, in particular at the points where short phosphate–phosphate distances are found.

Inspection of Table 3 indicates that dodecamers in the usual $P2_12_12_1$ unit cell have the lowest number of phosphate–phosphate interactions. Some dodecamers which pack in the $P2_12_12_1$ space group show a different rotational orientation in the unit cell and the phosphate–phosphate interactions will be different in such cases, although the overall organization of the oligonucleotides is similar. These structures have been reviewed by Vojtechovsky *et al.* (1995) and will not be discussed here. The B42 and B70 structures have many short phosphate–phosphate distances, some of which are associated with the presence of a divalent cation as reported in Table 3. The B42 structure in particular is very compact, as indicated by the low volume per base pair of 1210 Å^3, about 13% less than the equivalent volume in the reference B1 structure (Table 1). It is worth noting that in crystals the close P–P contacts are always found between infinite columns of duplexes with parallel packing. Such contacts usually indicate that important lateral interactions which stabilize crystal packing are present. The role of these contacts in the distortion of DNA fragments in the crystalline state is still unclear. For example, the initial explanation of bending in the B1 structure as a consequence of the presence of a short P2–P7 contact in the crystal should be reviewed, since a hydrated Mg^{2+} ion is found in the major groove of the high-resolution structure BDL084. On the other hand, the orientation and vertical displacement of oligonucleotide columns in the same space group $P2_12_12_1$ is not predetermined and may change without any significant influence on the cell parameters, as found by Vojtechovsky *et al.* (1995).

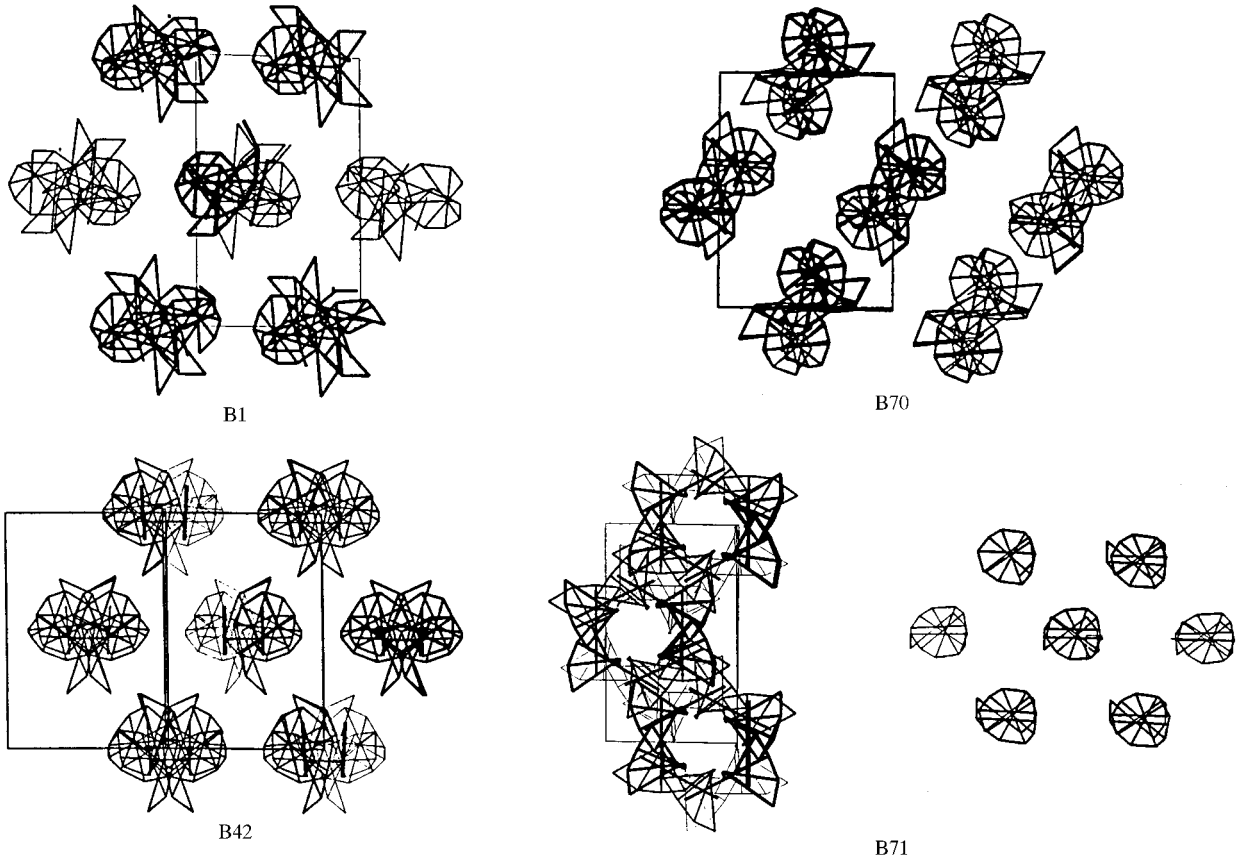


Figure 3
Projection of oligonucleotide columns. The projections are along the major *c* axis of the unit cell, except in B42 in which the projection is made onto a plane perpendicular to the *ac* diagonal. An octamer molecule (B71) and its nearest neighbours are also shown, projected onto the *ab* plane (left) and also onto a plane perpendicular to the molecular helix axis (right), showing the packing similarity with the dodecamers. The main phosphate–phosphate interactions occur between molecules along both diagonals in B1 and along the diagonal from top left to bottom right ion B42 and B70. In the latter case, the cytosine–cytosine interactions occur along the other diagonal.

Table 2
Displacement from the helical axis of the terminal base pairs.

The values given were calculated with *NEWHELIX*93 using the C6–C8 vector for each base pair. The two values given in each case correspond to both ends of each molecule. In the B70 structures cytosine 1 was disordered (Urpí *et al.*, 1996) and the corresponding value could not be calculated.

	Slide (Å)	Δx (Å)	Δy (Å)
B1	0.3/0.2	0.4/2.7	−0.1/1.6
B4	0.3/0.0	0.7/1.5	−0.7/1.3
B42	0.3/0.3	1.2/2.3	0.2/0.7
B47	2.0/−0.8	−3.9/−0.2	−1.3/−0.6
B70(1)	nd/0.9	nd/0.5	nd/−1.0
B70(2)	nd/1.0	nd/0.1	nd/−1.7
B71	0.2/0.7	0.8/1.8	1.1/−0.6
Ideal B	0.0	0.2	0.0

5. Influence of the Dickerson interaction on the average twist of DNA

The two terminal base pairs, related by the DI shown in Fig. 1, can be approximately superimposed by rotating the molecule by an angle τ, as indicated in Fig. 6. The angle τ can be determined with the program *CURVES* (Lavery & Sklenar, 1988). If we consider the column of dodecamers indicated in Fig. 6 as a continuous helix, we can determine a local helical axis and find the angle of twist between either base-pair steps G11·G14 and C1′·G′24 or base-pair steps G12·C13 and G2′·C′23. Both values of twist are rather similar and their average value τ is given in Table 1. It is always very close to 180°.

Since the Dickerson interaction is rather rigid, owing to the geometry of the guanine–guanine interactions shown in Fig. 1 which result in an almost constant value of τ, the average twist in dodecamers will also be rather constant. This is clearly apparent from Table 1: the maximum deviation from 36° is 0.2°. We reach the important conclusion that the average number of base pairs per helical turn in dodecamers crystallized with the Dickerson interaction and related by a screw or dyad axis will always be close to ten. All values shown in Table 1 fall in the range 10 ± 0.1. Dodecamers with a tendency to different average values of twist will not be able to crystallize in unit cells such as the ones analyzed in this paper. For example, it will not be possible to crystallize a dodecamer in the C form (nine base pairs per turn) in the unit cells found in dodecamers.

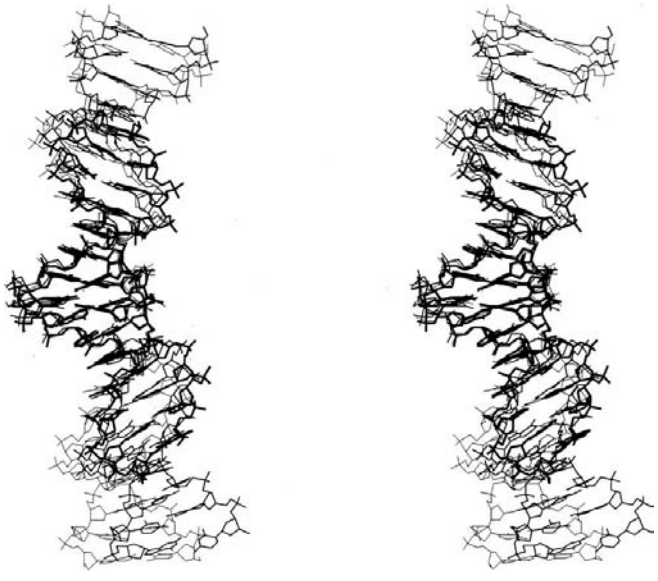


Figure 4
Stereopair showing the superposition of the B1 (thin lines), B4 (intermediate lines) dodecamers and the B71 octamer (heavy lines), taking as a reference the four guanines of the central DI.

Table 3
Close intermolecular P–P contacts between neighbouring molecules.

Phosphate–phosphate distances of 7.0 Å or less are included in the table. The distances indicated occur between phosphates in one oligonucleotide column in the crystal with those in another column. Thus, in the B70 dodecamer structure all short distances occur between the two different molecules in the asymmetric unit. In those cases in which an ion is involved in the interaction it is also indicated. Short phosphate–phosphate distances which occur in the DI are not included in the table. The values shown for the octamer are the average of those found in the three different molecules in the asymmetric unit.

Structure	P*i*–P*j*	Distance (Å)
B1	2–7	6.4
	10–18	6.6
B4	2–7	6.6
	10–19	6.8
B42	2–9 (Mg^{2+})	6.6
	3–15	6.9
	6–20	5.5
	7–20	5.3
	7–19	5.6
	10–11	6.2
	10–10	6.2
B70(1/2)	2–12 (Ca^{2+})	5.1
	2–24 (Ca^{2+})	6.2
	2–11	6.5
	4–9	6.9
	14–12 (Ca^{2+})	5.1
	15–12	6.7
	18–20	5.5
	18–19	5.6
	19–20	5.7
	20–10	6.6
	22–23	7.0
B71	2–7	7.0
	5–8	6.7
	5–10	5.9
	6–10	6.8
	8–16	6.4
	13–16	6.0

6. The octamer case

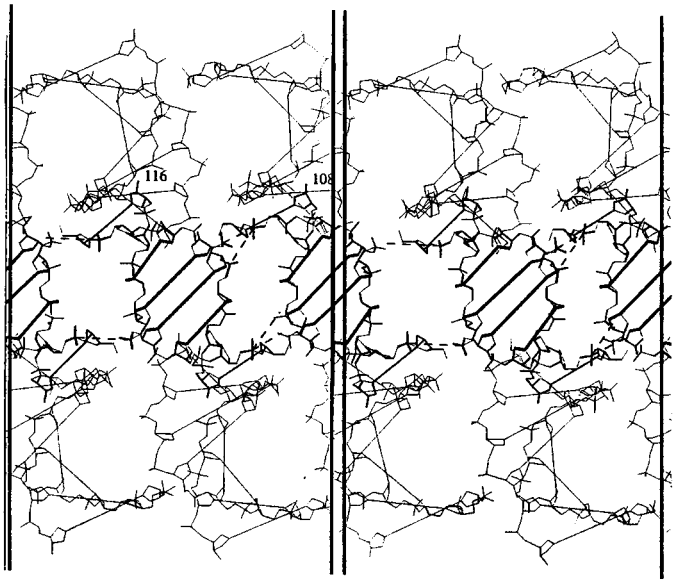
We have seen that in dodecamers the presence of the DI creates a strong relationship between the packing arrangement and the parameters of the double helix. It is interesting to note that the DI has never been found in decamers, although several crystals have sequences with a terminal CG sequence, as shown in Table 4. On the other hand, we have been able to crystallize an octamer (B71) with the sequence CGCTAGCG which also shows the DI (Urpí *et al.*, 1996; Tereshko *et al.*, 1996). The octamers also pack in continuous

columns, but they are considerably distorted, following a helical path as shown in Figs. 2 and 3. Three similar octamers are found in the asymmetric unit. The end-to-end interactions between consecutive molecules are also of the DI type. The whole octamer and its DI are very symmetric in this case, as is apparent in Fig. 1 and Table 1 (internal symmetry column). Packing on projection is quite different to the dodecamers, since consecutive molecules are related by a pseudohexagonal screw axis. However, when a view is obtained by projection onto a plane perpendicular to the helical axis of individual molecules, the mutual relationship of neighbouring molecules is similar to that found in the dodecamers, as seen in Fig. 3.

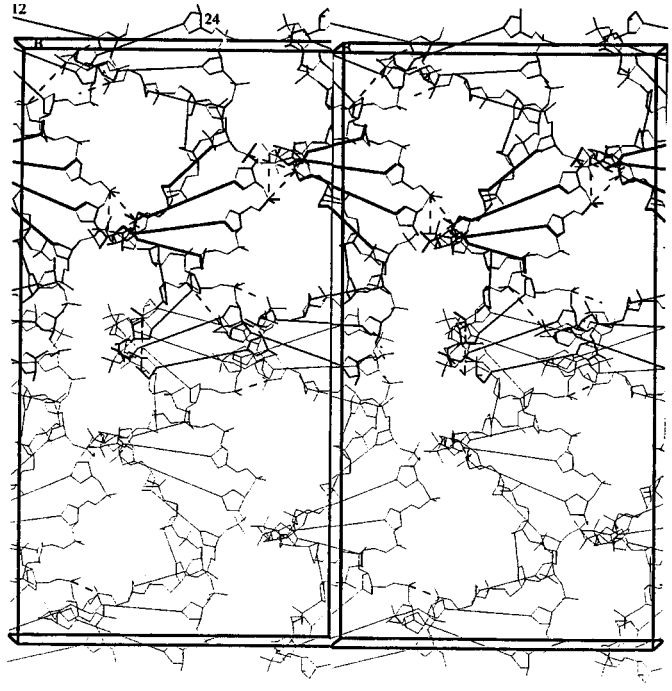
It is striking that the DI interaction is almost identical in octamer and the B1 dodecamer; the r.m.s. of the four guanine bases is only 0.17 Å. A superposition of the octamer and the B1 and B4 dodecamers is shown in Fig. 4. It can be seen that the four guanines which take part in the DI are practically superimposed, although the phosphodiester backbone is slightly different. The rest of the molecule deviates significantly, since the sequence and bending of each molecule is quite different (Fig. 4). The angle τ (defined in Fig. 6) can also be determined with the program *CURVES* in the octamer case. The value obtained for τ is practically identical to that found in dodecamers (Table 1), as would be expected (Fig. 4).

7. Packing in decamers

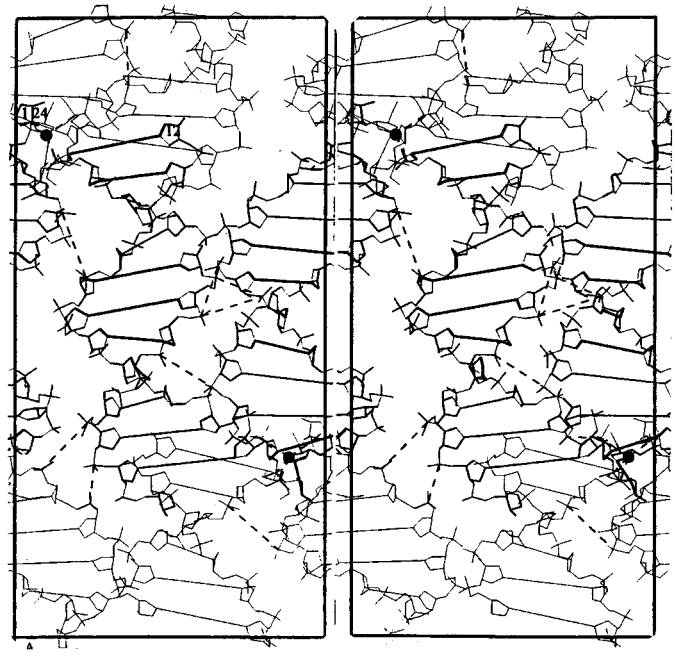
In all decamers (with one exception; Spink *et al.*, 1995) the terminal base pairs are stacked and long columns of a pseudocontinuous double helix are found in the crystal. One of the sides of the crystal coincides with the molecular length (about 34 Å). As a result, the average twist of the base-pair steps (including the step between contiguous decamers) coincides with the canonical 36° of B-DNA. Thus, the average



B71



B42



B70

Figure 5
Stereopairs showing the main phosphate–phosphate interactions present in some of the structures described in this paper. The same method of representation is used as in Fig. 3 in Dickerson *et al.* (1987), where B1 and B4 are compared. Most of the interactions occur approximately in the plane of the paper, except in B42 in which some interactions are perpendicular to the plane of the paper. The black dots in B70 correspond to the Ca^{2+} ions. The edges of the unit cells are indicated by heavy lines. In B71 only half of the unit cell in the *c* direction is shown.

Table 4
Structural parameters of decamers crystallized in the B form.

Class	Space group	NDB code	Sequence	$\langle\omega\rangle$ (°)	ω_T (°)	Decamer length (Å)
1	$C2$	BDJ017	CCAGGCCTGG	35.5	40.1	34.3
1	$C2$	BDJ019	CCAACGTTGG	35.5	40.6	34.3
1	$P6$	BDJ027	CCAGGC^{me}CTGG	35.5	40.6	34.8
1	$C2$	BDJ044	CCAACITTGG	35.7	38.4	34.2
1	$P2_12_12_1$	BDJ051	CATGGCCATG	35.5	40.0	34.7
1	$P6$	BDJ052	CCAAGCTTGG	35.4	41.3	34.3
1	$C2$	BDJ060	CTCTCGAGAG	35.5	40.4	33.4
1	$R3$	BDJ039	CCGGCGCCGG	36.0	35.7	34.6
2	$P2_12_12_1$	BDJ025	CGATCGATCG	37.1	25.8	33.3
2	$P2_12_12_1$	BDJ031	CGATTAATCG	37.2	25.3	33.1
2	$P2_12_12_1$	BDJ036	CGATATATCG	36.7	29.9	33.7
2	$P2_12_12_1$	BDJ037	CGATATATCG	37.0	26.8	33.6
3	$P3_221$	BDJB43	CCAACITTGG	34.4	50.8	33.2
3	$P3_221$	BDJ055	CCATTAATGG	34.2	52.6	33.2
3	$P3_221$	BDJB48	CGATCG^{me}ATCG	33.9	55.0	33.4

twist $\langle\omega\rangle$ of the nine base steps in the decamer will be related to the twist angle ω_T of the step between contiguous decamers by the formula

$$9\langle\omega\rangle + \omega_T = 360^\circ. \quad (1)$$

The values calculated for $\langle\omega\rangle$ with *NEWHELIX* and for ω_T using expression (1) are given in Table 4. The pitch of the pseudocontinuous helix, which coincides with one of the dimensions of the crystal, is also given in Table 4. Inspection of the table reveals that there are three structural classes of decamers, which show some striking structural features.

(i) All decamers which have been crystallized in the B-form have a sequence which starts with C.

(ii) Although there are only a few examples, when the second base is C, A or T the decamers show a similar behaviour, with a relatively long decamer length and an ω_T of about 40° (class 1).

(iii) On the other hand, when the decamer starts with the CG sequence $\langle\omega\rangle$ is larger and accordingly ω_T decreases (class 2). It is clear that the terminal base step has an influence on the average parameters found in the decamers.

(iv) An even more striking effect is related to the space group of crystallization. In $P3_221$, ω_T shows a very high value, 50–55° (class 3). The same sequences, crystallized in a different space group in which the helical axes are parallel, have different twist angles, as can be seen in Table 4 by comparing the data of BDJ043 and BDJ044, or BDJ025 and BDJ048. A detailed comparison of these four structures was presented in the original papers (Baikalov *et al.*, 1993; Lipanov *et al.*, 1993).

(v) The length of the decamer is less than 34 Å in the sequences which start with CG or are crystallized in $P3_221$. It is greater than 34 Å in the other cases (with the exception of BDJ060). Thus, there is no correlation between the twist parameters and the length of the decamer.

In summary, the features just discussed indicate that there is a strong influence of terminal sequence and packing conditions on the average helical parameters of decamers. The general features of the three different classes are summarized in Table 5. It is interesting to note that stacking between consecutive dodecamers corresponds to a G·C step and gives a very variable value of twist ω_T, whereas this step inside an oligonucleotide sequence is one of the least variable, with $\omega = 37.5 \pm 2.7^\circ$ (Subirana & Faria, 1997).

As an example of the different features of each decamer class, it is interesting to compare BDJ031 with BDJ055, which have the same sequence in the central hexamer. What is clear is that the stacking features of the terminal base pairs are quite variable (ω_T varies from 25.3 to 52.6°) and do not appear to be a determining factor in the crystallization of the oligonucleotide. The different crystallization conditions in both cases apparently favour a molecular conformation which determines packing in a different space group, with different intermolecular interactions. In one case (BDJ031) the helical axes of the decamers are parallel, whereas in the other case (BDJ055) they cross each other and a mutual interpenetration of crossed grooves takes place.

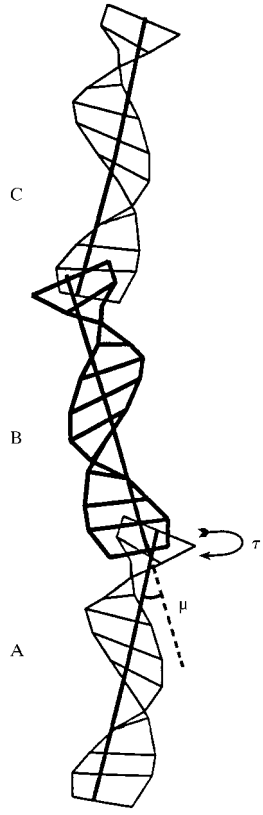


Figure 6
Three consecutive B71 dodecanucleotides (*A*, *B* and *C*), as crystallized in the $P2_12_12_1$ space group. The scheme indicates the angles μ and τ required to superimpose the terminal base pairs of the *A* and *B* dodecamers, as described in the text.

Table 5
Classification of decamers crystallized in the B form.

N corresponds to any base, except G. S represents either C or G.

Class	Space group	Sequence	Base pairs per turn	ω_T (°)	Decamer length (Å)
1	Several	CN . . . NG	10.1	36–41	34.4
2	$P2_12_12_1$	CG . . . CG	9.7	25–30	33.4
3	$P3_221$	CS . . . SG	10.5	51–55	33.3

8. Relationship between base-step variability and packing constraints

As shown by different authors (Gorin *et al.*, 1995; Subirana & Faria, 1997), each base step has a different average twist and a different flexibility. Sequences which contain rigid base steps and have an average twist very different from 36° may be unable to crystallize. In other cases, the presence of flexible base steps may help the molecule to adapt itself to the constraints of the lattice. For example the decamer CCAACGTTGG (BDJ019 in Table 1) is probably able to crystallize owing to the flexibility of the C·A/T·G steps, which compensates for the low twist and rigidity of most of the other steps (Subirana & Faria, 1997).

9. Discussion

In an analysis of some of the structures discussed in this paper, Dickerson *et al.* (1994) concluded that crystal packing or lattice forces are of secondary importance in the determination of the conformational features of B-DNA dodecamers. This is, in general, true for many features of DNA structure, but our study shows that the average twist is strongly related to the packing constraints.

The analysis presented in this paper shows that when dodecamers are crystallized in the B form, their average conformation is strongly restricted to remain in the range of 10 ± 0.1 base pairs per turn, as shown in Tables 1, 4 and 5. The lateral interactions of the oligonucleotides differ in each particular case, as illustrated in Figs. 2, 3, 5 and Table 3. Nevertheless, in all cases the average twist of the structure is the same. Packing is optimized by allowing base steps and space group to vary in order to improve phosphate–phosphate interactions, while the average conformation of the double helix is maintained. The conclusions do not apply to the few dodecamers which have been crystallized in the *R*3 space group (Timsit *et al.*, 1989; Berman *et al.*, 1992).

From the results summarized in Table 4, it is clear that decamers may show a wider range of variability than dodecamers, which have a practically constant average twist (Table 1). The same oligonucleotide may be crystallized in two different classes of those presented in Table 5 and vary from 9.7 to 10.6 base pairs per turn, as is the case for CGATCGATCG.

It is interesting to note that the DI has only been found in the dodecamers and in one octamer. Decamers which start with CG prefer either to form pseudocontinuous helices (Table 4) or to acquire a different type of end-to-end interaction, in which the terminal base pair is disrupted (Spink *et al.*, 1995).

The limitations imposed by packing may explain why some sequences are reluctant to crystallize or do so with a limited resolution, since they may not favour the average twist imposed by the lattice or may have inadequate phosphate–phosphate interactions.

In the case of dodecamers, infinite columns of duplexes related among themselves by either screw or dyad axes may pack in many different ways as shown in Table 1 and in Fig. 2. An intriguing feature of all the self-complementary dodecamers which start with the CG sequence and which have been crystallized is that they never use their chemical symmetry in crystal packing. The possibilities available suggest that the dodecamers may acquire their most stable structure compatible with sequence and crystallization conditions, although it is not clear what determines the space group chosen by a particular dodecamer.

Our study demonstrates that the crystal lattice has a strong influence on the average twist of DNA. As a result, oligonucleotides crystallized in the B form seldom have the average twist (10.6 base pairs per turn) usually found in solution (Niederweis *et al.*, 1992). Only in one particular case (class 3 in Table 5) does the twist found in crystals coincide with the solution value. Even in this case the average twist appears to be imposed by the crystal lattice.

This work has been supported in part by grants from the DGICYT (PB93-1067), the Comissionat de Recerca de la Generalitat de Catalunya and CESCA. VT acknowledges the support received from the Ministerio de Educación y Ciencia. We also thank Dr L. Malinina for valuable suggestions.

References

Baikalov, I., Grzeskowiak, K., Yanagi, K., Quintana, J. & Dickerson, R. E. (1993). *J. Mol. Biol.* **231**, 768–784.
Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelvin, A., Demeny, T., Hsieh, S.-H., Srinivasan, R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.
Cheng, X., Balendiran, K., Schildkraut, I. & Anderson, J. E. (1994). *EMBO J.* **13**, 3927–3935.
Dickerson, R. E. (1998). *Structure, Motion, Interaction and Expression of Biological Macromolecules*, edited by R. H. Sarma & M. H. Sarma, pp. 17–36. New York: Adenine Press.
Dickerson, R. E., Goodsell, D. S., Kopka, M. L. & Pjura, P. E. (1987). *J. Biomol. Struct. Dyn.* **5**, 557–579.
Dickerson, R. E., Goodsell, D. S. & Neidle, S. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 3579–3583.
DiGabriele, A. D. & Steitz, T. A. (1993). *J. Mol. Biol.* **231**, 1024–1039.
Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S. Itakura, K. & Dickerson, R. E. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 2179–2183.
El Hassan, M. A. & Calladine, C. R. (1997). *Philos. Trans. R. Soc. London Ser. A*, **355**, 43–100.
Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. (1982). *J. Biol. Chem.* **257**, 14686–14707.
Gorin, A. A., Zhurkin, V. B. & Olson, W. K. (1995). *J. Mol. Biol.* **247**, 34–48.
Heinemann, U. & Alings, C. (1989). *J. Mol. Biol.* **210**, 369–381.
Heinemann, U., Alings, C. & Hahn, M. (1994). *Biophys. Chem.* **50**, 157–167.

Hunter, W. N., Langlois d'Estaintot, B. & Kennard, O. (1988). *J. Mol. Biol.* **202**, 921–922.
Lavery, R. & Sklenar, H. (1988). *J. Biomol. Struct. Dyn.* **6**, 63–91.
Leonard, G. A. & Hunter, W. N. (1993). *J. Mol. Biol.* **234**, 198–208.
Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J. & Dickerson, R. E. (1993). *Biochemistry*, **32**, 1373–1389.
Niederweis, M., Lederer, T. & Hillen, W. (1992). *J. Mol. Biol.* **228**, 322–326.
Shui, X., McFail-Isom, L., Hu, G. G. & Williams, L. D. (1998). *Biochemistry*, **37**, 8341–8355.
Spink, N., Nunn, C. M., Vojtechovsky, J., Berman, H. M. & Neidle, S. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 10767–10771.
Subirana, J. A. & Faria, T. (1997). *Biophys. J.* **73**, 333–338.
Subirana, J. A., Salas, X., Urpí, L., Font, E. & Verdaguer, N. (1995). *Nuevas Tendencias en Cristalografía*, edited by F. H. Cano, C. Foces-Foces & M. Martínez-Ripoll, pp. 223–240. Madrid: Ediciones CSIC.
Tereshko, V., Urpí, L., Malinina, L., Huynh-Dinh, T. & Subirana, J. A. (1996). *Biochemistry*, **35**, 11589–11596.
Timsit, Y., Westhof, E., Fuchs, R. P. P. & Moras, D. (1989). *Nature (London)*, **341**, 459–462.
Urpí, L., Tereshko, V., Malinina, L., Huynh-Dinh, T. & Subirana, J. A. (1996). *Nature Struct. Biol.* **3**, 325–328.
Vojtechovsky, J., Eaton, M. D., Gaffney, B., Jones R. & Berman, H. M. (1995). *Biochemistry*, **34**, 16632–16640.
Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. & Dickerson, R. E. (1980). *Nature (London)*, **287**, 755–758.