Eur Biophys J (1987) 15: 13–26

European Biophysics Journal

Molecular mechanics calculations of $dA_{12} \cdot dT_{12}$ and of the curved molecule $d(GCTCGAAAAA)_4 \cdot d(TTTTTCGAGC)_4$

E. von Kitzing* and S. Diekmann

Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg, D-3400 Göttingen-Nikolausberg, Federal Republic of Germany

Received September 25, 1986/Accepted in revised form January 26, 1987

Abstract. Using the AMBER software package (Weiner and Kollman 1981) substantially modified for electrostatic contributions, the structural energies of the double-stranded oligonucleotides $dA_{12} \cdot dT_{12}$ and $d(GCTCGAAAAA)_4 \cdot d(TTTTTCGAGC)_4$ were minimized. Using various starting structures for the molecule $dA_{12} \cdot dT_{12}$, one final structure is obtained which possesses the experimentally determined properties of poly(dA) · poly(dT). This structure is an A-form-B-form-hybrid structure similar to that of Arnott et al. (1983). The dA-strand is similar to an A-form while the dT-strand is similar to normal B-form. This structure and separately optimized B-form sequence stretches were used to construct the double-stranded fragment $d(GCTCGAAAAA)_4$ which again was optimized. This sequence, when imbedded in a DNA fragment as contiguous repeats, shows a gel migration anomaly which has been interpreted as stable curvature of the DNA (Diekmann 1986). The calculated structure of this sequence indeed has a curved helix axis and is discussed as a model for curved DNA. A theoretical formalism is presented which allows one to calculate the structural parameters of any nucleic acid double helix in two different geometrical representations. This formalism is used to determine the parameters of the base-pair orientations of the curved structure in terms of wedge as well as cylindrical parameters. In the structural model presented here, the curvature of the helix axis results from an alternation of two different DNA structures in which the base-pairs possess different angles with the helix axis ('cylinder tilt'). Resulting from geometric restraints, a negative cylinder tilt angle correlates strongly with the closing of the minor groove ('wedge roll'). The blocks with different structure are not exactly coincident with the dA_5-blocks and the B-DNA stretches. Within the dA_5 block, base-pair tilt and wedge roll adopt large values which proceed into the 3′ flanking B-DNA sequence by about one base-pair. These properties of the structure calculated here are discussed in terms of different models explaining DNA curvature.

Key words: Molecular mechanics, symmetry constraints, structure parameters of DNA, poly(dA) · poly(dT), DNA curvature

I. Introduction

Recent experimental evidence indicates that poly(dA) · poly(dT) has a B′-form structure different from normal B-DNA. In aqueous solutions, poly(dA) · poly(dT) has a helical repeat of 10.1 base-pairs per turn compared to 10.5 for normal B-DNA (Peck and Wang 1981; Rhodes and Klug 1981; Prunell et al. 1984; Strauss et al. 1981). Raman measurements indicate that at salt concentrations lower than 0.1 *M* NaCl and below room temperature, poly(dA) · poly(dT) has a structure different from B-form (Thomas and Peticolas 1983; Jolles et al. 1985; Wartell and Harrell 1986). Under the same conditions, linear dichroism data suggest that poly(dA) · poly(dT) has a large tilt angle and a large propeller twist (Edmondson and Johnson 1985). At higher salt concentrations and elevated temperatures poly(dA) · poly(dT) might tend to adopt the B-form structure. This would explain that the classical B-form for poly(dA) · poly(dT) was found by a 500 MHz NMR study using one dimensional NOE (Sarma et al. 1985) and a recent Raman study (Katahira et al. 1986) which were performed at 30 °C and at salt concentrations higher than 0.1 *M* NaCl. Arnott et al. (1983) have suggested an A-form-B-form hybrid structure for poly(dA) · poly(dT) ('heteronomous DNA') to explain X-ray diffraction data. They propose that the dA-strand adopts an A-like struc-

* To whom offprint requests should be sent

ture while the dT-strand has a structure similar to the B-form. Recent theoretical energy minimization studies (Rao and Kollman 1985) also suggest a hybrid structure; however, they propose that the dA-strand is in the B-form and the dT-strand is in the A-form.

In addition, some unique features of poly(dA) · poly(dT) seem to be recognized by proteins. Immunoglobulins raised in rabbits specifically bind poly(dA) · poly(dT) (Diekmann and Zarling, submitted) while normal B-DNA is not recognized. The restriction endonuclease *Fok I* cleaves poly(dA) · poly(dT) cloned into plasmids with a low rate and only up to about 20% compared to normal B-form DNA sequences (Diekmann and Robert-Nicoud, in preparation). Also histone octamers do not bind to relatively long linear poly(dA) · poly(dT) polymers (Rhodes 1979; Simpson and Kunzler 1979) or to DNA molecules with sufficiently long tracts of poly(dA) · poly(dT) (Kunkel and Martinson 1981; Prunell 1982).

In this paper we present molecular mechanics calculations of the oligonucleotide $(dA)_{12} \cdot (dT)_{12}$ using the AMBER-software-package (Weiner and Kollman 1981) which was modified to account for the influence of counter-ions in the electrostatic contributions (Kitzing 1986). Our results support the A-form-B-form hybrid structure suggested by Arnott et al. (1983) with only slight modifications.

A restriction fragment of a minicircle of mitochondrial DNA of *Leishmania tarantolae* (kinetoplast DNA) exhibits unusually low electrophoretic mobility in acrylamide gels (Marini et al. 1982, 1983; Wu and Crothers 1984; Diekmann and Wang 1985) which is thought to result from a sequence-dependent curvature of the DNA fragment (Trifonov and Sussman 1980, for a review see Trifonov 1985 and Diekmann 1987b). The sequence responsible for this effect is a stretch of (C)AAAAA(T) repeated several times in phase with the helix turn (Wu and Crothers 1984; Diekmann and Wang 1985). The phasing requirement of this effect was rigorously tested by Hagerman (1985). Circularization (Ulanovsky et al. 1986) and electron microscopy (Griffith et al. 1986) experiments have shown that DNA fragments containing such sequences are indeed curved. Koo et al. (1986) and Diekmann (1986, 1987a) showed that the presence of dA_n stretches with $n \geqq 4$ repeated in phase with the helix turn is sufficient for strongly altered electrophoretic mobility of the DNA fragment while only a very small effect was observed for $n < 4$. This length dependence was interpreted by assuming that the dA_n tracts with $n \geqq 4$ adopt a B′-form structure which apparently cannot be formed for $n < 4$ in the analyzed sequences (Diekmann 1986).

The kinetoplast sequences containing repeated stretches of dA_{4-6} have a helical repeat of 10.4 base-pairs per turn (Diekmann and Wang 1985), in between the values for poly(dA) · poly(dT) and B-form DNA. In addition, the kinetoplast DNA fragment injected into rabbits, gave rise to immunoglobulins which specifically recognize poly(dA) · poly(dT) (Diekmann and Zarling, submitted). These results also support the contention that the short tracts of dA_n with $n \geqq 4$ adopt a B′-form structure, possibly the structure of poly(dA) · poly(dT). The structure alternation between B- and B′-form DNA might be the origin of DNA curvature (Wu and Crothers 1984; Levene and Crothers 1983; Diekmann 1986; Koo et al. 1986).

In this paper a model of a curved DNA fragment is constructed by joining short tracts of dA_5 in a B′-form structure calculated for poly(dA) · poly(dT) with the sequence d(GCTCG) in normal B-form. The oligonucleotides d(GCTCGAAAAA) and d(TTTTTCGAGC) had been synthesized and cloned. DNA fragments containing several repeats of this sequence display abnormal gel electrophoretic mobility (Diekmann 1986). The calculated A-form-B-form hybrid structure of the dA_5 tracts is found to have a considerable base-pair cylinder tilt. These blocks, when joined to the B-form (which is roughly without tilt) give rise to an effective curvature of the helix axis. However, the structure junctions do not coincide with the sequence blocks; the structure junctions are shifted in the 3′ direction with respect to the dA_5-block.

II. Molecular mechanics calculations

For the molecular mechanics calculations the basic AMBER-software package (Weiner and Kollman 1981) was used. A new optimization algorithm was introduced (Kitzing 1986). The electrostatic interaction is considered differently in the new program. In addition, molecular symmetries can now be constrained to reduce some optimization problems (see below).

In the original AMBER program the electrostatic interaction was introduced by a Coulomb potential. The dielectric constant was considered as either constant or proportional to the atomic distance. To include salt effects, here the electrostatic interaction is calculated by a Debye-Hückel potential. A given atom of the molecule is assumed to be surrounded either by other atoms of the same molecule or by solvent (water) molecules. The 'effective electrostatic diameter' of an atom is set to a certain value σ (see below). In the case of ions in aqueous solution, it is the diameter of the ion plus the first hydration shell.

Within this shell a Coulomb potential with a dielectric constant $\varepsilon = 4$ is used. Between atoms with a distance larger then σ, a Debye-Hückel potential is assumed with a dielectric constant of 80. The Debye-Hückel potential is introduced to take into account electrostatic long-range screening effects. This approach seems to be more appropriate than using a dielectric constant proportional to the distance of the respective atoms (Weiner and Kollman 1981; Warshel and Levitt 1976).

Debye-Hückel-Potential:

$$E_{ij\,D} = Q_i Q_j/(\varepsilon R_{ij}) \cdot e^{(\lambda(\sigma - R_{ij}))}/(1+\lambda\sigma)\,, \qquad (1.1)$$

where σ is the effective diameter of the hydrated counter-ions and ε (= 80) is the dielectric constant.

Debye-Hückel screening parameter:

$$\lambda = (4\pi/(\varepsilon k_B T)\,(\varrho_1 \cdot q_1^2 + \varrho_2 \cdot q_2^2))^{1/2}\,,$$

where ϱ_a is the ion number density in solution ($a \in \{1,2\}$), q_a is the charge of ion type a, k_B is Boltzmann's constant and T the absolute temperature.

The calculations were performed at rather low ion concentrations (100 mM). Considering the electrostatic interactions in a more complex way to account for the electrostatic effects at high ionic strengths (Soumpasis 1984; Kitzing 1986), the DNA B-Z structure transition at about 3 M NaCl could be correctly calculated (Klement et al. 1985). When these more complex electrostatic potentials were used in our calculations at low ion concentrations, only a negligible difference for the structural energy was obtained. Therefore, these complex potentials were not introduced here.

During the optimization, the following values were used: following Soumpasis (1984) $\sigma = 4.5$ Å, for the salt concentration in solution, $\varrho_1 = \varrho_2 = 0.1$ Mol/l, monovalent ions with $q_1 = +1.0$, $q_2 = -1.0$, and $T = 300$ K.

For the potentials of the non-bonded interactions (Weiner et al. 1984) distances are defined ('cut-offs') up to which these interactions are taken into account. In this paper the following cut-off radii were used: For hydrogen bonding $R_{hb} = 4.5$ Å, for van-der-Waals interactions $R_{vdw} = 10$ Å, and for electrostatic interactions $R_{el} = 25$ Å (Brooks et al. 1983; Levitt 1983; Remerie et al. 1985; Singh et al. 1985; Kitzing 1986).

Since the optimization procedure requires a smooth and continuously differentiable potential, all boundaries between different potential types and cut-offs are convoluted with a fifth order polynomial (Kitzing 1986). The Coulomb and the Debye-Hückel potential is smoothly connected in the interval $\sigma = 4.5$ Å $< r < 6.8$ Å. The smooth cut-off for the hydrogen bond interaction was placed in the interval 3.5 Å $< r < R_{hb} = 4.5$ Å, for the van-der-Waals interaction at 7.5 Å $< r < R_{vdw} = 10.0$ Å and for the electrostatic interaction at 20.0 Å $< r < R_{el} = 25.0$ Å.

Due to the existence of local minima, the structure of large RNA and DNA molecules tend to loose the helical symmetry of the start coordinates during optimization. Constraining helical symmetries by means of a penalty function E_s with the appropriate parameters for rise and angle per base-pair reduces this effect. At the last stage of optimization the constraints are omitted which leads to the final molecular structure with only slightly shifted minima due to end effects. We always found that the practice of first constraining helical symmetries leads to lower final molecular energies than having calculated without these constraints.

If the atom pair i and j with the coordinates x_i and x_j is supposed to satisfy a specific symmetry, then the symmetry constraining potential is given by:

$$E_s = \tfrac{1}{2} k_s \,|\, \boldsymbol{x}_i - \boldsymbol{T}\boldsymbol{x}_j + \boldsymbol{r}_t \,| \qquad (1.2)$$

using

$$\boldsymbol{T} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for the rotation around the z-axis. $\boldsymbol{r}_t = (0, 0, z_t)$ gives the translation along the z-axis.

During the calculation all vectors (coordinates, gradients, and search direction) are stored with single precision, although all calculations are done with double precision (Kitzing 1986). An optimization is assumed to have converged, if the norm of the gradient divided by $\sqrt{3n}$ is less then 0.05 kcal/(Mol Å) where n is the number of atoms in the molecule. The Shanno method of conjugated gradients (Shanno 1978) is used for optimization. All calculations were performed using the facilities (SPERRY 1108) of the Gesellschaft für wissenschaftliche Datenverarbeitung, Göttingen.

III. Results and discussion

III.1. Calculation of a structural model for $dA_{12} \cdot dT_{12}$

The structural energy of slightly more than one helical turn of the homopolymer $(dA)_n \cdot (dT)_n$, $n = 12$, was minimized. Longer sequence stretches would lead to extremely long computation times while the structure of shorter sequences might be influenced too strongly by end effects. Most of the published double helical DNA structures with given coordinates were used as start values to optimize the structure of the sequence $(dA)_{12} \cdot (dT)_{12}$: A-DNA and B-DNA (Arnott and Hukins 1972), heteronomous

DNA (Arnott et al. 1983), and right- and left-handed B-DNA proposed by Sasisekharan et al. (1983). In addition, two-modified versions of B-DNA (Arnott and Hukins 1972) with an additional tilt of $-20°$ and $+20°$ were used.

The resulting optimized structure is expected to have not only a low structural energy (see Table 1) but also to display the experimentally known properties of poly(dA) · poly(dT). Helical repeat measurements of poly(dA) · poly(dT) in solution (Peck and Wang 1981; Strauss et al. 1981) as well as bound on a surface (Rhodes and Klug 1981) determined the twist angle per base-pair to be about 36°. Measuring the linear dichroism, Edmondson and Johnson (1985) found that poly(dA) · poly(dT) has a large tilt angle and propeller twist. Raman spectra (Thomas and Peticolas 1983; Jolles et al. 1985; Wartell and Harrell 1986) suggest that poly(dA) · poly(dT) has a B′-form structure different from B-DNA. Analysing fibre diffraction data Arnott et al. (1983) suggested a hybrid structure for poly(dA) · poly(dT) with the dA-strand in an A-form and the dT-strand in a B-form. Antibodies raised against poly(dA) · poly(dT) recognize the poly(dA)-strand of that double helix (Diekmann and Zarling, submitted). This indicates that the poly(dA)-strand might have a structure different from normal B-DNA.

Only one of the seven optimized structures is in agreement with all of the known structural features determined experimentally (see Table 2). The start coordinates given by Arnott et al. (1983) yield an A-form-B-form hybrid structure with a rise of 3.4 Å, a helical twist angle of 36.1° per base-pair, and with a propeller twist angle of $+15.0°$. The cylinder tilt angle equals $-7°$. The wedge roll is about $-4.5°$ (the base-pairs close towards the minor groove) and has the opposite sign as that for standard A-DNA (see for example Arnott and Hukins 1972). The exact definitions of the cylinder and the wedge double helical parameters are given in the Appendix.

This value of wedge roll can be understood in terms of the structure of adenine. The amino-group at the C_6 of adenine is close to the global helix axis. These groups on neighbouring adenines repel each

Table 1. The different components of the energies (in kcal/Mol) of the studied structures of $dA_{12} \cdot dT_{12}$

	E_{tot}	E_b	E_{ang}	E_{dih}	E_{vdw}	E_{el}	E_{hb}	E_{14vdw}	E_{14el}	rms
A	−932.3	10.63	230.3	335.5	−427.1	167.3	−11.83	118.3	−1,355	0.048
B_1	−931.4	8.25	211.7	400.1	−408.6	99.1	−11.77	123.8	−1,354	0.057
AB	−932.5	8.66	221.3	366.2	−401.6	120.7	−11.99	117.2	−1,353	0.035
B_2	−948.4	8.18	217.2	400.5	−433.8	98.5	−11.26	124.4	−1,352	0.037
C	−990.4	8.33	269.1	313.5	−419.4	38.5	−12.03	147.0	−1,335	0.036
B_3	−938.6	8.34	218.0	398.6	−424.6	101.6	−10.85	123.5	−1,353	0.037
B_4	−927.4	8.44	209.7	404.3	−405.9	98.4	−11.97	123.4	−1,354	0.034

The different structures are named by their start coordinates:

A: A-DNA (Arnott and Hukins 1972)
B_1: B-DNA (Arnott and Hukins 1972)
AB: heteronomous DNA (Arnott et al. 1983)
B_2: right handed B-DNA (Sasisekharan et al. 1983)
C: left handed B-DNA (Sasisekharan et al. 1983), as proposed by Premilat and Albiser (1984), the left handed form of B-DNA is called 'C-DNA'.
B_3: B-DNA (Arnott and Hukins 1972) with an additional cylinder tilt $-20°$
B_4: B-DNA (Arnott and Hukins 1972) with an additional cylinder tilt $+20°$

The total energy E_{tot} is the sum of the different energy terms: bond energy E_b, bond angle energy E_{ang}, torsion energy E_{dih}, van-der-Waals energy E_{vdw}, electrostatic energy E_{el}, hydrogen bond energy E_{hb}, the van-der-Waals energy E_{14vdw} and the electrostatic energy E_{14el} between the first and the fourth atom in dihedral angles. The last term, rms, gives the standard deviation of the gradient from zero at the end of the optimization (in kcal/(Mol Å)). At this last stage of optimization no constraints were imposed on the structures.

Table 2. Some structural properties of the studied sequence $dA_{12} \cdot dT_{12}$

	Hel. twist [°]	Prop. twist [°]	Cyl. tilt [°]	Cyl. roll [°]	Wedge tilt [°]	Wedge roll [°]
A-DNA	32.0	1.0	9.0	−3.0	1.5	5.5
B-DNA's	>40.0	22.0	−10.0 to 4.0	−5.0 to 5.0	−4.0 to 3.0	−10.0 to 3.0
C-DNA	−41.0	−14.0	−2.0	0.5	−0.5	−1.0
AB-DNA	36.1	15.0	−7.0	−1.5	1.0	−4.5

For the definitions of the parameters see the Appendix

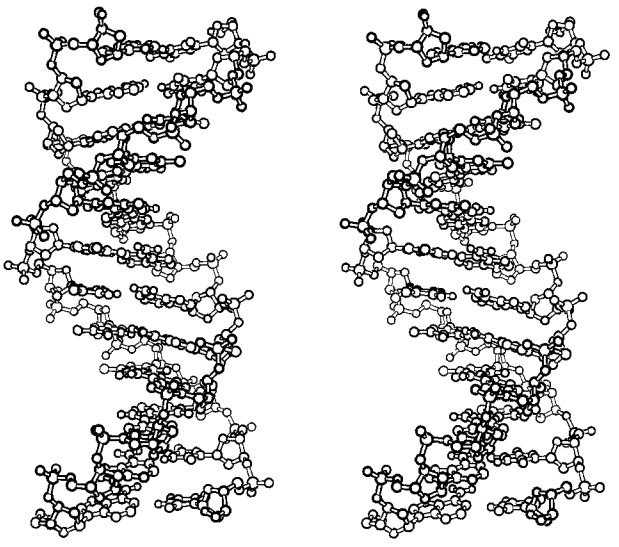

Fig. 1. The optimized structure of $dA_{12} \cdot dT_{12}$. The dA-strand adopts the $C_{3'}$-endo, the dT-strand the $C_{2'}$-endo sugar pucker. The cylinder tilt angle is $-7°$ equivalent to a wedge roll angle of $-4.5°$ (for the definitions of wedge and cylinder parameters see Appendix)

other. In addition and in contrast to guanine, the C_2 atom of adenine has no bulky group. Consequently, the adenines can slant into this 'niche' and close towards the minor groove. The base-pairs open with about 1° towards the thymidines (wedge tilt angle). The distance between the long base-pair-axis and cylinder twist axis is 1.6 Å. The base-pairs are moved with about 0.4 Å in the direction of the thymidines. Thus the adenines lie more in the centre of the helix while the thymidines lie more outside. The ribose of the dA-strand has $C_{3'}$-endo while the dT-strand has $C_{2'}$-endo sugar pucker. The resulting structure is similar to that suggested by Arnott et al. (1983). It is in agreement with the immunological data of Diekmann and Zarling (submitted). The structure is displayed in Fig. 1. Its coordinates can be obtained from the authors.

The other six structures differ in the helical twist angle from experimental data (see Table 2). The difference between the wedge twist and the cylinder twist angle is small in general (about 0.1°). Therefore, only one value for the twist angle is given in Table 2. In contrast to Haran et al. (1984) and Tung and Harvey (1986), using the molecular mechanics parameters of Weiner et al. (1984) for the energy function our calculations yield the correct sign and a reasonable value for the propeller twist. Although the standard conformational angles (compare IUPAC-IUB Notation 1983) differ up to 20°, the overall difference between 'heteronomous' DNA proposed by Arnott et al. (1983) and the optimized A-form-B-form DNA presented here is small. The

Table 3. The minimal rms-deviations between the studied structures of $dA_{12} \cdot dT_{12}$

	A	B_1	AB	B_2	C	B_3	B_4
A	1.45	5.13	4.37	6.62	9.69	5.97	5.15
B_1		2.62	2.65	1.98	8.62	1.10	0.57
AB			1.21	3.53	8.70	3.20	2.67
B_2				3.23	8.58	1.30	2.03
C					2.56	8.57	8.63
B_3						2.98	1.22
B_4							2.63

Table 3 displays the minimal rms-deviation (in Å) between the studied structures of the sequence $dA_{12} \cdot dT_{12}$. For the names of the structures see the legend of Table 1.

In the diagonal the deviations between the start and the optimized structures are given. The off-diagonal numbers are the deviations between the different optimized structures. The large values of the C-DNA with respect to the values of the other structures result from the left handedness of this helix

rms-deviation between the start and the optimized coordinates (1.21 Å) is lowest for this DNA-type compared to the other structures (see Table 3).

In contrast to our result, Rao and Kollman (1985) suggested a hybrid structure for poly(dA) · poly(dT) with the dA-strand in a B-form and the dT-strand in an A-form. Using the AMBER program they optimized $dA_6 \cdot dT_6$ in standard B-DNA but forcing the sugar puckering to flip from $C_{2'}$-endo to $C_{3'}$-endo by means of a constraining potential. They calculated the total energy of $(dA)_6 \cdot (dT)_6$ in different structures; in standard B-DNA with all adenine riboses having $C_{3'}$-endo or, alternatively, with all thymine riboses having $C_{3'}$-endo (the respective opposite strands remained with $C_{2'}$-endo sugar pucker). The model with dT in $C_{3'}$-endo and dA in $C_{2'}$-endo was found to have the lowest total energy. These structures with different sugar puckering in the two strands hardly differ from the starting B-DNA structure with respect to base stacking and overall shape. This might be due to the use of locally converging optimization methods; the strong base stacking forces keep the structure rigid. While the ribose and backbone structure is optimized, in most cases the reorientation of the stacking pattern will be prevented. Thus, it might be possible that these calculated structures are not in the global energy minimum.

III.2. Construction and calculation of a structural model for $d(GCTCGAAAAA)_4 \cdot d(TTTTTCGAGC)_4$

The large number of degrees of freedom of a long DNA molecule makes it difficult to optimize its structure by simply starting with standard B-DNA

coordinates (e.g. Arnott and Hukins 1972) and using a standard optimization algorithm (i.e. conjugated gradients). Of the many different modes of deformation those with high force constants (for example bond stretching) are far better optimized than modes with very low force constants (for example bending of the helix axis). By starting with a structure which requires the reorientation of large parts of the molecule, only local distortions of the structure are produced having an unfavourable energy minimum. Thus, the structure of interest, $d[(GCTCGAAAAA)_4] \cdot d[(TTTTTCGAGC)_4]$ was constructed from three sequence elements with shorter periodicity, each of them optimized independently as a first stage of optimization.

The sequence, $(dA)_5 \cdot (dT)_5$, was given the optimized B′-form structure of $dA_{12} \cdot dT_{12}$ presented above. The B-DNA part of the structure is obtained from the double-stranded sequence $d[C(GAGCTC)_4G]$. It contains the stretches dGCTCG and dCGAGC. The third sequence element was designed to display the junction between B-DNA and the B′-DNA: $d[A(GCGAA)_4G] \cdot d[C(TTCGC)_4T]$. These two last sequences are chosen such that the structural distortions due to different properties of the sequence elements are nearly compensated within one turn of the helix. Thus, during the optimization process only local reorientations of atoms are required. Because of the sequence symmetry, in both cases the global helix axis is nearly straight. The calculation of the junction structure was started using optimized standard B-DNA coordinates (Arnott and Hukins 1972) in ideal helix geometry for the dGCs, and B′-DNA coordinates of the hybrid structure presented above for the dAA · dTT parts of the helix.

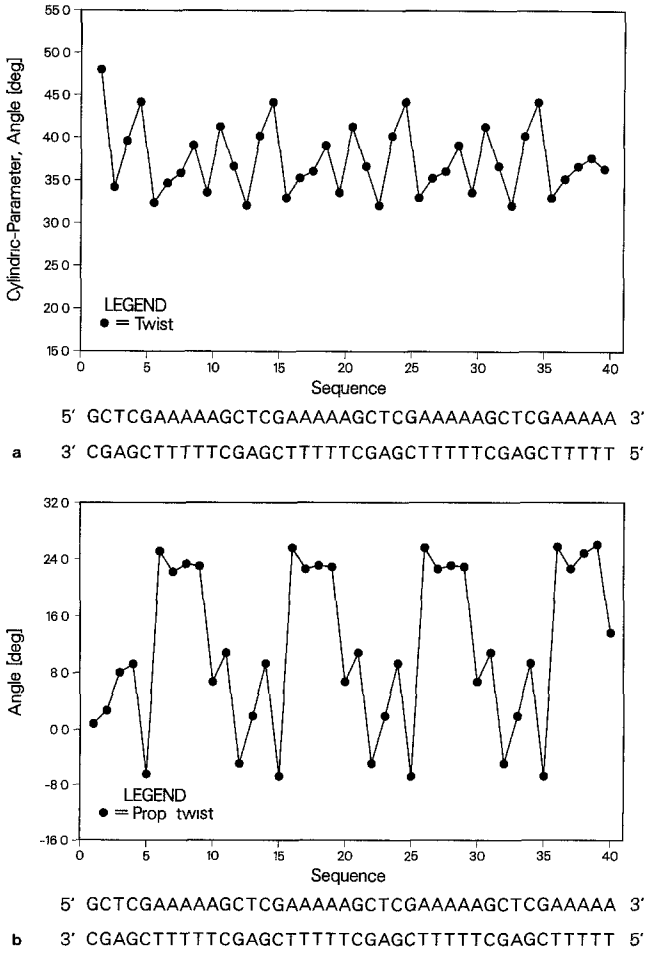

Fig. 2. **a** The cylinder twist per base-pair for the sequence indicated. **b** The propeller twist per base-pair for the sequence indicated

To get a reasonable starting structure for the sequence with mixed DNA blocks, the stacking of the bases and the orientation of the phosphate groups were monitored by means of an interactive program. During the optimization, helical constraints (described in Methods) of 20 kcal mol^{-1} $Å^{-2}$ were used to force the molecule $d[C(GAGCTC)_4G]$ to have 20.4 Å total rise and 216° total twist per 6 base-pairs and $d[A(GCGAA)_4G] \cdot d[C(TTCGC)_4T]$ to have 17 Å and 180° per 5 base-pairs (corresponding to mean values of 3.4 Å and 36° for the rise and twist per base-pair, respectively).

The starting structure of the sequence $d(GCTCGAAAAA)_2 \cdot d(TTTTTCGAGC)_2$ was built from the corresponding parts of the three sequences $d[C(GAGCTC)_4G]$ (1), $d[A(GCGAA)_4G] \cdot d[C(TTCGC)_4T]$ (2) and $dA_{12} \cdot dT_{12}$ (3) in the following way:

```
GCTCGAAAAAGCTCGAAAAA
11111      11111
    22    22    22
     33333      33333

* TTTTTCGAGCTTTTTCGAGC
      11111      11111
    22    22    22
 33333      33333
```

The numbers typed under the sequence indicate which of the three basic sequences were used for the construction. The parts were linked by maximum overlap of the respective purines which mainly determine the stacking interaction of DNA molecules (Bubienko et al. 1983). After removing doubly occurring atoms, this structure was optimized first with a constraint on every atom position; all atoms were bound to their start position by a harmonic potential. This fixes the overall geometry but allows a partial relaxation of distorted bond distances and bond angles. Finally, these constraints were removed and the structure was further optimized until

convergence. This procedure yields a structure which was used to construct d[(GCTCGAAAAA)$_4$] · d[(TTTTTCGAGC)$_4$] by simply tripling the middle part d(AAAAAGCTCG) of d[(GCTCGAAAAA)$_2$]. This sequence again was optimized. The result is displayed in Fig. 4a. The riboses of the dA-strand and the dT-strand still have $C_{3'}$-endo and $C_{2'}$-endo sugar pucker, respectively. The coordinates can be obtained from the authors.

Several structure parameters of this molecule scatter around mean values with no obvious sequence dependence, for example the rise per base-pair in cylinder parametrisation (3.3 Å with extremes at 3.0 Å and 3.5 Å) and the cylinder twist angle (see Fig. 2a). In contrast, the propeller twist angle adopts large positive values for the dA_5 block (with the exception for the last dA). For the B-form sequence stretches it scatters around zero values (see Fig. 2b). Figure 3a and b display the roll and tilt angles of the molecule in wedge angle and cylinder symmetry representation. Note that the roll and tilt exchange their contribution when changing from wedge to cylinder parameters. As an empirical estimate for larger values their relation is given by:

$$\gamma_c \cong \sqrt{2}\,\alpha_w\,, \quad \alpha_c \cong -\sqrt{2}\,\gamma_w\,.$$

Thus, base-pairs with a large wedge roll have a large tilt with respect to a local cylinder twist axis and vice versa. Because of the different sequences and the special properties of the B′-form DNA, only those wedge roll angles can be compared with crystal data of Fratini et al. (1982) which are in B-DNA like conformations. As with the models of Dickerson (1983) and Tung and Harvey (1986), the angles of the sequence analyzed here are in rough agreement with the crystal data.

Figure 3b displays the cylinder tilt and roll angles of the molecule. The figure indicates that the cylinder tilt angle adopts roughly only two different states. For 5 base-pairs the tilt is slightly positive while for the next 5 base-pairs it is ca. −15°. Thus, this molecule consists of alternating blocks having nearly zero and highly negative tilt, respectively. Since cylinder tilt is related to wedge roll (see above), in the wedge representation this alternation of blocks with large wedge roll and tilt is also found (see Fig. 3a). These structure blocks have their origin in the sequence variation but, surprisingly, these blocks of large cylinder tilt and wedge roll do not precisely coincide with the dA_5 sequences. The tilt blocks are shifted from 5′ to 3′ relative to the sequence elements (B-DNA and B′-DNA) of the first strand. The cylinder tilt (wedge roll) is built-up within the dA_5 block although already the first dA in this block adopts the B′-form structure as indicated by the propeller twist (see Fig. 2b). The

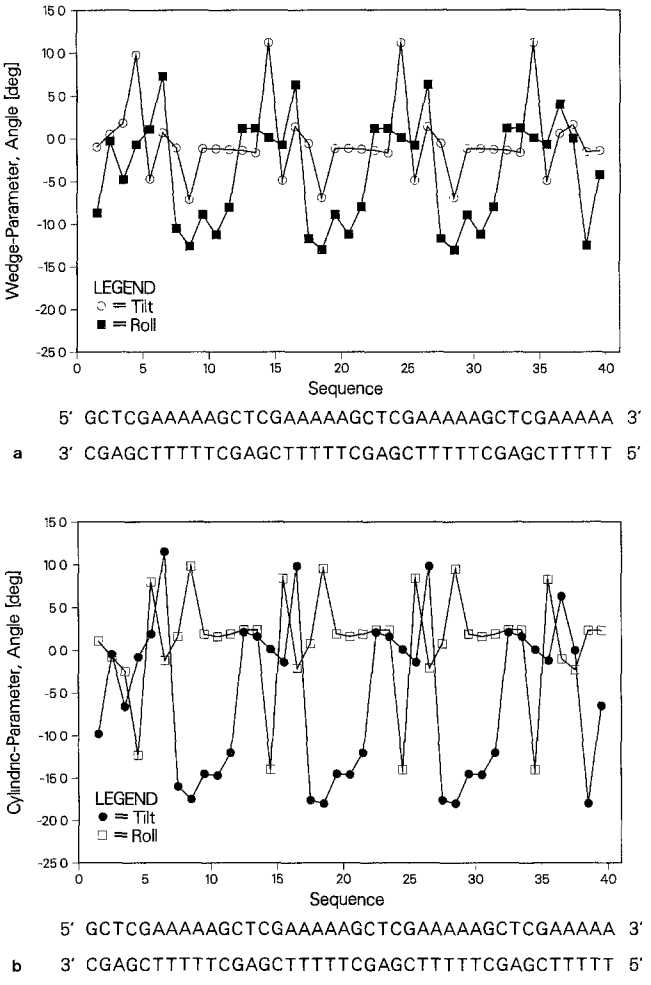

Fig. 3. a The wedge tilt angle α_w and the wedge roll angle γ_w of the sequence indicated. **b** The cylinder tilt angle α_c and the cylinder roll angle γ_c of the sequence indicated. The blocks of highly negative cylinder tilt are shifted approximately one base-pair in the 3′-direction of the first strand with respect to the sequence blocks

shift into the B′-form DNA block is found to be larger than the shift into the 3′ following B-DNA block. In general, this shift implies an anisotropic sequence-directed structure influence which is not in agreement with a simple nearest neighbour model on base-pair structure (Trifonov 1985, see below). The shift might not be related only to the B′-form DNA structure but to DNA structure in general and could be a purine stacking effect. The structure shift is expected to be sequence dependent. Such a sequence dependence would easily explain the small but observable effect of the dA_5-3′-neighbouring base-pair on the gel migration anomaly (Koo et al. 1986).

Figure 4a shows a stereo picture of the whole molecule. The long base-pair and the cylinder twist axes are displayed separately (Fig. 4b). Additionally, groups of 5 cylinder twist axes are connected to blocks representing adjacent base-pairs with similar

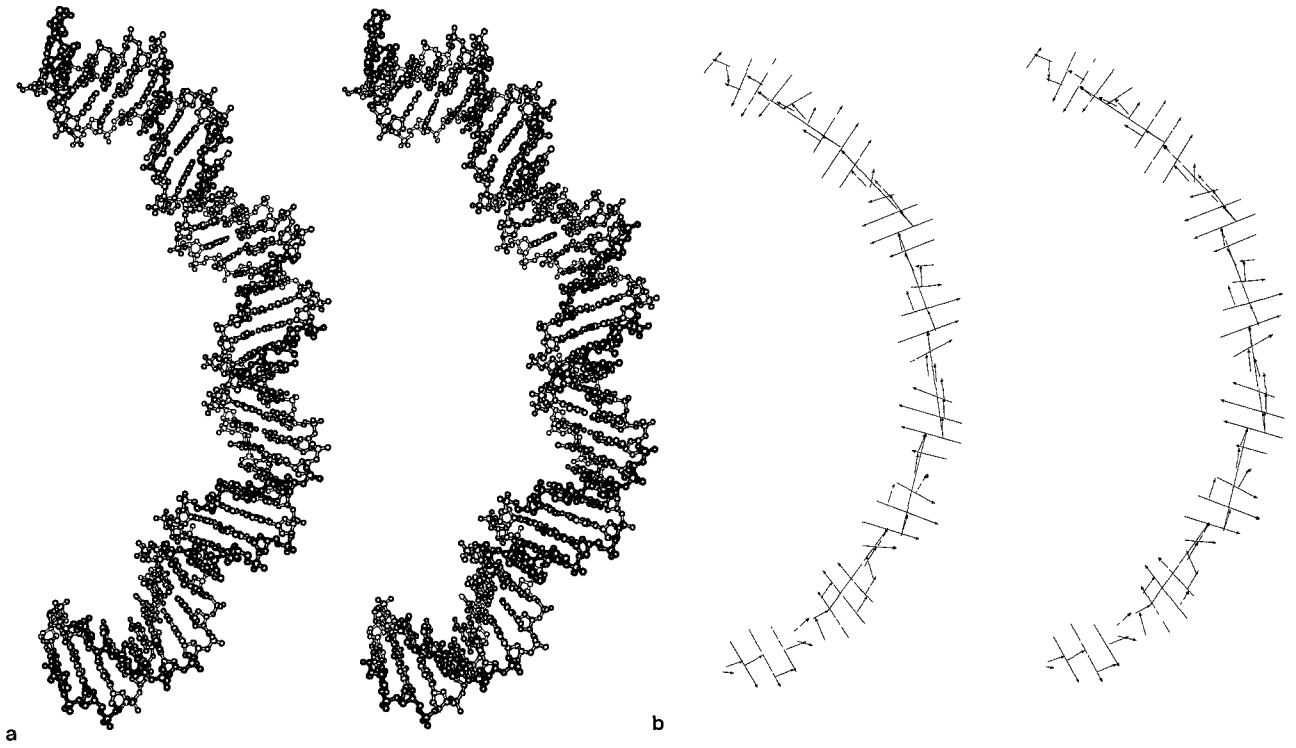

Fig. 4. **a** The optimized structure of $d[(GCTCGAAAAA)_4] \cdot d[(TTTTTCGAGC)_4]$. **b** The long base-pair axes ***a*** and the cylinder twist axes t_c (*short arrows*) of the molecule $d[(5'\text{-}GCTCGAAAAA\text{-}3')_4] \cdot d[(5'\text{-}TTTTTCGAGC\text{-}3')_4]$ (first strand * second strand). The t_c vectors show into the 3′ direction of the first strand, the vectors ***a*** from the first to the second strand. Groups of nearly parallel vectors t_c are indicated by *long arrows*

cylinder twist axes. Figure 4 b shows that the cylinder twist axes within each group are fairly parallel.

For a circularized DNA fragment Drew and Travers (1985) found that minor grooves of (dA+dT) rich sequences face to the inside of the curvature while the minor grooves of (dG+dC) rich sequences face to the outside. This can be rationalized by the sequence dependent structure parameters obtained from our calculation. The base planes of dAA and dTT dinucleotides close towards the minor groove whereas dG/dC rich sequences slightly open. For the curved molecule calculated here we find that the minor groove of the sequence blocks with large cylinder tilt angles is placed at the inside of the curvature. Of the $dA_5 \cdot dT_5$ sequences mainly the dT bases are inside. Thus, the properties of this structure are in agreement with the results of Drew and Travers (1985).

A DNA double helix with alternating structures was discussed earlier by Selsing et al. (1979) in a different context. They suggested that due to environmental changes or bent into a circle the DNA locally flipps into an A-form structure. The bent DNA then consists of straight stretches in standard B-DNA (nearly without tilt) alternating with standard A-DNA (with large positive tilt) with localized bending angles at the junctions of the structures.

III.3. Interpretation of the structure in terms of models of curved DNA

Several models have been proposed to explain the large sequence-directed DNA curvature discussed here. Models attributing the curvature only to the purine-pyrimidine sequence (Hagerman 1984; Zhurkin 1985) contradict experimental results (Diekmann 1986; Hagermann 1985, 1986; Koo et al. 1986). Two other models are able to explain the available experimental data; the 'nearest neighbour model' and the 'distal interaction model'.

Representing the curved structure in terms of the wedge parameters, blocks with alternating large and small wedge roll angles are observed (see Fig. 3 a). Thus, the observed curvature can be correlated with the large wedge angle of the dAA dinucleotides. The 'nearest neighbour model' explains DNA curvature by this large dAA/dTT wedge angle ('wedge model', Trifonov 1985). Following this model, wedge angles are assumed to be different for every dinucleotide and dependent only on nearest neigh-

bours, but otherwise sequence independent. The experimentally observed curvature is predicted to increase with the number of dAA dinucleotides until about dA_6 (as found by Koo et al. 1986); the wedge angles of additional dAA dinucleotides would cancel each other since they are then on opposite sides of the helix. The very small (however, clearly detectable) gel migration anomaly of dA_3 compared to dA_4 has been used as an argument against this model. However, the gel matrix might be relatively insensitive to small degrees of curvature.

When base-pairs with large wedge roll angles are consecutively packed into a double helix, a structure with a large cylinder tilt angle will be obtained (Prunell et al. 1984). From the helical repeat of poly(dA) · poly(dT) (10.1 bp/turn compared to 10.5 bp/turn for normal DNA) Prunell et al. (1984) calculated the wedge angle of dAA equal to 11° in agreement with our data (see Fig. 3a). The tilted structure appears different than classical B-form DNA. Thus, the large wedge angle of the dAA dinucleotide is suggested to lead to a B′-form structure for longer stretches of dAs. This becomes obvious when the curved DNA structure is discussed in terms of the cylinder parameters. The dA_5 block has a large cylinder tilt in contrast to the B-form DNA block with nearly no cylinder tilt (see Fig. 3b). Thus, according to this view, curvature is due to the alternation of sequence blocks having structures with different cylinder tilt. While individual sequence stretches might be straight, the tilt angle at the junction of the two structures deflects the DNA helix axis (the two block helix axes include a non-zero angle, 'junction bending model', Wu and Crothers 1984; Levene and Crothers 1983). The two different geometrical descriptions of the local DNA structure lead to two different views explaining curvature. Since it is one and the same structure, the two views are equivalent.

The view of the curved DNA in terms of alternating structures leads to the 'distal interaction model'. Following the distal interaction model, some wedge angles are assumed to be strongly sequence dependent; sequence interactions (for example from dA · dT base-pairs) more distant than the nearest dA neighbours are considered to have a strong influence on the wedge angle of dAA tracts. Thus, the wedge angle of dAA dinucleotides in longer dA blocks would be clearly different from the wedge angle of a single dAA dinucleotide resulting in a B′-form structure and leading to DNA curvature. The sequence dependence, or independence, of the angular structure of the dAA wedges might be used in the future to discriminate between the two models. The distal interaction model can explain all the available data (Diekmann 1986; Koo et al. 1986; Hagerman 1985, 1986). The experimental finding that a strong gel migration anomaly is observed only for dA_n sequences with $n \geqq 4$, is an inherent property of this model. If not only nearest but also next to nearest neighbours influence the local structure, it would change for $n \geqq 4$. The specific structure of longer runs of dAs is not expected to form for $n < 4$.

For the analyzed sequence we observe an additional effect which distinguishes the nearest neighbour from the distal interaction model. The tilt of the base-pairs with respect to the cylinder twist axis and the large wedge roll angles do not coincide with the dA_5 sequences; the structurally modified blocks are shifted in the 3′ direction of the strand containing the dA_5. Thus, the wedge angle of the first dAA dinucleotide in the dA stretch is clearly different from the other dAA wedge angles in the same stretch. Since the nearest neighbour model assumes the wedge angles to be dependent only from next neighbours but otherwise sequence independent, the shift effect contradicts the nearest neighbour model.

Hagerman (1986) found that the sequences $d(^{C}_{G}A_4T_4^{G}_{C})_n$ show a strong gel migration anomaly while $d(^{C}_{G}T_4A_4^{G}_{C})_n$ do not. The data can be explained in terms of wedge angles, if additionally to the large wedge roll angle of the dAA dinucleotide a non-negligible wedge tilt angle is assumed (note that wedge roll and wedge tilt have different phasing properties). Indeed, Ulanovsky and Trifonov (1987) use the data of Hagerman (1986) to estimate the absolute value of the dAA wedge-tilt (8.4°) and the wedge-roll (2.4°) component. The total dAA wedge angle had been estimated from circularization experiments (Ulanovsky et al. 1986). According to the nearest neighbour model, with this set of angles also all other data on gel migration anomaly (Koo et al. 1986; Hagerman 1985; Diekmann 1986) can be explained (Trifonov, personal communication). In the view of alternating structures, a cylinder roll together with the cylinder tilt has to be assumed to obtain the same structural properties as for the wedge considerations.

Thus, we present two equivalent methods to obtain structural parameters for adjacent base-pairs in double helical nucleic acids from atomic coordinates, the wedge angle and the cylindrically symmetrical representation. It is shown that because of geometrical restraints the cylinder tilt angle is strongly correlated to the wedge roll angle. Molecular mechanics calculations were used to build a model for $dA_{12} \cdot dT_{12}$ which displays the known experimental properties of poly(dA) · poly(dT). The structure obtained was combined with independently optimized B-DNA structures to construct and minimize the energy of a curved DNA molecule of the se-

quence $d(GCTCGAAAAA)_4 \cdot d(TTTTTCGAGC)_4$ which experimentally shows a strong gel migration anomaly. Two different models explaining DNA curvature, the 'nearest neighbour' and the 'distal interaction' model, are discussed in relation to the structure obtained. The strong negative cylinder tilt (wedge roll) induced by the dA_5-sequences is shifted in 3′-direction resulting in a structural anisotropy, a non-coincidence of alternated DNA structure with sequence blocks. This shift contradicts the nearest neighbour model.

Note added in proof: Recently, an X-ray fiber diffraction analysis of poly(dA) · poly(dT) was published by Alexeev et al. (1987, Nature 325: 821–823). In their structure the sugar puckers of both strands are close to $C_{2'}$-endo. Their structure is only 'slightly' heteronomous. Other structural features seem to be similar to the structure calculated here. In both structures the base-pairs fold into the minor groove. We consider base-pair stacking properties to be essential for the formation of the B′-form structure of poly(dA) · poly(dT).

Acknowledgements. Helpful discussions with Dr. E. N. Trifonov and Dr. R. Clegg are greatfully acknowledged. We thank Dr. E. N. Trifonov for providing data prior to publication, Dr. H. Kuhn and Dr. M. Kahlweit for support, and Mrs. Z. Wassermann, IE Dupond, for installing the AMBER program in Göttingen. The computations were performed using the SPERRY 1108 of the Gesellschaft für wissenschaftliche Datenverarbeitung, Göttingen. For the molecular plots we used the SCHAKAL plot program written by Dr. E. Keller, Freiburg, modified by R. Klement.

Appendix

Definitions of the structural parameters in cylindrical symmetry and by wedge angles

In this section we introduce definitions for the symmetry parameters describing the structure of double stranded nucleic acids. The definitions outlined by Arnott et al. (1969), Fratini et al. (1982) and Dickerson (1983) are not sufficient to describe the structural features of DNA. They have to be generalized to enable a complete description for double stranded structures including curved DNA.

The base-planes of purines and pyrimidines are defined by the atoms of the six membered rings, with the numbers 1 to 6 proposed by the IUPAC-IUB (1983) notation. Following Fratini et al. (1982) the vector connecting the atoms C_8 of purines and C_6 of pyrimidines are used to define the direction of the long base-pair axis. The 'ends' of a base-pair are defined by the projection of the $C_{1'}$-atoms on the long base-pair-axis. Starting with the base at the 5′-end of one strand, the long base-pair axis vector $\boldsymbol{a}$ points from the base of this strand to the complementary counterpart of the opposite strand. The absolute values of the structural parameters (as well as the values of other parameters derived from them) do not depend on which strand is chosen to be the 'first'.

The short base-pair-axis of a base-pair is defined to be orthogonal to the long base-pair-axis $\boldsymbol{a}$ and the base-pair plane normal vector $\boldsymbol{b}$:

$$\boldsymbol{c}^t = \boldsymbol{a} \times \boldsymbol{b} . \tag{A1}$$

(The following convention is used: The 't' transposes a column vector $\boldsymbol{a}$ into a row vector $\boldsymbol{a}^t$ and vice versa. Thus, $\boldsymbol{a}^t\boldsymbol{b}$ is the scalar product between $\boldsymbol{a}$ and $\boldsymbol{b}$. The cross product between two column vectors results in a row vector).

The signs of $\boldsymbol{b}$ and $\boldsymbol{c}$ are chosen such that $\boldsymbol{b}$ points into the 3′ direction of the first strand and $\boldsymbol{c}$ from the major into the minor groove. Because of these definitions $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are unit vectors:

$$|\boldsymbol{a}| = |\boldsymbol{b}| = |\boldsymbol{c}| = 1 \quad \text{with} \quad (\boldsymbol{a} \times \boldsymbol{b})\,\boldsymbol{c} = 1 . \tag{A2}$$

Since for curved DNA the notion of a global helix axis is meaningless, we introduce definitions for the local helical axis depending only on coordinates of atoms of two adjacent base-pairs. Two types of definitions of the helix axis are discussed. The first focuses on the angles between the planes of two neighbouring base-pairs. The helical twist axis, called 'local twist axis', is normal to the plane which bisects the two base-pair planes. If the calculation starts from the opposite end of the double helix, by this definition only the sign of the local twist axis between two adjacent base-pair planes will change. The parameters obtained relative to this axis are called 'wedge' parameters. The second definition uses the approximate local cylinder symmetry of the double helix. The cylinder twist axis is defined such that base-pair plane 1 is transformed into base-pair plane 2 only by rotation around and translation along this axis.

1. Wedge parameters

The wedge twist axis $\boldsymbol{t}_w$ shall be normal to the mean base-pair plane. $(\boldsymbol{a}_1 + \boldsymbol{a}_2)$ and $(\boldsymbol{c}_1 + \boldsymbol{c}_2)$ span this mean plane, thus $\boldsymbol{t}_w$ satisfys:

$$0 = \boldsymbol{t}_w^t (\boldsymbol{a}_1 + \boldsymbol{a}_2) \quad \text{and} \quad 0 = \boldsymbol{t}_w^t (\boldsymbol{c}_1 + \boldsymbol{c}_2) . \tag{A3}$$

Therefore, the direction of the wedge twist axis can be calculated as

$$\boldsymbol{t}_w^t = ((\boldsymbol{c}_1 + \boldsymbol{c}_2) \times (\boldsymbol{a}_1 + \boldsymbol{a}_2)) / |(\boldsymbol{c}_1 + \boldsymbol{c}_2) \times (\boldsymbol{a}_1 + \boldsymbol{a}_2)| . \tag{A4}$$

The wedge tilt angle α_w is twice the angle between the mean plane of base-pairs and the long base-pair-axis.

$$\sin(\alpha_w/2) = \boldsymbol{t}_w^t \boldsymbol{a}_2 = -\boldsymbol{t}_w^t \boldsymbol{a}_1 . \tag{A5}$$

Similarly, the wedge roll angle γ_w is given by:

$$\sin(\gamma_w/2) = \boldsymbol{t}_w^t \boldsymbol{c}_2 = -\boldsymbol{t}_w^t \boldsymbol{c}_1 . \quad \text{(A6)}$$

α_w and γ_w are positive if the two base-pairs open towards the opposite strand and towards the minor groove, respectively. This implies that α_w changes its sign if the calculation of the parameters starts from the 'second' strand, whereas γ_w does not. The rotation angle between two neighbouring base-pairs with respect to the wedge twist axis is described by the wedge twist angle ω_w. It is given by:

$$\cos(\omega_w) = (\boldsymbol{a}_1 \times \boldsymbol{t}_w)\,(\boldsymbol{a}_2 \times \boldsymbol{t}_w)^t / (|\boldsymbol{a}_1 \times \boldsymbol{t}_w|\,|\boldsymbol{a}_2 \times \boldsymbol{t}_w|) . \quad \text{(A7)}$$

These definitions of base-pair wedge roll and wedge tilt are closely related to those of Fratini et al. (1982) and Dickerson (1983). However, in our presentation the helix axis is calculated from the molecular coordinates and has not to be known a priori. α_w, γ_w, and ω_w are orthogonal to each other. Thus, in our definition the wedge roll and tilt angles are independent of the twist angle in contrast to the definition of Fratini et al. (1982). γ_F and α_F are the roll and tilt angles defined as proposed by Fratini et al. (1982). For a typical twist angle $\omega_w = 36°$ and roll angle $\gamma_F = 5°$ (Fratini et al. 1982), the difference between the two tilt angles α_F and α_w is found to be ca. 2°, the difference in roll angles is calculated to be only 1°. Since the roll angles of both representation are relatively large with smaller differences, they might be comparable; whereas, the tilt angles are smaller with larger differences and thus, might not be comparable in the two representations.

The following six wedge parameters (Fig. 5a and b) uniquely describe the relative orientation of two base-pairs with respect to each other:

- Wedge *twist angle* ω_w: the rotation of the second base-pair around the local twist axis with respect to the first,
- wedge *tilt angle* α_w: the opening of the long axes towards the opposite strand,
- wedge *roll angle* γ_w: the opening of the short axes towards the minor groove,
- wedge *long axis distance* D_w of the local twist axis from the long base-pair-axis,
- wedge *short axis distance* M_w of the local twist axis from the short base-pair-axis,
- wedge *rise per base-pair* R_w: the translation of the second base-pair in the direction of the local twist axis with respect to the first one.

Although these six parameters describe the orientation of two base-pairs uniquely, the choice of these parameters themselves is not unique. One may prefer the parameter base sliding S_w (see also Tung and Harvey 1986) instead of the distance D_w between wedge twist axis $\boldsymbol{t}_w$ and long base-pair-axis $\boldsymbol{a}$:

$$S_w/2 = (D_w - D_w^*) \sin(\omega_w/2) . \quad \text{(A8)}$$

D_w^* has then to be defined as a standard distance between long base-pair-axis and wedge twist axis.

2. *Cylinder parameters*

The ideal cylindrical symmetry of the double helix is mainly used in fibre diffraction analysis (Arnott et al. 1969). Even the structure of a non-linear DNA molecule can locally be described by cylindrical symmetry. Therefore, we discuss such a local axis, here called cylinder twist axis. The cylinder tilt angle α_c is the angle between the plane normal to the cylinder twist axis $\boldsymbol{t}_c$ and the long base-pair-axis $\boldsymbol{a}$. The cylinder roll angle, γ_c, is the angle between the plane normal to cylinder twist axis $\boldsymbol{t}_c$ and the short base-pair-axis $\boldsymbol{c}$:

$$\sin(\alpha_c) = -\boldsymbol{t}_c^t \boldsymbol{a}_1 = -\boldsymbol{t}_c^t \boldsymbol{a}_2 \quad \text{(A9)}$$

$$\sin(\gamma_c) = -\boldsymbol{t}_c^t \boldsymbol{c}_1 = -\boldsymbol{t}_c^t \boldsymbol{c}_2 . \quad \text{(A10)}$$

These angles should not be mistaken for tilt and roll angles according to the wedge angles. In an ideal cylindrically symmetrical DNA molecule the plane of the second base-pair can be obtained by rotation around and translation along a common axis. As a consequence, the orientation given by the cylinder roll and tilt angles with respect to the common axis must be the same for the two base-planes. This geometrical property is used to locally define the cylinder twist axis. From (A9) and (A10) we get:

$$0 = \boldsymbol{t}_c^t (\boldsymbol{a}_1 - \boldsymbol{a}_2) \quad \text{and} \quad 0 = \boldsymbol{t}_c^t (\boldsymbol{c}_1 - \boldsymbol{c}_2) . \quad \text{(A11)}$$

Comparing (A11) with (A3) we see that the cylinder twist axis is normal to the difference of the directions of the long and short axes, respectively, instead of their sums as in the case for the wedge twist axis. Thus, the direction of the cylinder twist axis can be calculated as with (A4) by changing the '+' into '−' signs:

$$\boldsymbol{t}_c^t = ((\boldsymbol{c}_1 - \boldsymbol{c}_2) \times (\boldsymbol{a}_1 - \boldsymbol{a}_2)) / |(\boldsymbol{c}_1 - \boldsymbol{c}_2) \times (\boldsymbol{a}_1 - \boldsymbol{a}_2)| . \quad \text{(A12)}$$

The remaining parameters, the cylinder twist angle ω_c, the long axis D_c between cylinder twist axis and the long base-pair-axis, the short axis distance M_c between cylinder twist axis and the short base-pair axis, and the cylinder rise per base-pair R_c are calculated in the same way as the respective wedge parameters by using the cylinder twist axis instead of the wedge twist axis.

Thus, the following six cylinder parameters (Fig. 5c and b) are the same as given by Arnott et al. (1969), if the cylinder twist axis coincides with the

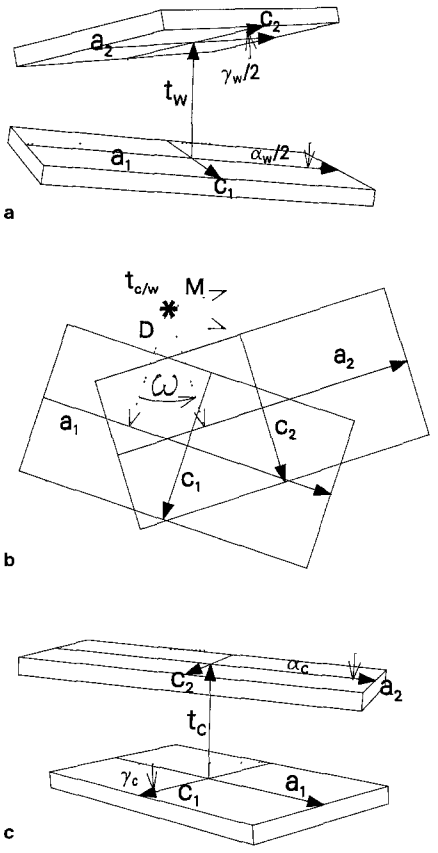

Fig. 5. **a** The wedge tilt angle $\alpha_w/2$, the wedge roll angle $\gamma_w/2$ (i.e. the opening of the base-pair-planes towards the second strand and the minor groove, respectively), the wedge twist axis t_w and the two long (a_1, a_2) and two short (c_1, c_2) base-pair axes are shown for two symbolized neighbouring base-pairs.
b The helical twist angle ω (i.e. the rotation of the second base pair around the helical twist axis with respect to the first) is displayed as well as the long axis distances D and short axis distances M of the helical twist axis from the long (a_1, a_2) and short (c_1, c_2) base-pair axes, respectively. These definitions are identical in the wedge and cylinder parameter representation, however they refer to the cylinder (t_c) and wedge (t_w) twist axis, respectively.
c The cylinder tilt angle α_c, the cylinder roll angle γ_c, the cylinder twist axis t_c and the two long (a_1, a_2) and two short (c_1, c_2) base-pair axes are displayed

z-axis of their system. The possibility that the short base-pair-axis does not intersect the cylinder twist axis has been neglected in their approach.

– cylinder *twist angle* ω_c: the rotation of the second base around the cylinder twist axis with respect to the first,
– cylinder *tilt angle* α_c: the angle between the plane orthogonal to the cylinder twist axis and the long base-pair-axis,
– cylinder *roll angle* γ_c: the angle between the plane orthogonal to the cylinder twist axis and the short base-pair-axis,
– cylinder *long axis distance* D_c of the cylinder twist axis from the long base-pair-axis,
– cylinder *short axis distance* M_c: the distance of the cylinder twist axis from the short base-pair-axis.
– cylinder *rise per base-pair* R_c: the translation of the second base-pair in the direction of the cylinder twist axis with respect to the first.

Both, wedge and cylinder descriptions only use two neighbouring base-pairs for the definition of the corresponding angles and distances. The wedge parameters describe the orientation of the two base-pairs with respect to each other. The wedge roll and tilt are the angles between the planes of the base-pairs and a 'mean plane'. The wedge twist axis will only coincide with the global helix axis of an ideal cylindrically symmetric double helix, if the wedge roll and tilt angles are zero. In contrast, the cylinder parameters reflect the cylindrical symmetry of the double helix. The cylinder twist axis will coincide with the global helix axis of an ideal cylinder symmetrical double helix. The cylinder roll and tilt angles are defined with respect to the cylinder twist axis.

Specific roll and tilt angles will influence the global structure of the helix. In the wedge representation the angles between two consecutive wedge twist axes are approximately (see Fig. 5a):

$$\alpha_{w12} \cong (\alpha_{w1} + \alpha_{w2})/2$$
$$\gamma_{w12} \cong (\gamma_{w1} + \gamma_{w2})/2\,. \qquad \text{(A13)}$$

α_{w12} and γ_{w12} are the angles between two consecutive wedge twist axes with respect to the long and short base-pair-axes of the common base-pair, respectively. For normal DNA, the wedge twist axes will be ordered in the form of a spiral. Alternatively, in the cylinder representation the angles between two consecutive cylinder twist axes are approximately (see Fig. 5b):

$$\alpha_{c12} \cong \alpha_{c1} - \alpha_{c2} \quad \text{and} \quad \gamma_{c12} \cong \gamma_{c1} - \gamma_{c2}\,. \qquad \text{(A14)}$$

Consequently DNA (or RNA) stretches with nearly the same cylinder tilt and roll angles will have nearly parallel cylinder twist axes and are therefore called 'straight'. If several consecutive cylinder twist axes are nearly parallel, they can be represented by a 'block helix axis' (see Fig. 4b). If the cylinder twist axes are nearly parallel for the whole molecule, the block helix axis is the 'global helix axis'. In general, such a global helix axis will not exist. Additionally, for special sequences the local or block helix axes might add up to a curved DNA trajectory. For such a discussion, the wedge twist axes are not convenient.

Here, roll and tilt angles are properties with respect to the local helix axis defined between two

neighbouring base-pairs and thus, are addressed to a dinucleotide. The cylinder roll and tilt angles of a *single* base-pair can be defined only if a global helix axis is given. While the wedge parameters display the local sequence properties, the cylinder parameters allow a better insight into the global helical properties of the sequence. The formalism applied here for wedge and cylinder representation is of general nature. It can be used also if parameters are defined differently than those discussed here. The two geometrical parametrizations describe the structure with parameters of intuitive meaning. Analyzing the structure in terms of Euler angles would not have the same intuitive interpretation; on the other hand, this parametrization might be helpful when the structural correlation between distant base-pairs is studied.

References

Arnott S, Hukins DWL (1972) Optimized parameters for A-DNA and B-DNA. Biochem Biophys Res Commun 47: 1504–1510

Arnott S, Dover SD, Wonacott AJ (1969) Least-squares refinement of the crystal and molecular structure of DNA and RNA from X-ray data and bond length and angles. Acta Crystallogr B25: 2192–2206

Arnott S, Chandrasekaran R, Hall IH, Puigjaner LC (1983) Heteronomous DNA. Nucl Acid Res 11: 4141–4151

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: A program for macromolecular energy, minimization and dynamics calculations. J Comp Chem 4: 187–217

Bubienko E, Cruz P, Thomason JF, Borer PN (1983) Nearest-neighbour effects in the structure and function of nucleic acids. Prog Nucl Acids Res Mol Biol 30: 41–90

Dickerson RE (1983) Base sequence and helix structure variation in B- and A-DNA. J Mol Biol 166: 419–441

Diekmann S (1986) Sequence specificity of curved DNA. FEBS Lett 195: 53–56

Diekmann S (1987a) Temperature and salt dependence of the migration anomaly of curved DNA fragments. Nucl Acids Res 15: 247–265

Diekmann S (1987b) DNA curvature. In: Eckstein F, Lilley DM (eds) Nucleic acids and molecular biology, vol 1. Springer, Berlin Heidelberg New York

Diekmann S, Wang JC (1985) On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities. J Mol Biol 186: 1–11

Drew HR, Travers AA (1985) DNA bending and its relation to nucleosome positioning. J Mol Biol 186: 773–790

Edmondson SP, Johnson WC (1985) Base tilt of poly(dA) · poly(dT) and poly(dAT) · poly(dAT) in solution determined by linear dichroism. Biopolymers 24: 825–841

Fratini AV, Kopka ML, Drew HR, Dickerson RE (1982) Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATT(Br)CGCG. J Biol Chem 257: 14686–14707

Griffith J, Bleyman M, Rauch CA, Kitchin PA, Englund PT (1986) Visualization of the bent helix in kinetoplast DNA by electron microscopy. Cell 46: 717–724

Hagerman PJ (1984) Evidence for the existence of stable curvature of DNA in solution. Proc Natl Acad Sci USA 81: 4632–4636

Hagerman PJ (1985) Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. Biochemistry 24: 7033–7037

Hagerman PJ (1986) Sequence-directed curvature of DNA. Nature 321: 449–450

Haran TE, Berkovich Z, Shakked Z (1984) Base-stacking interactions in double-helical DNA structures: experiment versus theory. J Biomol Struct Dyn 2: 397–412

IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1983) In: Pullman B, Jortner J (eds) Nucleic acids: the vectors of life. Reidel, Dordrecht, pp 559–565

Jolles B, Laigle A, Chinsky L, Turpin PY (1985) The poly(dA) strand of poly(dA) · poly(dT) adopts an A-form in solution: a UV resonance Raman study. Nucl Acid Res 13: 2075–2085

Katahira M, Nishimura Y, Tsuboi M, Sato T, Mitsui Y, Iitaka Y (1986) Local and overall conformations of DNA double helices with the A·T base pairs. Biochim Biophys Acta 867: 256–267

Kitzing Ev (1986) Molekülsimulation mit Hilfe von Kraftfeldrechnungen am Beispiel der Aggregation von Nukleinsäuren verschiedener Konformation zu einem Komplex mit Übersetzungsfunktion. Edition Herodot/Rader, Aachen

Klement R, Kitzing Ev, Jovin T, Soumpasis DM (1985) Molecular mechanical simulations of the DNA B-Z transition compared to the RNA A-Z transition. In: Sarma RH (ed) Book of abstracts, 4th Conversation in Biomolecular Stereodynamics. Adenine Press, New York

Koo HS, Wu HM, Crothers DM (1986) DNA bending at adenine-thymine tracts. Nature 320: 501–506

Kunkel GR, Martinson HG (1981) Nucleosomes will not form on double-stranded RNA or over poly(dA) · poly(dT) tracts in recombinant DNA. Nucl Acid Res 9: 6869–6888

Levene SD, Crothers DM (1983) A computer graphics study of sequence-directed bending in DNA. J Biomol Struct Dyn 1: 429–435

Levitt M (1983) Protein folding by restrained energy minimization and molecular dynamics. J Mol Biol 170: 723–764

Marini JC, Levene SD, Crothers DM, Englund PT (1982) Bent helical structure in kinetoplast DNA. Proc Natl Acad Sci USA 79: 7664–7668

Marini JC, Englund PT (1983) Correction. ibid. 80: 7678

Peck LJ, Wang JC (1981) Sequence dependence of the helical repeat of DNA in solution. Nature 292: 375–378

Premilat S, Albiser G (1984) Conformation of C-DNA in agreement with fiber X-ray and infrared dichroism. J Biomol Struct Dyn 3: 607–613

Prunell A (1982) Nucleosome reconstitution on plasmid inserted poly(dA) · poly(dT). EMBO J 1: 173–179

Prunell A, Goulet I, Jacob Y, Goutorbe J (1984) The smaller helical repeat of poly(dA) · poly(dT) relative to DNA may reflect the wedge property of the dA · dT base pair. Eur J Biochem 138: 253–257

Rao SN, Kollman PA (1985) On the role of uniform and mixed sugar puckers in DNA double-helical structures. J Am Chem Soc 107: 1611–1617

Remerie K, Gunsteren WF van, Engberts JBFN (1985) Molecular dynamics computer simulation as a tool for analysis of solvation. A study of dilute aqueous solution of 1,4- and 1,3-dioxan. Recl Trav Chim Pays-Bas 104: 79–89

Rhodes D (1979) Nucleosome cores reconstituted from poly(dA-dT) and the octamer of histones. Nucl Acid Res 6: 1805–1816

Rhodes D, Klug A (1981) Sequence-dependent helical periodicity of DNA. Nature 292:378–380

Sarma MH, Gupta G, Sarma RH (1985) Untentability of the heteronomous DNA-model for $dA_n \cdot dT_n$ in solution. J Biomol Struct Dyn 2:1057–1084

Sasisekharan V, Bansal M, Gupta G (1983) Structures of DNA: a case study of right and left handed Duplex in the B-form. In: Pullman B, Jortner J (eds) Nucleic acids: the vectors of life. Reidel, Dordrecht, pp 101–111

Selsing E, Wells RD, Alden CJ, Arnott S (1979) Bent DNA: visualization of a base-paired and stacked A-B conformational junction. J Biol Chem 254:5417–5422

Shanno DF (1978) On the convergence of a new conjugate gradient algorithm. SIAM J Numer Anal 15:1247–1252

Simpson RT, Kunzler P (1979) Chromatin and core particles formed from the inner histones and synthetic polydeoxyribonucleotides of defined sequence. Nucl Acid Res 6: 1387–1415

Singh UC, Weiner SJ, Kollman P (1985) Molecular dynamics simulations of d(CGCGA) · d(TCGCG) with and without hydrated counter ions. Proc Natl Acad Sci USA 82: 755–759

Soumpasis MD (1984) Statistical mechanics of B-Z-transition of DNA: Contribution of diffuse ionic interaction. Proc Natl Acad Sci USA 81:5116–5120

Strauss F, Gaillard C, Prunell A (1981) Helical periodicity of DNA, poly(dA) · poly(dT), and poly(dA-dT) · poly(dA-dT) in solution. Eur J Biochem 118:215–222

Thomas GA, Peticolas WL (1983) Fluctuations in nucleic acid conformations. 2. Raman spectroscopic evidence of varying ring pucker in A-T polynucleotides. J Am Chem Soc 105:993–996

Trifonov EN (1985) Curved DNA. CRC Crit Rev Biochem 19:89–106

Trifonov EN, Sussman JL (1980) The pitch of chromatin DNA is reflected in its nucleotide sequence. Proc Natl Acad Sci USA 77:3816–3820

Tung CS, Harvey SC (1986) Base sequence, local helix structure, and macroscopic curvature of A-DNA and B-DNA. J Biol Chem 261:3700–3709

Ulanovsky L, Trifonov EN (1987) Estimation of wedge components in curved DNA. Nature 326:720–722

Ulanovsky L, Bodner M, Trifonov EN, Choder M (1986) Curved DNA: design, synthesis, and circularization. Proc Natl Acad Sci USA 83:862–866

Warshel A, Levitt M (1976) Theoretical studies of enzymatic reactions. J Mol Biol 103:227–249

Wartell RM, Harrell JT (1986) Characteristics and variations of B-type DNA conformations in solution: a quantitative analysis of Raman band intensities of eight DNAs. Biochemistry 25:2664–2671

Weiner PK, Kollman PA (1981) AMBER: Assisted model building with energy refinement. J Comput Chem 2: 287–303

Weiner JS, Kollman PA, Case DA, Singh UC, Ohio C, Alogerma G, Profeta jr S, Weiner PK (1984) A new force field for molecular mechanical simulations of nucleic acids and proteins. J Am Chem Soc 106:765–784

Wu HM, Crothers DM (1984) The locus of sequence-directed and protein-induced DNA bending. Nature 308:509–513

Zhurkin VB (1985) Sequence-dependent bending of DNA and phasing of nucleosomes. J Biomol Struct Dyn 2:785–804