ELSEVIER

Biophysical Chemistry 50 (1994) 157–167

Biophysical Chemistry

Crystallographic studies of DNA helix structure

Udo Heinemann *[a,b], Claudia Alings [b], Michael Hahn [a,b]

[a] *Forschungsgruppe Kristallographie, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle Strasse 10, 13122 Berlin, Germany*
[b] *Institut für Kristallographie, Freie Universität, 14195 Berlin, Germany*

(Received 5 January 1994)

Abstract

X-ray crystallography can reveal the three-dimensional structure of short fragments of DNA or RNA with unique precision. It provides information concerning both the global helical structure and the geometry of local features such as base-pair stacking patterns and backbone conformation. An analysis of the structures of a family of DNA decamers with related sequences, crystallizing in a number of different lattices, defines the ranges in which conformational parameters can vary in B-DNA helices and shows the correlations between them. Thus, these studies show the static structures and give insight into the mechanics of DNA helices by showing how a change of one local conformational parameter will influence others. Crystal structures are also used to assess the competing influences of nucleotide sequence and environment on the three-dimensional DNA structure. To extrapolate from DNA crystal structures to physical characteristics and function of these molecules in solution or embedded into a defined sequence context remains a major challenge.

Key words: X-ray diffraction; DNA double helix; Sequence-dependent conformation; Crystal packing; Base-pair stacking

1. Introduction

It is now about 15 years since the first crystal structures of short synthetic DNA fragments in the canonical A, B and Z forms were reported [1–4]. These analyses radically changed the perception of DNA double helices. They are no longer considered regular and static polymers but irregular and dynamic molecules with sequence and environment dependent local conformations and deformabilities instead. The new view of DNA structures, summarized in a number of excellent reviews [4–9], raises the possibility that local structural features of DNA helices are important for their interaction with large and small ligands such as proteins and drug molecules. An understanding of the influence of the nucleotide sequence and the solution environment on the helix conformation would thus permit the prediction of functional aspects from the DNA sequence.

Here we shall limit ourselves to a discussion of a subset of the crystallographic analyses of DNA fragments observed in the biologically most significant B form. Early studies of B-DNA have fo-

* Corresponding author.

SSDI 0301-4622(94)00017-E

Table 1
Crystallographic data for self-complementary DNA decamers. Only those structures are listed for which atomic coordinates are in the public domain

Sequence	Space group	a (Å)	b (Å)	c (Å)	β (°)	NDB code [a]	Ref.
C-C-A-A-G-A-T-T-G-G	C2	32.52	26.17	34.30	118.90	BDJ008	[16]
C-C-A-G-G-C-C-T-G-G	C2	32.15	25.49	34.82	116.71	BDJ017	[17]
C-C-A-A-C-G-T-T-G-G	C2	32.25	25.53	34.38	113.40	BDJ019	[18]
C-G-A-T-C-G-A-T-C-G	$P2_12_12_1$	38.93	39.63	33.30	90.00	BDJ025	[19]
C-G-A-T-T-A-A-T-C-G	$P2_12_12_1$	38.60	39.10	33.07	90.00	BDJ031	[20]
C-G-A-T-A-T-A-T-C-G · Ca^{2+}	$P2_12_12_1$	38.76	40.06	33.73	90.00	BDJ036	[21]
C-G-A-T-A-T-A-T-C-G · Mg^{2+}	$P2_12_12_1$	38.69	39.56	33.64	90.00	BDJ037	[21]
C-C-G-G-C-G-C-C-G-G	R3	54.07	54.07	44.59	90.00	BDJ039	[22]
C-C-A-A-C-I-T-T-G-G	$P3_221$	33.23	33.23	94.77	90.00	BDJ043	[23]
C-C-A-A-C-I-T-T-G-G	C2	31.87	25.69	34.21	114.10	BDJ044	[23]
C-C-A-A-I-A-T-T-G-G	C2	32.21	25.14	34.14	114.70	BDJ045	[30]
C-A-T-G-G-C-C-A-T-G	$P2_12_12_1$	36.60	42.49	34.69	90.00	BDJ051	[24]
C-C-A-G-G-C-m^5C-T-G-G	P6	53.77	53.77	34.35	90.00	BDJB27	[25]
C-G-A-T-C-G-m^6A-T-C-G	$P3_221$	33.38	33.38	98.30	90.00	BDJB48	[26]

[a] NDB is the Rutgers Nucleic Acids Database [29].

cussed on duplexes with sequences related to that of the classical Dickerson–Drew dodecamer C-G-C-G-A-A-T-T-C-G-C-G [2,10–15]. More recently the emphasis has shifted to the study of self-complementary decamer duplexes [16–28]. These molecules offer the advantages of crystallizing in a variety of sequences and crystal forms and diffracting to better than 2 Å resolution in most cases (Tables 1 and 2). They may thus be used to evaluate the experimental limits of DNA crystal-

Table 2
Structure refinement of self-complementary DNA decamers

Sequence	a.u. [a]	d_{lim} [b] (Å)	F_0/bp	R [c] (%)	n_{solv} [d]	n_{ion} [e]
C-C-A-A-G-A-T-T-G-G	strand	1.30	681	18.5	138	6
C-C-A-G-G-C-C-T-G-G	strand	1.60	484	16.9	84	4
C-C-A-A-C-G-T-T-G-G	strand	1.40	880	16.0	118	8
C-G-A-T-C-G-A-T-C-G	duplex	1.50	511	16.1	142	2
C-G-A-T-T-A-A-T-C-G	duplex	1.50	379	15.7	108	1
C-G-A-T-A-T-A-T-C-G · Ca^{2+}	duplex	1.70	368	17.8	87	1
C-G-A-T-A-T-A-T-C-G · Mg^{2+}	duplex	2.00	182	16.5	55	1
C-C-G-G-C-G-C-C-G-G	duplex	2.20	202	16.7	47	–
C-C-A-A-C-I-T-T-G-G	duplex	2.20	172	16.2	36	1
C-C-A-A-C-I-T-T-G-G	strand	1.30	1005	15.2	144	1
C-C-A-A-I-A-T-T-G-G	strand	2.00	239	13.4	100	–
C-A-T-G-G-C-C-A-T-G	duplex	2.00	322	19.6	49	–
C-C-A-G-G-C-m^5C-T-G-G	duplex	1.75	380	17.4	80	–
C-G-A-T-C-G-m^6A-T-C-G	duplex	2.00	228	17.2	45	1.5

[a] Content of the asymmetric unit, either a single strand or a self-complementary duplex.
[b] Limit resolution to which observations were included in the refinement.
[c] The crystallographic residual $R = 100 \times (\Sigma | F_o - F_c | / \Sigma F_o)$, where F_o and F_c are observed and calculated structure amplitudes, respectively.
[d] Number of solvent molecules.
[e] Number of ions included in the refinement.

lography, the ranges of variations in global helix structure and in local conformational features, the interdependence between local conformational states and the influence of crystal environment on double helix structure.

2. Experimental limits

Before considering structural details of DNA molecules it is necessary to examine the crystallographic methods they are based upon. The reason for this is twofold. First, in nearly all cases the DNA structures are based on model helices derived from X-ray fiber diffraction [31] which are used to solve the crystallographic phase problem by molecular replacement techniques. Second, although B-DNA decamers often yield very good X-ray diffraction patterns, the resolution is never high enough to resolve individual atoms making the use of stereochemical restraints in the refinement process necessary. We must therefore assess to which degree a DNA crystal structure is determined by the experimental data and how far the influence of additional factors goes.

This problem has been addressed several times by comparing DNA models based on the same set of X-ray diffraction data and derived from different structure refinement algorithms [27,32,33]. The result of these studies is that, at a nominal resolution of 2 Å or better, the DNA model is only slightly influenced by the structure refinement protocol used. A least-squares superposition of 3 models for the decamer duplex C-C-A-G-G-C-m^{5}C-T-G-G [27] derived from 3 different structure refinement schemes [34–36] shows only minor differences between the models (Fig. 1). Whereas base-pair geometries and stacking is virtually identical, notable deviations occur in sugars and phosphate groups towards the ends of the helix. This corresponds well with thermal motions of the double helical fragment as indicated by the crystallographic temperature factors which increase from the bases over the sugars to the phosphates and from the center to the end of the duplex. In spite of the satisfactory agreement between the DNA models, some of the derived structural parameters, such as backbone torsion angles, have been shown to be sensitive to the restraints applied in the refinement. Significant differences between models have been observed in the hydration spheres which are in partial agreement only. The study with C-C-A-G-G-C-m^{5}C-T-G-G was carried out using X-ray diffraction data to 1.6 Å resolution. As the experimental data become more scarce going to lower resolution, the influence of the restraints employed in structure refinement and of other factors such as crystal packing and partial disorder will become more pronounced.

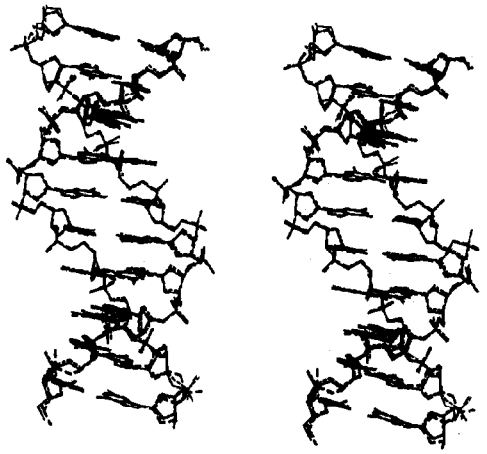

Fig. 1. Least-squares superposition of 3 models for C-C-A-G-G-C-m^{5}C-T-G-G derived from stereochemically restrained least-squares refinement with the programs NUCLIN/NUCLSQ [34], TNT [35] and X-PLOR [36] against 1.6 Å X-ray diffraction data. The models are drawn with solid bonds, thin dashed bonds and heavy dashed bonds, respectively. This drawing as well as fig. 2 was prepared with MOLSCRIPT [37].

3. Global helix structure

In their crystal lattices all B-DNA decamers investigated here stack end-to-end to form quasi-continuous double helices with a periodicity of 10.0 base pairs per turn (Fig. 2). This value is also found for B-DNA in fibers [31], but differs clearly from the helical periodicity of 10.6($\pm$0.1) base pairs per turn observed in solution studies of general-sequence B-DNA [38,39]. Another property shared by the studied decamers is a 5′-termi-

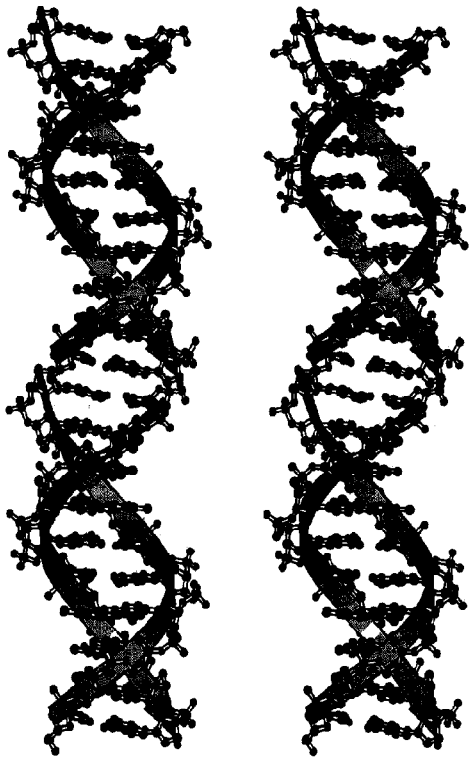

Fig. 2. Stereo drawing of C-C-A-G-G-C-m^{5}C-T-G-G. Two duplex molecules are stacked in the crystal lattice to form a fragment of a quasi-continuous double helix traversing the crystal. The continuous portions of the sugar–phosphate backbones are highlighted by curved arrows penetrating the deoxyribose moieties and pointing in 5′ to 3′ direction.

nal cytosine followed in all but one cases by either C or G. In the molecules crystallizing in space group C2 the two chemically identical strands of the duplex are related by exact dyad symmetry, whereas in the other DNA fragments they are not.

Although all fall within the well established B-helical domain with about equally deep major and minor grooves, relatively straight helix axes and sugar puckers centered around C2′-*endo*, the 14 decamers display distinct conformational differences (Fig. 3). The accessibility of bases in the helical grooves which is of great importance for the sequence-specific and unspecific interaction of DNA with ligands is modulated significantly. Since there is no large variation in *rise* per base pair, a wide major groove implies a narrow minor groove and *vice versa*. Near the center of the helix, the minor groove can be as narrow as 2.6 Å (in C-G-A-T-T-A-A-T-C-G) and as wide as 8.9 Å (in C-C-G-G-C-G-C-C-G-G) after subtracting 5.8 Å from the P–P separation to account for the phosphate group van del Waals radii. Consecutive purine–purine mismatches as present in C-C-A-A-G-A-T-T-G-G and in C-C-A-A-I-A-T-T-G-G cause a widening of the minor groove, but are accommodated easily in a standard B-DNA environment. Base variations such as the exchange of inosine for guanine or of 5-methylcytosine or of 6-methyladenine for their unmethylated counterparts have no major effect on global helix structure. There is a tendency of AT-rich stretches of DNA to have narrowed minor grooves and of GC-rich stretches to have widened minor grooves. In this view, Gm^{5}C base pairs behave like AT and not like GC base pairs.

It is worthwhile to examine the geometry of the minor groove more closely, since this is where many types of drug molecules bind and where distinct hydration patterns are observed [16,41,42]. In B-DNA the width of the groove is limited by the approach of phosphate groups from the two strands. The closest approach is usually between the phosphate groups belonging to base pairs i and $i-4$. In the standard fiber model of B-DNA [31] this P–P distance is uniformly 5.9 Å. In the crystals structures of B-DNA decamers there are significant deviations towards either widened or narrowed grooves (Fig. 4). In addition, the groove width may vary drastically over the length of the decamer duplex, e.g. from 3.0 to 8.2 Å over four base pairs in C-G-A-T-C-G-A-T-C-G. Due to the stacking of duplexes in the crystals, the grooves are continuous beyond the decamer fragments represented in the figure. However, since the sugar–phosphate backbones are less restricted at the fragment termini due to the missing phosphodiester links, groove widths have not been analyzed across these breaks. In general, the opposite trends to those described for minor groove geometry are observed for the major groove of B-DNA.

4. Range of variation

In the same way as the global structures and groove geometries differ, local conformational parameters vary in the 14 B-DNA decamers. These variations become obvious upon a systematic analysis of the relative orientation and positioning of (a) the two bases of a pair, (b) a base pair with respect to a helix axis and (c) two adjacent base pairs in the helix. Eliminating the two trivial parameters describing the rotation of a base pair about and its translation along the helix axis, this analysis yields the 16 helical parameters which have been defined in the 1988 Cambridge EMBO Workshop [43]. Here we shall limit the discussion to the five important parameters *propeller*, defining the relative rotation of two bases about the long axis of the pair, *buckle*, defining the relative rotation of two bases about the pseudodyad axis of a pair, *twist*, defining the relative rotation of two adjacent base pairs about the helix axis, *roll*, defining the relative rotation of two adjacent base pairs about their common long axis, and *slide*, defining the relative translation of two adjacent base pairs along their common long axis.

In the crystalline B-DNA decamers all base pairs deviate from coplanarity of the two bases by showing *propeller* twisting with a negative sign, i.e. when viewing the base pair edge-on along its long axis passing through purine C8 and pyrimidine C6, the distant base is rotated anti-clockwise with respect to the near base. The *propeller* is thought to increase intra-strand stacking efficiency in all double-helical forms of DNA [44] at the expense of hydrogen bonding energy. Therefore, AT base pairs are expected to display a higher degree of *propeller* than GC pairs since only two and not three hydrogen bonds are distorted. In principle, this is in fact observed with mean values of *propeller* of $-11.4°(\pm 4.6°)$ for AT and m^6AT and $-8.7°(\pm 5.4°)$ for GC, Gm^5C and IC base pairs. The large standard deviations for both samples indicate, however, that there is no clear separation in the *propeller* twisting of AT and GC base pairs. Since the *propeller* contributes to a closing of the minor groove, here is one factor, but probably not the only one, explaining the tendency of narrow minor grooves at AT stretches of B-DNA. GC base pairs in the test sample are more distorted by *buckle* than AT base pairs with mean magnitudes of $5.7°(\pm 3.8°)$ versus $3.2°(\pm 2.5°)$.

The base pair stacking parameters *twist, roll* and *slide* can be analyzed in terms of the 10 unique base-pair steps in duplex DNA [45]. This has been done for those 14 B-DNA decamers containing only Watson–Crick base pairs and no mismatches (Fig. 5). With the exception of ApC/GpT (present 3 times in the sample) all base-pair steps occur at least 5 times in these crystal structures after elimination of symmetry-related steps. CpG/CpG is observed 18 times in the sample structures. In spite of the uneven distribution of stacks, some trends are discernible.

An immediately obvious fact is that every base-pair step can adopt a fairly wide range of stacking geometries as indicated by the vertical bars giving plus or minus one standard deviation of the sample. Mean values differ considerably between 26.1° and 44.0° for the *twist* of ApG/CpT and CpA/TpG, between $-7.0°$ and 3.6° for the *roll* of ApC/GpT and CpC/GpG and between -0.24 Å and 1.73 Å for the *slide* of ApT/ApT and CpA/TpG, respectively. There is no clear separation between stacking preferences of purine–purine, purine–pyrimidine and pyrimidine–purine steps. Therefore, attempts to predict the three-dimensional structure of short pieces of DNA based on dinucleoside phosphate or purine/pyrimidine patterns have met with very limited success [44,47]. It may be possible, however, to use parameters derived from stacking geometries as seen in crystal structures to predict the curvature of double helical DNA in solution [48].

Very large standard deviations and unusual mean values are observed with $44.0°(\pm 6.7°)$, $-5.8°(\pm 8.7°)$ and $1.73(\pm 0.85)$ Å for *twist, roll* and *slide* of the CpA/TpG stacking step. A closer look reveals 7 CpA/TpG steps with a *twist* of $48.8°(\pm 2.2°)$, a *roll* of $-12.2°(\pm 2.2°)$ and a *slide* of $2.39(\pm 0.30)$ Å contrasting 4 such steps with a *twist* of $35.5°(\pm 1.4°)$, a *roll* of $5.5°(\pm 1.1°)$ and a *slide* of $0.74(\pm 0.18)$ Å. Thus, there is a pro-

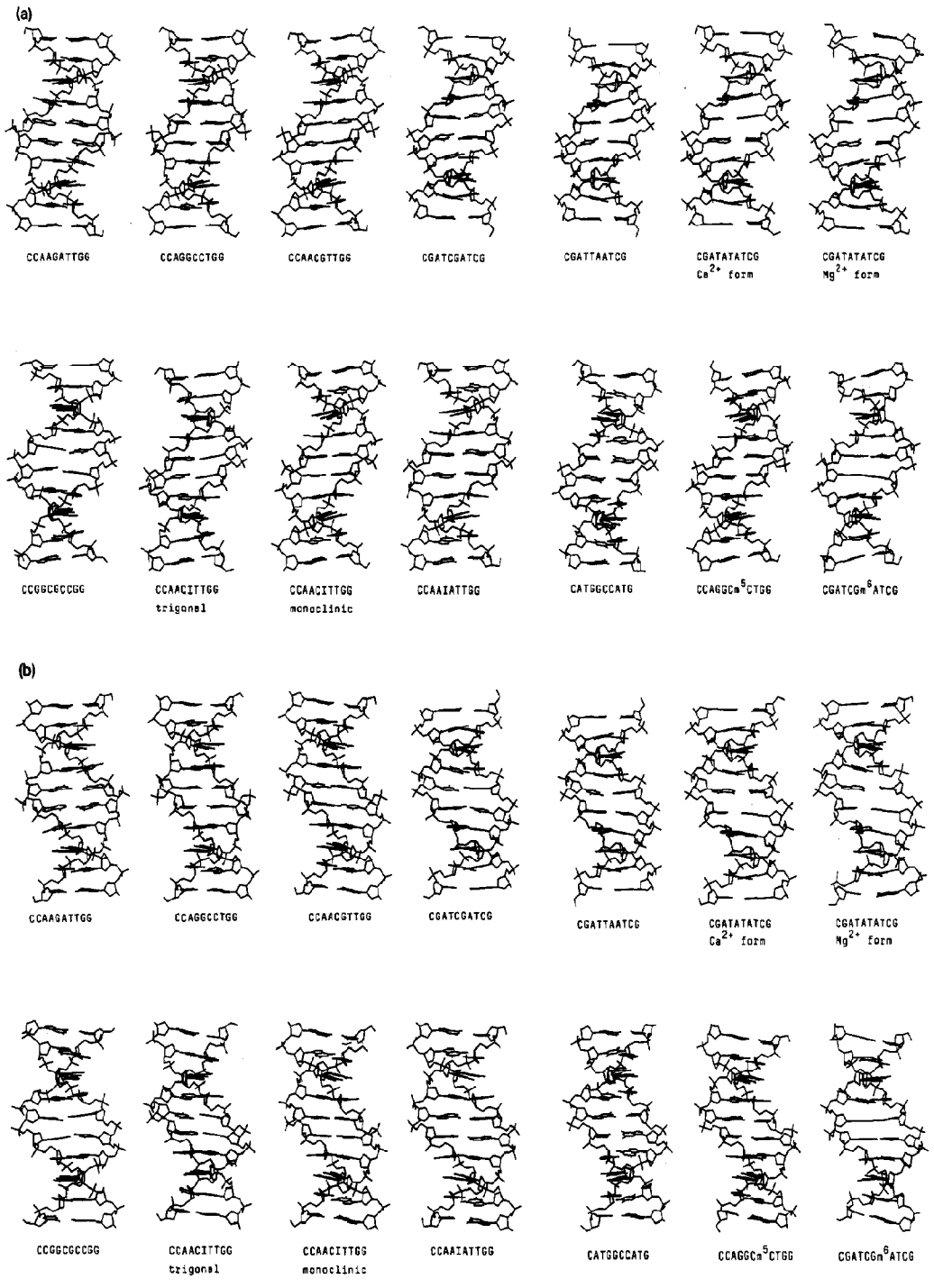

Fig. 3. Wire diagrams of the 14 B-DNA decamers. The views are into the major groove (a) and into the minor groove (b). Note the differences in groove widths and in base-pair geometry and stacking. This drawing as well as fig. 6 was prepared with SCHAKAL [40].

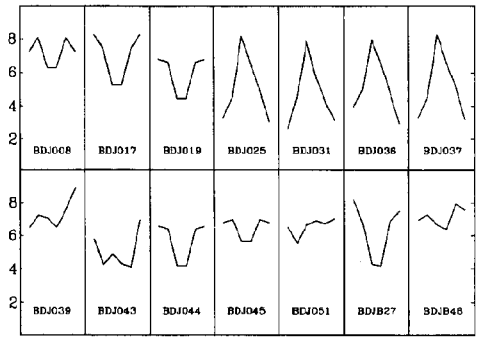

Fig. 4. Minor groove width in angstrom units of the 14 B-DNA decamers expressed as the shortest phosphorus-phosphorus separation across the groove minus 5.8 Å to account for the phosphate group van der Waals radii. The structures are identified by their Nucleic Acid Database codes. Note the different degrees of molecular asymmetry displayed in the diagrams of those DNA fragments not crystallizing in space group C2.

nounced bistability of two rather closely defined conformational states with either large *twist*, negative *roll* and large *slide* or medium *twist*, positive *roll* and medium *slide*. Such bistability is observed for no other base-pair stack in B-DNA.

5. Mechanics of B-DNA structure

We have seen that the peculiar high-*twist* conformation of the CpA/TpG base-pair step involves a combination of *twist*, *roll* and *slide*. Obviously, these parameters are not totally independent of each other; changing one will affect others. This is an example of the mechanics of DNA structure which is slowly being unravelled by X-ray crystallography. The most important results emerging from a new crystal structure are not the static image of the double helix it provides but information about the interdependence of structural parameters, i.e. the underlying structural principle. Analyzing the geometry of the CpA/TpG step more closely [17] it becomes clear that further structural adjustments are required for the occurrence of the high-*twist* conformation: on both chains the connecting sugar–phosphate backbone has to adopt the B_{II} conformation which itself is defined by a correlated motion of backbone torsion angles ϵ and ζ [18].

A systematic analysis of the mechanics of B-DNA structure can be obtained by computing pairwise correlations between conformational parameters. This approach has been pioneered [10] and perfected [49] by Dickerson and co-workers who have used linear regression analysis to obtain a large number of structural correlations in B-DNA. Interestingly, some of these correlations are not only valid for B-DNA but for the entire superfamily of right-handed antiparallel double helices. For instance, the base pair *slide* is structurally linked with the *displacement* of base pairs

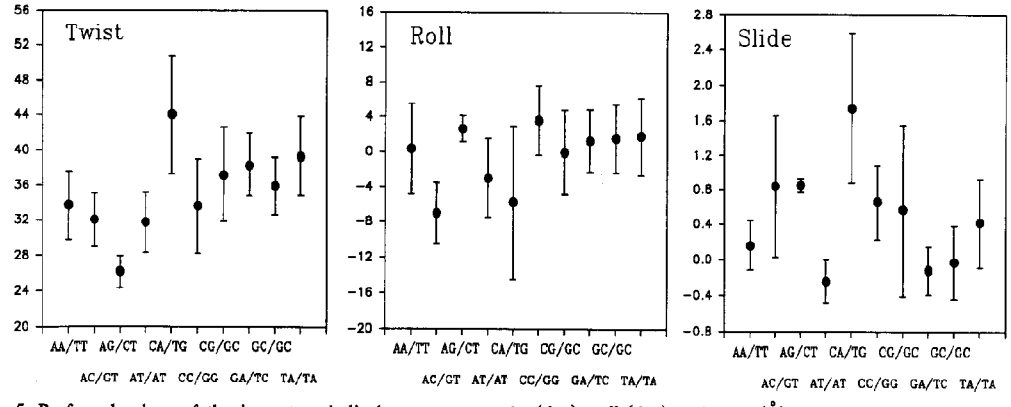

Fig. 5. Preferred values of the important helical parameters *twist* (deg), *roll* (deg) and slide (Å) in the 10 unique dinucleoside phosphate steps for those 12 B-DNA decamers which have strictly Watson–Crick base pairing. Parameter values given as mean plus/minus one standard deviation were computed with NEWHEL91 [46] and agree with the Cambridge convention [43]. Note the unusual values of *twist* and *slide* for the CpA/TpG base-pair step and the large variability displayed by this step.

Table 3
Comparison of the crystal structures of B-DNA decamers. The C1′ positions of the duplexes were superimposed by the least-squares procedure of Kabsch [54]. Values above the diagonal are the root-mean-square deviations (Å) between duplexes, values below the diagonal are the maximum distances between corresponding C1′ positions. Structures are identified by their NDB codes

	BDJ008	BDJ017	BDJ019	BDJ025	BDJ031	BDJ036	BDJ037	BDJ039	BDJ043	BDJ044	BDJ045	BDJ051	BDJB27	BDJB48
BDJ008 [a]	–	0.779	1.095	1.972	2.000	1.909	2.000	1.377	1.808	1.133	0.700	1.549	1.373	2.256
BDJ017	1.83	–	0.678	1.976	1.989	1.854	1.956	1.279	1.656	0.757	0.613	1.376	0.941	2.159
BDJ019	1.70	1.44	–	1.765	1.747	1.652	1.738	1.192	1.330	0.250	0.569	1.376	0.617	1.867
BDJ025	3.07	2.80	2.76	–	0.238	0.409	0.405	1.425	1.094	1.677	1.763	1.541	1.727	1.273
BDJ031	2.99	2.78	2.73	0.49	–	0.439	0.420	1.434	1.052	1.657	1.777	1.552	1.709	1.278
BDJ036	2.87	2.47	2.41	0.72	0.71	–	0.339	1.278	1.012	1.575	1.676	1.370	1.576	1.168
BDJ037	3.06	2.80	2.84	0.62	0.77	0.78	–	1.412	1.124	1.660	1.752	1.488	1.676	1.319
BDJ039	2.10	2.03	1.87	3.16	3.17	2.59	2.87	–	1.139	1.155	1.199	0.897	1.157	1.432
BDJ043 [b]	2.51	2.53	2.16	1.88	1.93	1.40	1.86	2.30	–	1.245	1.508	1.330	1.153	0.999
BDJ044 [b]	1.77	1.45	0.52	2.56	2.53	2.20	2.67	1.96	2.17	–	0.662	1.379	0.644	1.787
BDJ045 [a,b]	1.03	1.24	1.17	2.49	2.47	2.20	2.47	1.88	2.43	1.30	–	1.415	0.876	2.002
BDJ051	2.08	2.05	2.34	2.19	2.26	2.04	2.12	1.55	2.05	2.32	2.27	–	1.267	1.622
BJDB27 [c]	2.03	1.44	1.08	2.87	2.88	2.52	2.71	1.86	1.67	1.19	1.49	2.23	–	1.634
BDJB48 [c]	3.46	3.53	3.17	2.08	2.06	1.94	2.47	2.56	2.14	3.13	3.41	2.47	2.75	–

[a] Structures with purine–purine base mismatches.
[b] Structures containing inosine.

from the helix axis, and this holds true for the entire range of conformations from A- to B-form DNA [50].

6. Influence of environment on conformation

It has long been suspected that the crystal structures of short DNA duplexes are strongly influenced by crystal packing forces which may affect their conformations more than the nucleotide sequence [51,52]. If this were so, one would predict their solution conformation to be quite different from the crystal structure and their biological significance would be very limited. The family of self-complementary B-DNA decamers provides a unique opportunity to assess the competing influence of nucleotide sequence and crystal environment on double helix conformation. The molecules have different, but related sequences and crystallize in five different lattices.

Whenever asymmetry between the chemically identical strands of a self-complementary duplex is observed, this is an effect of crystal packing [53]. The effect can clearly be seen in Figs. 3 and 4 in all molecules except for those crystallizing in space C2 where the two strands are related by exact (crystallographic) dyad symmetry. To examine further the influence of crystal environment of DNA conformation all crystal structures have been compared by matching their C1′ positions in a least-squares fit (Table 3). A close match with r.m.s. deviations below 0.5 Å is found between C-C-A-A-C-G-T-T-G-G and C-C-A-A-C-I-T-T-G-G crystallizing in the same space group (C2) and differing only in the presence of either guanine or inosine in position 6, as well as among the structures of C-G-A-T-C-G-A-T-C-G, C-G-A-T-T-A-A-T-C-G and the two forms of C-G-A-T-A-T-A-T-C-G, all crystallizing in $P2_12_12_1$. This does not imply, however, that the same crystal lattice will necessarily yield the same global conformation: C-A-T-G-G-C-C-A-T-G also crystallizes in $P2_12_12_1$ but bears no similarity with the other orthorhombic structures. There are three examples of identical or near identical oligonucleotides crystallizing in different lattices: C-C-A-G-G-C-C-T-G-G (space group C2) and C-C-A-G-G-C-m^{5}C-T-G-G (P6), C-G-A-T-C-G-A-T-C-G ($P2_12_12_1$) and C-G-A-T-C-G-m^{6}A-T-C-G ($P3_221$), and C-C-A-A-C-I-T-T-G-G ($P3_221$ and C2). In all cases, structural differences between these molecules as judged from r.m.s. deviations of C1′ atoms exceed differences between duplexes of different sequence within the same space group. It must be concluded therefore that lattice contacts influence the global structure of short synthetic DNA duplexes at least as much as the underlying nucleotide sequence.

Crystal packing not only influences the global helix structure, but also local conformational features. The unique CpA/TpG dinucleotide step may serve as a test case. From a structural comparison of C-C-A-G-G-C-C-T-G-G with its meth-

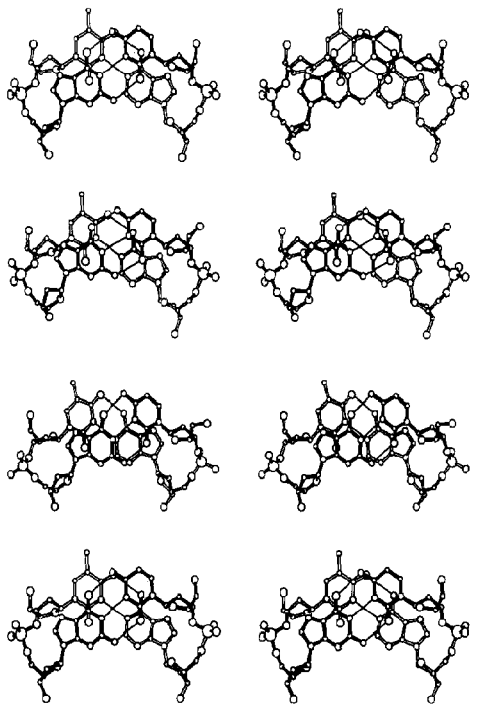

Fig. 6. Different effects of crystal packing on the conformation of the CpA/TpG dinucleoside phosphate steps in C-C-A-G-G-C-C-T-G-G (space group C2), C-C-A-G-G-C-m^{5}C-T-G-G (P6) and C-C-A-A-C-I-T-T-G-G ($P3_221$ and C2), from top to bottom. With the exception of the trigonal crystal structure, the CpA/TpG steps show unstacking of the bases accompanied by B_{II} backbone conformation on both connecting strands.

ylated counterpart where all CpA/TpG steps were found in the high-*twist* geometry it was concluded that salient structural features of DNA remain unaffected by changing crystal environment [55]. However, when the comparison could be repeated with C-C-A-A-C-I-T-T-G-G [23] the same steps were found in high-*twist* arrangement only in space group C2 but not in $P3_221$ (Fig. 6). Thus, the interaction energy between DNA molecules in a crystal lattice may suffice to overcome the energetical barrier between these two well-defined conformations.

The strong influence of crystal packing is strikingly demonstrated by the coexistence of A- and B-form DNA in the same lattice [56–59]. It should also causes structural differences in DNA molecules between the crystal and the solution state. This has been demonstrated by Raman [60,61] as well as NMR spectroscopy [62–65]. The picture of DNA emerging from these studies is that of a highly flexible and deformable molecule.

The forces acting on a DNA molecule in a crystal lattice are not principally different from those present in a protein–DNA complex. Here as well significant distortions of the double helix may be caused by sequence-specific protein binding. Prominent examples for protein-induced structural alterations are the DNA complexes with the restriction endonuclease EcoRV [66] and with eukaryotic TATA box binding proteins [67,68]. Especially in the latter case the DNA molecules are distorted to a degree not observed previously.

7. Conclusions

X-ray crystallography can reveal the three-dimensional structure of short helical fragments of nucleic acids with unmatched precision. Although usually not extending to atomic resolution, the available X-ray diffraction data permit reliable structure determinations nearly unaffected by technical artefacts. However, the crystal structure is of limited use only for prediction of the structure of a DNA fragment in solution or implications of its biological function when embedded in flanking sequences *in vivo*, since the conformation is strongly influenced by crystal packing. This does not imply that the crystal structure is arbitrary. It merely draws attention to the fact the process of crystallization selects one out of an unknown number of energetically favorable conformations. By studying these conformations and the mechanics of DNA structure revealed by them an understanding of the structural principles of DNA can be achieved.

Acknowledgement

We are grateful to Professor W. Saenger for many years of generous support. This work was funded by the Deutsche Forschungsgemeinschaft through SFB 344/D3 and the Fonds der Chemischen Industrie.

References

[1] Z. Shakked, D. Rabinovich, W.B.T. Cruse, E. Egert, O. Kennard, G. Sala, S.A. Salisbury and M.A. Viswamitra, Proc. Roy. Soc. B 213 (1981) 479.
[2] R.M. Wing, H.R. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R.E. Dickerson, Nature 287 (1980) 755.
[3] A.H.-J. Wang, G.J. Quigley, F.K. Kolpak, J.L. Crawford, J.H. van Boom, G.A. van der Marel and A. Rich, Nature 282 (1979) 680.
[4] R.E. Dickerson, H.R. Drew, B.N. Conner, R.M. Wing, A.V. Fratini and M.L. Kopka, Science 216 (1982) 475.
[5] Z. Shakked and D. Rabinovich, Progr. Biophys. Mol. Biol. 47 (1986) 159.
[6] O. Kennard and W.N. Hunter, Quart. Rev. Biophys. 22 (1989) 327.
[7] O. Kennard and W.N. Hunter, Angew. Chem. Intern. Ed. Engl. 30 (1991) 1254.
[8] R.E. Dickerson, in: Structure and methods, Vol. 3. DNA and RNA, eds. R.H. Sarma and M.H. Sarma (Adenine Press, Albany, 1990) p. 1.
[9] R.E. Dickerson, Methods Enzymol. 211 (1992) 67.
[10] R.E. Dickerson and H.R. Drew, J. Mol. Biol. 149 (1981) 761.
[11] H.C.M. Nelson, J.T. Finch, B.F. Luisi and A. Klug, Nature 330 (1987) 221.
[12] A.D. DiGabriele, M.R. Sanderson and T.A. Steitz, Proc. Natl. Acad. Sci. USA 86 (1989) 1816.
[13] C. Yoon, G.G. Privé, D.S. Goodsell and R.E. Dickerson, Proc. Natl. Acad. Sci. USA 85 (1988) 6332.
[14] T.A. Larsen, M.L. Kopka and R.E. Dickerson, Biochemistry 30 (1990) 4443.

[15] N. Narendra, S.L. Ginell, I.M. Russu and H.M. Berman, Biochemistry 30 (1990) 4450.
[16] G.G. Privé, U. Heinemann, S. Chandrasegaran, L.-S. Kan, M.L. Kopka and R.E. Dickerson, Science 238 (1987) 498.
[17] U. Heinemann and C. Alings, J. Mol. Biol. 210 (1989) 369.
[18] G.G. Privé, K. Yanagi and R.E. Dickerson, J. Mol. Biol. 217 (1991) 177.
[19] K. Grzeskowiak, K. Yanagi, G.G. Privé and R.E. Dickerson, J. Biol. Chem. 266 (1991) 8861.
[20] J.R. Quintana, K. Grzeskowiak, K. Yanagi and R.E. Dickerson, J. Mol. Biol. 225 (1992) 379.
[21] H. Yuan, J. Quintana and R.E. Dickerson, Biochemistry 31 (1992) 8009.
[22] U. Heinemann, C. Alings and M. Bansal, EMBO J. 11 (1992) 1931.
[23] A. Lipanov, M.L. Kopka, M. Kaczor-Grzeskowiak, J. Quintana and R.E. Dickerson, Biochemistry 32 (1993) 1373.
[24] D.S. Goodsell, M.L. Kopka, D. Cascio and R.E. Dickerson, Proc. Natl. Acad. Sci. USA 90 (1993) 2930.
[25] U. Heinemann and M. Hahn, J. Biol. Chem. 267 (1992) 7332.
[26] I. Baikalov, K. Grzeskowiak, K. Yanagi, J. Quintana and R.E. Dickerson, J. Mol. Biol. 231 (1993) 768.
[27] M. Hahn and U. Heinemann, Acta Cryst. D 49 (1993) 468.
[28] Y.-G. Gao, G.A. van der Marel, J.H. van Boom and A.H.-J. Wang, Biochemistry 30 (1991) 9922.
[29] H.M. Berman, W.K. Olson, D.L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.H. Hsieh, A.R. Srinivasan and B. Schneider, Biophys. J. 63 (1992) 751.
[30] A.A. Lipanov, M.L. Kopka, M. Kaczor-Grzeskowiak and R.E. Dickerson, to be published.
[31] R. Chandrasekaran and S. Arnott, in: Landolt-Börnstein, New Series, Group VII, Vol. 1b, ed. W. Saenger (Springer, Berlin, 1989) p. 31.
[32] S.R. Holbrook, R.E. Dickerson and S.-H. Kim, Acta Cryst. B 41 (1985) 255.
[33] E. Westhof, J. Biomol. Struct. Dyn. 5 (1987) 560.
[34] E. Westhof, P. Dumas and D. Moras, J. Mol. Biol. 184 (1985) 119.
[35] D.E. Tronrud, L.F. TenEyck and B.W. Matthews, Acta Crystallogr. A 43 (1987) 489.
[36] A.T. Brünger, J. Kuriyan and M. Karplus, Science 235 (1987) 458.
[37] J.P. Kraulis, J. Appl. Cryst. 24 (1991) 946.
[38] L.J. Peck and J.C. Wang, Nature 292 (1981) 375.
[39] D. Rhodes and A. Klug, Nature 292 (1981) 378.
[40] E. Keller, in: Molecular modelling, eds. B.M. Bussian and E. Weber (Bundesministerium für Forschung und Technologie, Bonn, 1988) p. 87.
[41] C. Zimmer and U. Wähnert, Progr. Biophys. Mol. Biol. 41 (1986) 31.
[42] V.P. Chuprina, U. Heinemann, A.A. Nurislamov, P. Zielenkiewicz, R.E. Dickerson and W. Saenger, Proc. Natl. Acad. Sci. USA 88 (1991) 593.
[43] R.E. Dickerson, M. Bansal, C.R. Calladine, S. Diekmann, W.N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H.C.M. Nelson, W.K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D.M. Soumpasis, C.-S. Tung, A.H.-J. Wang and V.B. Zhurkin, EMBO J. 8 (1989) 1.
[44] R.E. Dickerson, J. Mol. Biol. 166 (1983) 419.
[45] U. Heinemann, C. Alings and H. Lauble, Nucleosides Nucleotides 9 (1990) 349.
[46] R.E. Dickerson, NEWHELIX user manual, University of California at Los Angeles (1992).
[47] C.R. Calladine, J. Mol. Biol. 161 (1982) 343.
[48] D. Mohanty and M. Bansal, J. Biomol. Struct. Dyn. 9 (1991) 127.
[49] K. Yanagi, G.G. Privé and R.E. Dickerson, J. Mol. Biol. 217 (1991) 201.
[50] U. Heinemann, J. Biomol. Struct. Dyn. 8 (1991) 801.
[51] Z. Shakked, G. Guerstein-Guzikevich, M. Eisenstein, F. Frolow and D. Rabinovich, Nature 342 (1989) 456.
[52] S. Jain, G. Zon and M. Sundaralingam, Biochemistry 28 (1989) 2360.
[53] R.E. Dickerson, D.S. Goodsell, M.L. Kopka and P.E. Pjura, J. Biomol. Struct. Dyn. 5 (1987) 557.
[54] W. Kabsch, Acta Cryst. A 32 (1978) 922.
[55] U. Heinemann and C. Alings, EMBO J. 10 (1991) 35.
[56] J. Doucet, J.-P. Benoit, W.B.T. Cruse, T. Prange and O. Kennard, Nature 337 (1989) 190.
[57] J. Liquiers, E. Taillandier, W.L. Peticolas and G.A. Thomas, J. Biomol. Struct. Dyn. 8 (1990) 295.
[58] H. Lauble, R. Frank, H. Blöcker and U. Heinemann, Nucl. Acids Res. 16 (1988) 7799.
[59] H. Fabian, W. Hölzer, U. Heinemann, H. Sklenar and H. Welfle, Nucl. Acids Res. 21 (1993) 569.
[60] J.M. Benevides, A.H.-J. Wang, A. Rich, Y. Kyogoku, G.A. van der Marel, J.H. van Boom and G.J. Thomas Jr., Biochemistry 25 (1986) 41.
[61] W.L. Peticolas, Y. Wang and G.A. Thomas, Proc. Natl. Acad. Sci. USA 85 (1988) 2579.
[62] D.G. Reid, S.A. Salisbury, S. Bellard, Z. Shakked and D.H. Williams, Biochemistry 22 (1983) 2019.
[63] L.J. Rinkel, M.A. Sanderson, G.A. van der Marel, J.H. van Boom and C. Altona, Eur. J. Biochem. 159 (1986) 85.
[64] M.A. Keniry, C. Levenson and R.H. Shafer, J. Biomol. Struct. Dyn. 4 (1987) 745.
[65] S. Wolk, W.N. Thurmes, W.S. Ross, C.C. Hardin and I. Tinoco Jr., Biochemistry 28 (1989) 2452.
[66] F.K. Winkler, D.W. Banner, C. Oefner, D. Tsernoglou, R.S. Brown, S.P. Heathman, R.K. Bryan, P.D. Martin, K. Petratos and K.S. Wilson, EMBO J. 12 (1993) 1781.
[67] Y. Kim, J.H. Geiger, S. Hahn and P.B. Sigler, Nature 365 (1993) 512.
[68] J.L. Kim, D.B. Nikolov and S.K. Burley, Nature 365 (1993) 520.