Nucleic Acids Research, 1993, Vol. 21, No. 3 569–576

Conformation of d(GGGATCCC)$_2$ in crystals and in solution studied by X-ray diffraction, Raman spectroscopy and molecular modelling

Heinz Fabian, Wolfgang Hölzer[1], Udo Heinemann[2], Heinz Sklenar and Heinz Welfle*
Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Str. 10, D-1115 Berlin-Buch, [1]Pedagogic University Erfurt, Nordhäuser Str. 63, D-O-5010 Erfurt and [2]Institute of Crystallography, Free University Berlin, Takustr. 6, D-1000 Berlin 33, Germany

Received October 16, 1992; Revised and Accepted January 4, 1993

ABSTRACT

In the crystal, d(GGGATCCC)$_2$ forms an A-DNA double helix as known from a single crystal X-ray diffraction study. Accordingly, in the Raman spectra of crystals the A-family marker bands at 664, 705, 807 and 1101 cm^{-1} and the spectral characteristics in the region 1200 to 1500 cm^{-1} clearly demonstrate the A-form as the dominant conformation. Bands at 691, 850, and 1080 cm^{-1}, however, indicate that a minor fraction of the octamer molecules in the crystal is in an unusual, still not unequivocally identified conformation possibly belonging to the B-family. In solution, the octamer is in B-like conformation as shown by the presence of B-DNA Raman marker bands at 685, 837, 1094 and 1421 cm^{-1}. Molecular modelling techniques lead to three structures with slightly different B-form geometries as the lowest energies models when a sigmoidal dielectric function with the bulk dielectric constant ϵ = 78 and the value q = −0.5e for the effective phosphate charges was used in the calculations. An A-form structure bearing a strong resemblance to the experimentally determined crystal structure becomes the lowest energy model structure when the electrostatic parameters are changed to ϵ = 30 and q = −0.25e, respectively.

INTRODUCTION

Upon changes of sequence and environmental conditions, DNA can adopt different conformations or may undergo conformational transitions like the B-to-A transition. Under rather similar solution conditions, oligonucleotides may crystallize in the A, B or Z forms depending on their specific sequences. On the other hand, the same oligonucleotide may adopt different conformations depending on the buffer conditions. Surprisingly, evidence for the coexistence of two different helical structures in a single crystal has been obtained very recently for d(GG^{Br}UA^{Br}UACC)$_2$ by X-ray diffraction (1) and for d(GGTATACC)$_2$ by infrared and Raman spectroscopy (2). d(GG^{Br}UA^{Br}UACC)$_2$ (1) and d(GGTATACC)$_2$ (3) crystallize in the hexagonal space group P6$_1$. In the crystals of d(GG^{Br}UA^{Br}UACC)$_2$, six A-DNA helices, related by the 6$_1$ screw axis, surround a channel with a diameter of 22 Å which is large enough to accommodate additional DNA molecules. The observed diffuse X-ray scattering indeed suggests that additional octamer molecules are occluded in this channel and that they adopt the B conformation. In the infrared and Raman spectra of the d(GGTATACC)$_2$ crystal, conformation marker bands were found indicating the A-DNA conformation for the majority of the molecules and pointing to the presence of 15% (infrared) or 20% (Raman) of B-form DNA in the crystals (2). These findings fit to a model with six octamer molecules in A-form surrounding one octamer molecule in B-form similar to a clathrate structure as proposed for d(GG^{Br}UA^{Br}UACC)$_2$ (1).

The self-complementary octamer d(GGGATCCC)$_2$ also crystallizes in the space group P6$_1$ (4). In the crystal, d(GGGATCCC)$_2$ forms an A-DNA double helix with slightly varying conformation of the two strands. As in the related octamers d(GGTATACC)$_2$ and d(GG^{Br}UA^{Br}UACC)$_2$, the d(GGGATCCC)$_2$ molecules pack in spirals around the crystallographic *c*-axis leaving wide continuous solvent channels in the crystal (5). Here we present and discuss the Raman spectra of d(GGGATCCC)$_2$ in the crystal and in solution. There is spectroscopic evidence for the presence of a minority of octamers in the crystal in an unusual conformation possibly belonging to the B-family besides a majority of A-form molecules as a further example for the coexistence of oligonucleotides in two different conformations in a single crystal lattice as described for the first time for d(GG^{Br}UA^{Br}UACC)$_2$ (1). Furthermore, according to the Raman spectrum, the conformation of d(GGGATCCC)$_2$ in solution belongs clearly to the B-family in contrast to the A-form found in the crystal (4).

The application of molecular modelling techniques for deriving energetically optimized oligonucleotide structures allows to address further questions: How is the experimentally observed conformational polymorphism of a DNA duplex reflected in low-energy structural models stabilized by sequence-dependent

*To whom correspondence should be addressed

intramolecular interactions under different environmental conditions? And which conformers can be expected to contribute to the conformational equilibrium in solution? The results of the calculations are in general agreement with the experimental data and support the view that d(GGGATCCC)$_2$ is an impressive example for the conformational polymorphism possible for certain oligonucleotide sequences. The possible functional significance of the polymorphism of specific DNA sequences justifies the continued interest in such structures.

METHODS

DNA synthesis and crystallization

The octamer d(GGGATCCC)$_2$ was synthesized utilizing an Applied Biosystem Model 380B DNA Synthesizer and purified by HPLC. Samples were desalted by repeated lyophilizations and prepared at the desired buffer conditions through Sephadex G-10 filtration. Before the Raman measurements the samples were dissolved in appropriate volumes of H_2O. Crystallization was performed as described earlier (4) by microdialysis of the DNA solution against 20 mM sodium cacodylate buffer, pH 7.5, and 5 mM magnesium chloride supplemented with 24% 2-methyl-2,4-pentanediol. Finally, crystals were transferred into 10 mM Tris−HCl, pH 7.5, 10 mM $MgCl_2$ containing 24% isopropanol in order to reduce the disturbance of the crystal spectra by mother liquor bands.

Raman spectroscopy

Raman spectra were excited with the 488 nm line of an argon laser with a radiation power at the sample of about 30 mW for crystals and about 150 mW for solutions. The Raman scattering at 90° was collected on a computer-coupled Raman spectrometer as described earlier in more detail (6). Crystal spectra were measured from large single crystals as well as from very small crystals (larger crystals were cracked in the Raman tube) in order to avoid possible anisotropy effects due to defined orientations of the laser beam to the crystallographic axes of the crystals. The spectra obtained were very similar and therefore the averaged spectra are given. Spectral data were recorded at intervals of 1 cm^{-1} using a slit width of 4 cm^{-1}. About 15 scans were collected for each sample, averaged and prepared for analysis by a base-line adjustment procedure based on a polynomial curve.

Data treatment, including smoothing, subtraction and curve fitting was performed using a PC Dell 325D. Curve fitting was performed using a program written in Turbo Pascal by O. Ristau, Berlin-Buch, and based on the Marquardt algorithm (7). Solvent background correction for Raman spectra of crystals was performed by subtracting the normalized mother liquor spectrum recorded at the same conditions as the crystal spectrum.

Molecular modelling

Our theoretical studies on conformational properties of d(GGGATCCC)$_2$ are based on methodological developments carried out in collaboration with Richard Lavery (Paris). This work has led to the versatile modelling program JUMNA (JUnction Minimization of Nucleic Acids) and to the helical analysis program CURVES, both particularly suited to the study of nucleic acid structures. These programs were quite successful in reproducing experimentally established data and parameter correlations describing the mechanics of oligonucleotide chains. The detailed description of the algorithms, including the energy formulation termed 'FLEX', are given in references (8−13).

In contrast to the classical molecular mechanics techniques, the modelling algorithms in JUMNA make direct use of the helical parameters describing global features of oligomer conformations. This gives new possibilities for a thorough search of conformational subspaces with much better chances to locate structures with low intramolecular energies. In the initial step we use an automatic procedure which applies different distortions to the individual nucleotide and dinucleotide conformations in the A and B allomorphs. These structures are the starting points, subsequently, for deriving energy optimized conformations by means of the conjugate gradient method.

In the present studies, the distortions have included the helical twist (from 25 to 40 degrees) and rise (from 2.5 Å to 3.6 Å) at the different dinucleotide steps as well as the sugar pucker parameters (between 2′-endo and 3′-endo) and the glycosidic angles (between syn- and anti-conformations).

The dielectric model used in the 'FLEX' parametrization of the force field mimics the damping of electrostatic interactions due to the solvent by introducing a sigmoidal function $\epsilon(R)$ for the distance dependence of the dielectric constant as proposed by Hingerty et al. (14). In the studies of solution structures counter-ion effects were taken into account by reducing the net phosphate charge to $q = -0.5e$, and a small slope of $\epsilon(R)$ was chosen to reach the bulk dielectric conditions $\epsilon = 78$ at a distance of about 80 Å. As an attempt to approach crystal conditions the value $q = -0.25e$, and a slope of $\epsilon(R)$ approaching the bulk value $\epsilon = 30$ at 30 Å were used, assuming a more effective shielding of phosphate charges by counter-ions and a lower dielectric constant reduced by the organic solvent.

RESULTS

Evidence for the presence of a second conformation in crystals of d(GGGATCCC)$_2$

Ordered DNA octamers of d(GGGATCCC)$_2$ are clearly in the A-form in the crystal (4). However, the six d(GGGATCCC)$_2$ duplexes in the unit cell do not fully account for the observed diffraction pattern. Precession photographs of the h0l projection show diffuse scattering perpendicular to the crystallographic c* and *c* axes (Figure 1). This scattering corresponds to a spacing of 3.4 Å and is therefore reminiscent of the strong meridional reflection of B-DNA fiber diffraction arising from the base pairs stacking at 3.4 Å distance perpendicular to the fiber axis. Hence, the presence of rotationally disordered octamer molecules in the solvent channel along *c* might explain this feature of the diffraction. Formal proof for this hypothesis, however, requires a more complete fiber diffraction pattern than recorded for d(GGGATCCC)$_2$ or additional evidence from other techniques. It should be noted that, assuming linearly stacked B-DNA octamers in the channel, the unit cell dimensions with $c = 44.49$ Å put an upper limit of 27% to the fraction of B-DNA molecules in crystals of d(GGGATCCC)$_2$.

d(GGGATCCC)$_2$ has B-form geometry in solution

Figure 2 (bottom) shows the Raman spectrum of the octamer in solution in the region 600 to 1800 cm^{-1}. Significant peaks are marked by their wavenumbers, and correlations to base, backbone, and sugar vibrations are indicated by abbreviations as explained in the legend of Figure 2. The spectrum contains all the typical bands which can be expected for an oligonucleotide of such base composition. Conformation marker bands at 685, 837, 1094 and 1421 cm-^{1} clearly demonstrate that the

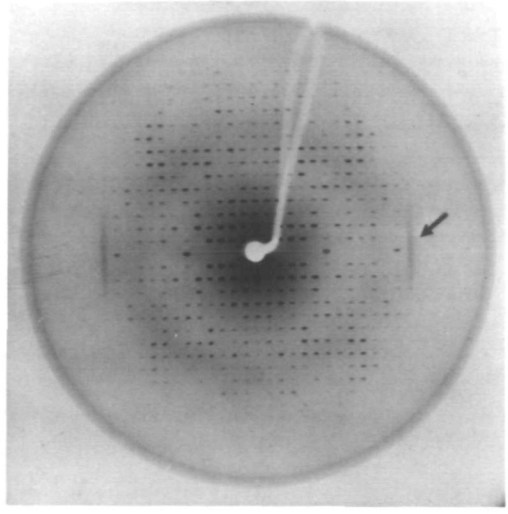

Figure 1. 15° precession photograph of the h0l layer of a d(GGGATCCC)$_2$ crystal. The space group is P6$_1$, a = 46.83 Å, c = 44.49 Å. The arrow points to a diffuse reflection which is not part of the Bragg pattern and unaccounted for by the crystal structure of A-DNA molecules spiralling around a large solvent channel along c.

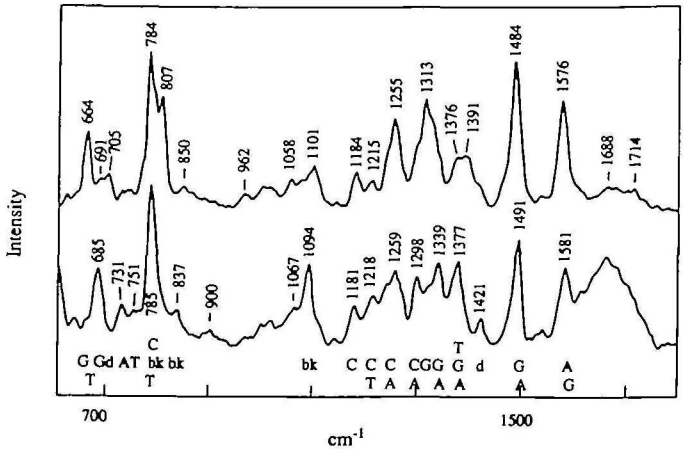

Figure 2. Raman spectra of d(GGGATCCC)$_2$ in solution (2 mM sodium cacodylate buffer, 150 mM NaCl, pH 7.5, c ≈ 50 mg·ml^{-1}) (bottom) and of d(GGGATCCC)$_2$ crystals (top) in the region 600 to 1800 cm^{-1} measured at 15°C. Raman frequencies of prominent bands are listed in cm^{-1} and assignments are indicated as follows: guanine (G), cytosine (C), adenine (A), thymine (T), backbone (bk), and deoxyribose (d).

conformation of the octamer in solution at near physiological salt concentration belongs to the B-family according to known spectrum-structure correlations (15–19).

A more detailed view of the fingerprint spectrum region 620 to 900 cm^{-1} is given in Figure 3 together with results of a curve fitting procedure which resolved the experimental data into a sum of Lorentzian bands. With respect to the sugar backbone conformation of the octamer in solution, the bands at 838 cm^{-1} and at 807 cm^{-1} are most interesting. A band at 830 to 840 cm^{-1} is observed in all B-form nucleic acids and is correlated to the presence of furanose rings in the C2′-endo conformation (15–19). The C2′-endo sugar conformation is associated with

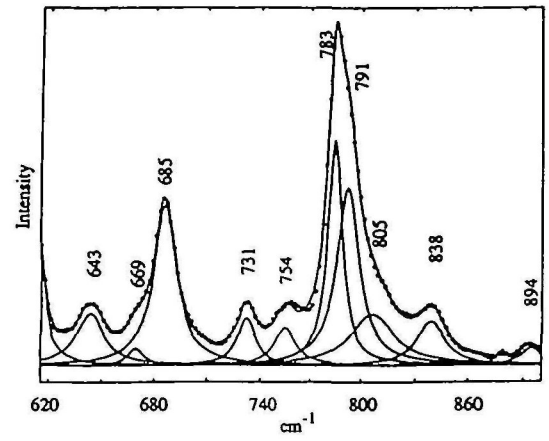

Figure 3. Fingerprint region 620 to 900 cm^{-1} of the solution Raman spectrum of d(GGGATCCC)$_2$. Individual band components were obtained by curve fitting and are represented by thin lines. The calculated spectrum obtained by curve fitting is superimposed on the dots of the experimental spectrum.

the classical B-form structure. Thus the strong band at 838 cm^{-1} indicates B-family conformation of the octamer in solution. In addition, the weak band at 805 cm^{-1} in Figure 3 can be attributed to furanose rings in a C3′-endo conformation (16–18). The C3′-endo furanose ring conformation is associated with A-form nucleic acids. The amount of C3′-endo conformation present in the population can be calculated based upon the intensity ratio I_{805}/I_{1094}. Taking an I_{805}/I_{1094} of 1.8 for pure A-form structures (20, 21), a C3′-endo population of about 16% in the octamer in solution can be estimated from its I_{805}/I_{1094} value of 0.29. This corresponds to one residue per strand in d(GGGATCCC)$_2$ or to one octamer molecule out of eight.

B-family conformation is also indicated by the 685 cm^{-1} band assigned to a guanine coupled furanose vibration. This band is sensitive to the orientation (syn or anti) of the guanine residue relative to the furanose residue and to the conformation of the furanose ring and usually found at 682 cm^{-1} in B-form polymers and native DNA with guanine in anti orientation attached to furanose rings in the C2′-endo conformation (17). The slight deviation of 3 cm^{-1} of the 865 cm^{-1} band from the usual position at 682 cm^{-1} of this band in native, random sequence DNA (20) points to differences in the furanose pucker and/or glycosyl orientation of dG residues of the octamer in comparison to random sequence DNA.

Spectroscopic evidence for the coexistence of two different conformations in the d(GGGATCCC)$_2$ crystal

Figure 2, top, shows the Raman spectrum in the region 500 to 1800 cm^{-1} of the d(GGGATCCC)$_2$ crystal as the difference spectrum of the crystal spectrum in mother liquor minus the spectrum of mother liquor. Unfortunately, the position of a strong band of the mother liquor at 818 cm^{-1} is very close to the position of the A-form and B-form backbone marker bands at about 807 and 835 cm^{-1}, respectively. Further mother liquor bands are localized at 948 and 1455 cm^{-1}. Careful compensation of the mother liquor bands by spectral subtraction is very important and confirmed by the absence of the mother liquor Raman band at 948 cm^{-1} in the crystal spectrum. At the first glance it is obvious that crystal and solution spectra differ extensively and reflect different conformations of d(GGGAT-

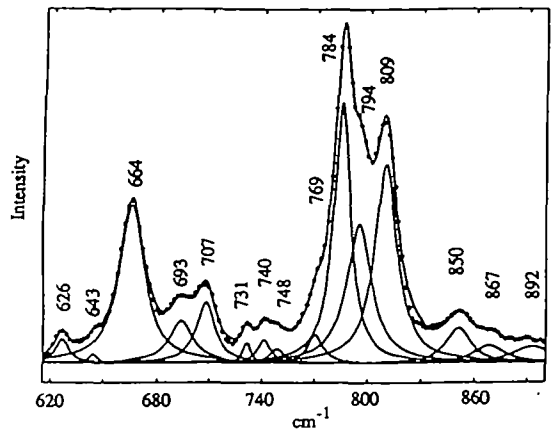

Figure 4. Fingerprint region 620 to 900 cm^{-1} of the Raman spectrum of d(GGGATCCC)$_2$ crystals and results of curve fitting.

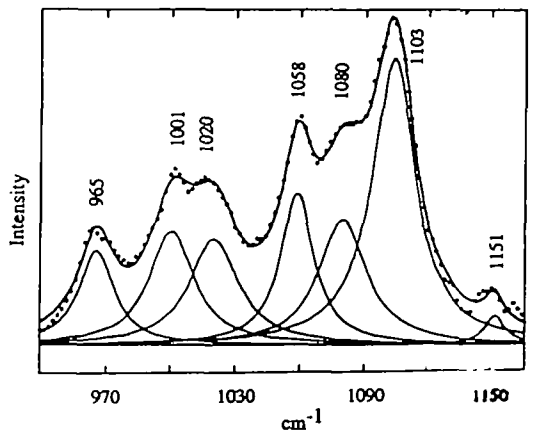

Figure 5. Raman spectrum of d(GGGATCCC)$_2$ crystals in the region 940 to 1165 cm^{-1} and results of curve fitting.

CCC)$_2$ in crystal and solution. In the crystal spectrum bands at 664, 705, 807, and 1101 cm^{-1} typical for the A-form (15–19) are to be found. A band around 665 cm^{-1} is assigned to guanosine residues in C3′-endo,anti conformation. Bands at 706 ± 5 cm^{-1} belong to a ribose-phosphate vibration in A-DNA. At 807 ± 3 cm^{-1} the backbone marker band of A-DNA is positioned and at 1101 cm^{-1} the symmetric stretching vibration of the PO_2^- group of A-DNA is found. The presence of these bands in the crystal spectrum of d(GGGATCCC)$_2$ presented in Figure 2 (top) clearly indicates the A-form for the octamer as the dominant conformation in the crystal in agreement with the crystal structure analysis (4).

An interesting feature of the octamer crystal spectrum can be seen in Figure 3, top in the fingerprint region between the bands at 664 and 705 cm^{-1}. The intensity in this region is unusually high for an A-form nucleic acid and a band is marked at 691 cm^{-1}. The position of this band in the crystal spectrum is near to the G-specific band at 682 ± 5 cm^{-1} typical for B-form DNA and found as a very pronounced band at 685 cm^{-1} in the solution spectrum of the octamer (Figure 2, bottom). This spectral region is more clearly represented in Figure 4 showing the Raman spectrum of the crystal in the wavelength region 620 to 900 cm^{-1} (dots) together with the results of a curve-fitting procedure. A band at 693 cm^{-1} is obtained and presumed to be indicative of backbone geometry within the B-family because of its proximity to the band at 682 cm^{-1} assigned to a vibration of guanosine in C2′-endo, anti conformation.

Another feature of the octamer crystal spectrum concerns the presence of a minor band at 850 cm^{-1} correlated with ribose phosphate backbone vibrations. A phosphodiester marker near 850 to 860 cm^{-1} was assigned to an all-trans conformation (trans/trans/trans) of the α, β, and γ torsions (23) within an overall B-genus conformation (24). Thus, this band gives evidence for a B-form backbone conformation besides the A-form backbone indicated by the strong marker band at 807 cm^{-1}.

A surprisingly low intensity of the band of the phosphate diester stretch at 1101 cm^{-1} was found in the octamer crystal spectrum (Figure 2, top). As already mentioned before, the intensity ratio of the strong A-form marker band at about 810 cm^{-1} and of the band at 1098 cm^{-1} can be used for an estimate of the amount of A-form DNA in nucleic acids (20, 21). The maximum value for the ratio of the peak heigths is usually taken about 1.8 for 100% A-form (20, 21). In the case of d(GGGATCCC)$_2$ crystals, however, a ratio of 2.4 is obtained. This ratio is even higher than the highest previously observed value of 2.0 in the Raman spectrum of d(GGTATACC)$_2$ crystals (2). In addition to the band at 1101 cm^{-1} a band at about 1080 cm^{-1} is observed in the Raman spectrum of the d(GGGATCCC)$_2$ crystal. Figure 5 shows the region 940 to 1165 cm^{-1} of the d(GGGATCCC)$_2$ crystal spectrum (dots) and the sum of Lorentzian component bands obtained by curve-fitting. There is no band component at 1094 cm^{-1} in the usual position of the symmetric PO_2^- vibration of B-DNA besides the 1101 cm^{-1} band (at 1103 cm^{-1} after curve fitting) of the dominating A-form. Instead, a band component at 1080 cm^{-1} is found and it is tempting to correlate this component with a PO_2^- vibration of an unusual B-form conformation.

Molecular mdelling of d(GGGATCCC)$_2$

The calculations were started from more than 100 differently distorted structures of the canonical A- and B-conformations of the d(GGGATCCC)$_2$ duplex structure and were performed for different values of the dielectric constant, of the slope of the sigmoidal function $\epsilon(R)$, and of the effective phosphate charge. These structures relaxed during energy minimization finally to only four conformations belonging each to a local minimum of the energy surface and exhibiting significant conformational differences. Three of the final structures belong to the B-family and one shows the characteristics of the A allomorph. Under solution conditions, with an effective phosphate charge q = $-0.5e$ and a sigmoidal dielectric function $\epsilon(R)$ approaching the bulk value $\epsilon = 78$ at 80 Å, the three B-form structures have only slightly varying energies and are more stable than the A-form by about 4 kcal·mol^{-1}. Comparative studies with other sets of parameters have shown that strong and long range electrostatic interactions stabilize the B-forms compared to the A-form, and vice versa, any electrostatic damping shifts the equilibrium towards the A conformation. Thus, a continuous change of the relative stability in favor of the A-form is obtained by lowering the effective phosphate charge and by increasing the slope of $\epsilon(R)$.

Our crude electrostatic model does not allow to choose a priori parameters that correspond to crystal conditions. For this reason, we have compared for a wide range of parameters the calculated

Table 1. Structural parameters of the lowest energy B-form model of d(GGGATCCC)$_2$

	Sugar pucker Phase	Ampli	Glyc. angle chi	Xdisp dx	Inclin eta	Rise Dz	Twist Omega
G1	159.0	38.9	−141.1	−1.69	−20.2		
G2	157.1	43.6	−122.5	−1.61	−17.3	4.58	35.6
G3	155.5	43.2	−120.8	−1.57	−12.4	4.08	34.7
A4	163.2	39.5	−114.4	−1.53	−10.2	3.39	37.4
T5	160.0	41.2	−114.2	−1.53	−10.2	3.32	34.9
C6	159.9	41.4	−116.7	−1.57	−12.4	3.39	37.4
C7	153.6	43.2	−132.3	−1.61	−17.3	4.08	34.7
C8	154.1	39.3	−136.0	−1.69	−20.2	4.58	35.6

The helix parameters Xdisp (X displacement), Inclin (inclination), Rise, and Twist are base pair parameters with reference to the curved global axis of the duplex. They were calculated by means of the CURVES algorithm (12, 13). Lengths are given in Å, angles in degrees.

Table 2. Comparison of structural parameters of three low-energy B-form models

B-form Model		Mean values Xdisp	Inclin	Rise	Twist	Twist G3/A4	A4/T5	T5/C6
all	2′-endo	−1.60	−15.0	3.91	35.7	37.4	34.9	37.4
C6:	O4′-endo	−1.91	−12.1	3.84	34.2	32.3	36.2	32.2
T5:	O4′-endo	−1.20	−15.2	3.88	35.4	38.7	28.0	38.7

For explanations see Table 1.

A-form structures with the experimental X-ray structure. The best fit was obtained for a phosphate charge q = −0.25e, and a sigmoidal function $\epsilon(R)$ which approaches the asymptotic value $\epsilon = 30$ at 30 Å, i.e. under highly damped electrostatic interactions. For these parameters the calculated energy of the A-form is about 10 kcal·mol^{-1} lower than the energies of the B-forms. This energy difference, however, seems not to be very meaningful because changes of the electrostatic conditions in the calculations have only minor effects on the structural parameters of the A-form but reduce the energy difference. Thus, the fit to the experimental structure is only slightly worse when the phosphate charge is increased to about q = −0.40e where the energies of the A- and B-forms are almost the same.

Models of the solution conformation of d(GGGATCCC)$_2$

In the lowest energy model all sugars are in C2′-endo conformation. The two other structures have an O4′-endo sugar pucker at cytidine in position 6 and at thymidine in position 5 and increased energies of 1.0 kcal·mol^{-1} and 1.6 kcal·mol^{-1}, respectively. The most important structural parameters characterizing the all C2′-endo conformation are given in Table 1. The overall structure of the calculated model is depicted in Figure 6 (left) together with a schematic representation (right) showing the slight curvature of the helix axis as a result of the CURVES analysis (12, 13).

Recent crystallographic (25) and NMR studies (26, 27) have shown the existence of O4′-endo puckering in oligonucleotides containing cytosine and/or thymine in their sequences. O4′-endo sugars are always sandwiched between nucleotides with C2′-endo sugars. Thus, the induced structural changes are absorbed by local environments of the nucleotides concerned and have only minor effects on the overall structure. In Table 2 the mean values of several helical parameters and the helical twist angles in the central part of the octamer are summarized for a comparison of the essential differences of the three low-energy B-form model structures.

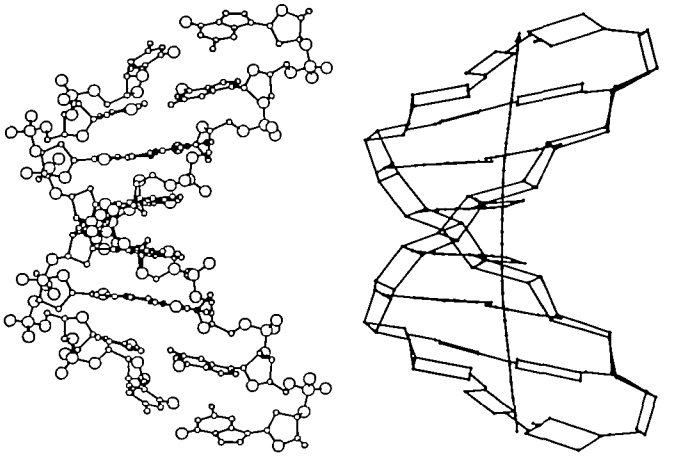

Figure 6. Lowest energy B-form model of d(GGGATCCC)$_2$ with all sugars in C2′-endo conformation. Ball and stick (left) and schematic representation of the model including the curved helical axis of the duplex structure (right).

A-form model structure of d(GGGATCCC)$_2$

Under the condition of damped electrostatic interactions the structure with lowest energy has A-form characteristics. A reasonable choice of electrostatic parameters led to a model with all essential features observed in the crystal structure of d(GGGATCCC)$_2$ (4). For comparison, in Table 3 several structural parameters of the calculated A-form model and of the experimentally determined crystal structure are given. Low twist angles corresponding to 12 base pairs per complete turn indicate a considerably underwound helix compared with canonical A-DNA. Other significant deviations from the canonical structure, such as a small base pair inclination, an increased rise, and an opened major grove, are also in agreement with the crystallographic results. The molecular graphics shown in

Table 3. Comparison of calculated and experimentally determined structural parameters of the A-form of d(GGGATCCC)$_2$

	Mean values				Twist		
	Xdisp	Inclin	Rise	Twist	G3/A4	A4/T5	T5/C6
Calc.	−4.76	9.6	2.91	30.2	31.3	30.0	31.3
X-ray	−4.78	9.6	2.90	30.8	35.5	26.4	33.0

For explanations see Table 1.

Figure 7 illustrate the overall similarity of the calculated A-form model (left) and the crystal structure (right) of d(GGGATCCC)$_2$.

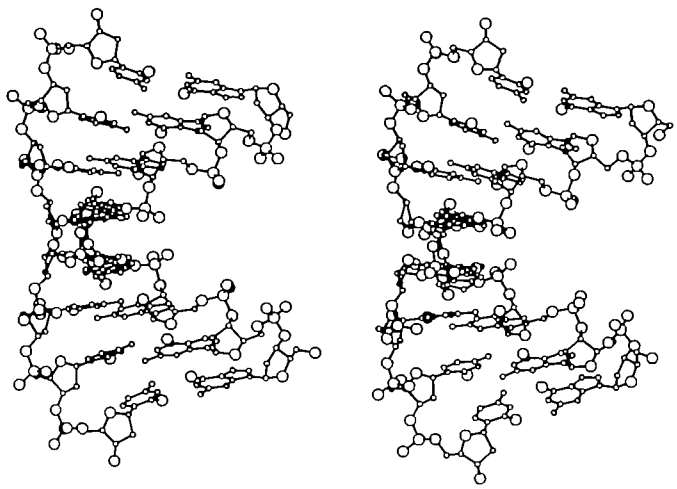

Figure 7. Comparison of calculated (left) and experimentally determined (right) A-form model of d(GGGATCCC)$_2$

DISCUSSION

The crystal structure analysis of d(GGGATCCC)$_2$ revealed an A-form DNA with structural parameters clearly distinct from those of ideal A-DNA (4). Incomplete evidence for the coexistence of B-form DNA in the crystals of d(GGGATCCC)$_2$ had been observed by diffuse scattering perpendicular to the orientation of large solvent channels in the crystals (Figure 1). In the absence of a detailed fiber pattern, however, this observation did not prove the coexistence of A-DNA and B-DNA in the crystal.

The Raman spectroscopic data presented here confirm the A-form as the dominant conformation of d(GGGATCCC)$_2$ and give clear evidence for the coexistence of a second conformation of a minor fraction of the octamer molecules in the crystal. This second conformation is not unequivocally identified but probably an unusual B-form DNA as discussed below.

Typical B-form marker bands would be expected at 682 ± 2 cm^{-1} for the guanosine C2′-endo,anti conformation, at about 830 cm^{-1} for the coupled sugar phospate backbone vibration and at 1094 cm^{-1} for the symmetric PO_2^- phosphodiester vibration (15−19) as found in the solution spectrum of d(GGGATCCC)$_2$ (Figure 2, bottom). Bands demonstrating a second conformation besides the dominating A-form in the crystal, however, were found at 693 cm^{-1}, at 850 cm^{-1}, and at 1080 cm^{-1}. Raman spectra with comparable characteristics have been observed already earlier and interpreted in different manner.

A band at 691 cm^{-1} was observed in the Raman spectra of calf thymus DNA fibers at 85% ethanol (29) and considered to reflect a shift of the 682 cm^{-1} band during the B-to-A transition of the DNA with increasing ethanol concentrations but the presence of residual B-form molecules is not excluded in these experiments.

In the Raman spectrum of crystals of d(GGTATACC)$_2$ these bands are located at 688 cm^{-1}, and at 854 cm^{-1} (2). At about 1080 cm^{-1} a weak shoulder of the 1098 cm^{-1} band can be seen in the spectrum but a band was not resolved in this position by the curve fitting procedure (2). In the Raman spectra of d(CCCCGGGG)$_2$ crystals bands at 686, 850, and about 1075 cm^{-1} were found (20). The band at 686 cm^{-1} indicates the C2′-endo,anti conformation of guanosine or a closely related C1′-exo or another minor variant of the dG sugar ring pucker (20). The weak Raman line near 850 cm^{-1} is discussed as indicative of a backbone geometry within the B-family and correlated (17, 20, 23) to an all-trans (trans/trans/trans) conformation of the α, β, and γ torsions at the P-O5′-C5′-C4′ bonds. This conformation was found for the central CpG step of d(CCCCGGGG)$_2$ which otherwise has a backbone conformation in the crystal similar to other crystal structures of A-DNA fragments (30). The all-trans conformation favors interstrand G-G base stacking and interstrand purine-purine stacking was also supposed to be correlated with a Raman band at 857 cm^{-1} found in the spectrum of the Lambda operator site O_L1 (23).

A surprisingly low intensity of the band at about 1100 cm^{-1} is observed in the crystal spectra (Figure 2, top) in comparison to reference spectra and an additional component at about 1080 cm^{-1} is identified by the curve fitting procedure (Figure 5). A comparable observation is described for the d(GGTATACC)$_2$ crystal infrared spectra (2). These spectra also show important differences in the region of the phosphate group vibrations which are discussed with respect to possible pecularities of phosphate-phosphate interactions in d(GGTATACC)$_2$ crystals in comparison to canonical A-form DNA (2). For both d(GGGATCCC)$_2$ and d(GGTATACC)$_2$, a more open major groove was observed in the crystals (3, 4) in comparison to the value of only 2.3 Å (closest phosphorus-phosphorus approach minus 5.8 Å to account for the van der Waals radii of the phosphate groups) obtained from fiber diffraction studies of A-form helices (31).

Taken together, the observed positions, intensities and shapes of the band of the symmetric PO_2^- vibration in the region of 1080 to 1100 cm^{-1} as well as the unusual positions of the bands at 693 and 850 cm^{-1} seem to indicate a specific, up to now not precisely defined conformation within the B-family of a fraction of the octamer molecules in the crystal which is connected with

peculiar phosphate characteristics, variants of the dG sugar pucker, and an all-trans (trans/trans/trans) conformation of the α, β, and γ torsions of the P-O5'-C5'-C4'-bonds.

Since vibrational spectroscopy yields an integral information on all components of the system, the question arises whether the bands indicating an unusual conformation in the crystals reflect the occurrence of some C2'-endo sugars in A-DNA molecules or the coexistence of duplexes that are either all A-like or all B-like. Deoxyriboses with C2'-endo pucker are not present in the A-form molecules found in the crystal structure analysis of $d(GGGATCCC)_2$ (4). The only deviation from normal A-DNA C3'-endo conformation is seen at one of the two 3'-terminal cytosine residues which shows a generally reduced pucker due to crystal contacts. In A-form oligonucleotides rigid C3'-endo pucker is normally observed with only occasional exceptions for 3'-terminal nucleosides (28). Therefore, it is considered unlikely that the Raman spectra reflect mixed sugar conformations in A-DNA duplexes. Instead, they indicate the coexistence of a majority of A-form octamers with some B-form octamers in the $d(GGGATCCC)_2$ crystals. The same conclusion was drawn for $d(GGTATACC)_2$ from the spectroscopic evidence of B-form conformation in its crystals (2). For a quantitative calculation of the amount of B-forms present in the crystals, the intensity ratios $I_{693}/I_{664+693}$ as well as the area ratios $A_{693}/A_{664+693}$ of the bands at 664 and 693 cm^{-1} can be taken. Comparable results of about 20% B-form in the crystals were obtained from these data pointing to the presence of one B-form octamer per four A-form octamers in the crystal. For $d(GGTATACC)_2$ comparable values of 20 to 24% were found (2). From the crystal structure analysis of $d(GG^{Br}UA^{Br}UACC)_2$ a model of six A-form oligonucleotides occluding one B-form oligonucleotide was derived (1). In view of the assumptions necessary for the quantification of the two conformations in the crystal from the Raman spectrum the agreement between these results seems to be resonably well.

The solution structure of $d(GGGATCCC)_2$ belongs to the B-family as demonstrated by the presence of typical B-form marker bands in the Raman spectrum (Figure 2). Under the influence of high salt concentrations, $d(GGGATCCC)_2$ is going to the A-form. This structural transition could not be analyzed by Raman measurements because the turbidity of the octamer solutions prevented such experiments, but was shown earlier by infrared measurements which can be performed also with turbid samples (14).

Molecular modelling of the octamer structure with systematic search procedures based on energy minimization led to only four low-energy conformations exhibiting significant structural differences, three belonging to the B-family and one showing the characteristics of the A allomorph. The relative stability of the A- and B-conformations depends on parameters describing the electrostatic interactions.

To mimic solution conditions with a bulk dielectric constant $\epsilon = 78$, an effective phosphate charge $q = -0.5e$ and a small slope of $\epsilon(R)$ approaching the bulk value at 80 Å were used. Then, in agreement with the Raman data, the structure with the lowest energy is a B allomorph. The structural parameters of the calculated B-DNA models deviate significantly from those of the canonical B-DNA. The negative X displacement (−1.5 Å) shifting the base pairs into the minor groove and, in particular, the large inclination (−15 degrees) were never found in B-DNA crystal structures. However, since the experimentally observed local conformations vary considerably with the base sequences of the studied oligonucleotides, it seems reasonable to explain the peculiarities of the calculated B-DNA models by the specific sequence of four purines followed by four pyrimidines in $d(GGGATCCC)_2$.

The calculated A-form structure can be directly compared with the X-ray results. A very good agreement was obtained in the calculation with the bulk dielectric constant $\epsilon = 30$, the effective phosphate charge $q = -0.25e$, and a slope of $\epsilon(R)$ approaching the bulk value at 30 Å. Under these conditions the relative stability of the A-form over the B-form amounts 10 $kcal \cdot mol^{-1}$. By increasing the effective phosphate charge equal energies of the A- and B-forms can be obtained. This leads to only slightly changed structural parameters of the calculated A allomorphs and supports the possible coexistence of A- and B-forms in the $d(GGGATCCC)_2$ crystals.

Conformational differences between the calculated model structures and the canonical A- and B-forms are in accordance with the finding of spectroscopic peculiarities of $d(GGGATCCC)_2$. A detailed interpretation, however, and especially a correlation of conformational properties of the model structures with the band positions of the Raman spectra is still not possible and would obviously lead to an overinterpretation of the available data.

ACKNOWLEDGEMENTS

We thank Dr Otto Ristau for writing and support in the implementation of programs for data treatment, Dr G.Herrmann for the preparation of an oligonucleotide sample, and Mrs B.Kannen for skilful technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (He 1318/6-1), by the Bundesminister für Forschung und Technologie under grant number BEO 021/0319799A (H.S.), and by the Fonds der Chemischen Industrie. The collaboration with Molecular Simulations Inc. enabling us to use the Quanta software for graphical visualizations is also gratefully acknowledged.

REFERENCES

1. Doucet,J., Benoit,J.-P., Cruse,W.B.T., Prange,T. and Kennard,O. (1989) *Nature* **337**, 190–192
2. Liquiers,J., Taillandier,E., Peticolas,W.L. and Thomas,G.A. (1990) *J. Biomol. Struct. Dyn.* **8**, 295–302
3. Shakked,Z., Rabinovich,D., Kennard,O., Cruse,W.B.T.,Salisbury,S.A. and Viswamitra,M.A. (1983) *J. Mol. Biol.* **166**, 183–201
4. Lauble,H., Frank,R., Blöcker,H. and Heinemann,U. (1988) *Nucl. Acid. Res.* **16**, 7799–7816
5. Heinemann,U., Alings,C. and Lauble,H. (1990) in: Sarma,R.H. and Sarma,M.H. (eds.), Structure and Methods, Vol.3: DNA and RNA, Adenine Press, Albany, pp. 39–53
6. Fabian,H., Hölzer,W., Herrmann,G., Ristau,O., Sklenar,H. and Welfle,H. (1990) *J. Mol. Struct.* **217**, 99–114
7. Marquardt,D.W. (1963) *J. Soc. Ind. Appl. Math.* **11**, 431–581
8. Sklenar,H., Lavery,R. and Pullman,B. (1986) *J. Biomol. Struct. Dyn.* **6**, 967–987
9. Lavery,R., Sklenar,H., Zakrzewska,K. and Pullman,B. (1986) *J. Biomol. Struct. Dyn.* **3**, 989–1014
10. Lavery,R., Parker,I. and Kendrick,J. (1986) *J. Biomol. Struct. Dyn.* **4**, 443–458
11. Lavery,R. (1988) in: Olson,W.K., Sarma,M.H., Sarma,R.H. and Sundaralingam,M. (eds.), Structure and Expression, Vol. 3, DNA Bending and Curvature, Adenine Press, pp. 191–211
12. Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dyn.* **6**, 63–91
13. Lavery,R. and Sklenar,H. (1989) *J. Biomol. Struct. Dyn.* **6**, 655–667
14. Hingerty,B.E., Richie,R.H., Ferrell,T.L. and Turner,J.E. (1985) *Biopolymers* **24**, 427–439

15. Nishimura,Y. and Tsuboi,M. (1986) in: Clark,R.J.H. and Hester,R.E. (eds.), Spectroscopy of Biological Systems, Vol. 13, Wiley, Chichester, pp. 177–232
16. Peticolas,W.L., Kubasek,W.L., Thomas,G.A. and Tsuboi,M. (1987) in: Spiro,T.G. (ed.), Biological Applications of Raman Spectroscopy, Vol. 1, Wiley, New York, pp. 81–133
17. Nishimura,Y., Tsuboi,M., Sato,T. and Aoki,K. (1986) *J. Mol. Struct.* **146**, 123–153
18. Thomas,G.J.,Jr., Benevides,J.M. and Prescott,B. (1986) in: Sarma,R.H. and Sarma,M.H. (eds.) Biomolecular Stereodynamics 4, Adenine Press, Albany, pp. 227–254
19. Thomas,G.J.,Jr. and Wang,A.H.-J. (1988) in: Eckstein,F. and Lilley,D.M.J. (eds.) Nucleic Acids and Molecular Biology, Vol. 2, Springer-Verlag Berlin Heidelberg, pp. 1–30
20. Benevides,J.M., Wang,A.H.-J., Rich,A., Kyogoku,Y., Van der Marel,G.A., Van Boom,J.H. and Thomas,G.J.,Jr. (1986) *Biochemistry* **25**, 41–50
21. Thomas,G.A., Kubasek,W.L., Peticolas,W.L., Greene,P,. Grable,J. and Rosenberg,J.M. (1989) *Biochemistry* **28**, 2001–2009
22. Wartell,R.M. and Harrell,J.T. (1986) *Biochemistry* **25**,2664–2671
23. Benevides,J.M., Weiss,M.A. and Thomas,G.J.,Jr. (1991) *Biochemistry* **30**, 5955–5963 24. Torigoe,C., Nishimura,Y., Tsuboi,M., Matsuzaki,J., Hotoda,H., Sekine,M. and Hata,T. (1986) *Spectrochim. Acta* **42A**, 1101–1106
25. Privé,G.G., Yagani,K. and Dickerson,R.E. (1991) *J. Mol. Biol.* **217**, 177–199
26. Metzler,W.J., Wang,C., Kitchen,D.B., Levy,R.M. and Pardi,A. (1990) *J. Mol. Biol.* **214**, 711–736
27. Schmitz,U., Zon,G. and James,T.L. (1990) *Biochemistry* **29**, 2357–2368
28. Heinemann,U., Lauble,H., Frank,R. and Blöcker,H. (1987) *Nucleic Acids Res.* **15**, 9531–9550
29. Martin,J.C. and Wartell,R.M. (1982) *Biopolymers* **21**, 499-512)
30. Haran,T.E., Shakked,Z., Wang,A.H.-J. and Rich, A. (1987) *J. Biomol. Struct. Dyn.* **5**, 199- 217
31. Chandrasekaran,R. and Arnott,S. (1989) In Saenger,W. (ed.), Landoldt-Börnstein, New Series, Group VII. Springer, Berlin, Heidelberg, New York, Vol 1b, pp. 31–170.