# Conformational Flexibility of the DNA Backbone

**Geertrui S. Schuerman[†,‡] and Luc Van Meervelt*[,†]**

*Contribution from the Laboratorium voor Macromoleculaire Structuurchemie, Departement Scheikunde, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium, and Laboratorium voor Analytische Chemie en Medicinale Fysicochemie, Faculteit Farmaceutische Wetenschappen, Katholieke Universiteit Leuven, Van Evenstraat 4, B-3000 Leuven, Belgium*

*Received June 25, 1999*

**Abstract:** An unprejudiced final model of the daunomycin−d(CGCGCG) complex was obtained at 1.1 Å resolution and at 100K after unrestrained *SHELXL* refinement, which was possible thanks to the high data-to-parameter ratio and which provided us with true standard deviations on positions and distances. The structural pattern that emerges from the refinement proves that the sugar phosphate backbone is considerably more conformationally flexible than was previously observed [*Biochemistry* **1991**, *30*, 3812−3815; *J. Biol. Chem.* **1993**, *268*, 10095−10101]. All phosphates (as well as one sugar moiety) that are not rigidified by intercalation of the drug molecule are found to adopt two or more distinct conformations. Furthermore, the high quality of the data enabled us to find double conformations for several waters associated with the flexible phosphates or with flexible groups of the daunomycin molecule. These present findings suggest that the crystal lattice still allows for a certain conformational freedom of the DNA in the crystal. This freedom refers to the multiple conformations observed in the group of final models resulting from an NMR structure determination.

## Introduction

For any sequence of bases along DNA, the global conformation of the double stranded form of the polymer is usually predefined in one of the helical families. However, it is the sequence-dependent structural variations that provide the required specificity for DNA recognition by proteins and drugs. Hence, it is the fine details of base, sugar, and phosphate morphology to which all of our attention is attracted in DNA crystallographic studies. Consequently, computational tools were developed for the description of helical parameters and backbone torsion angle morphology in both linear and curved helices.[1−3] Comparison of these carefully determined parameters with the sequences crystallized for a sufficient amount of oligonucleotide structures would then unveil mutual relationships.

That until now only a very simple set of such relationships between the parameters determined and the sequences crystallized could be established could be due to the fact that all presently available DNA structures which are measured in the nonfrozen state, might report an averaged structure. The reported values for the $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, and $\xi$ backbone torsion angles and sugar puckering parameters, although used to describe static sugar−phosphate backbone conformations for room-temperature structures, could be the resulting average of a range of plausible conformers.

Evidence for this view is now provided by the structure determination of the daunomycin−d(CGCGCG) complex (DAU−d(CGCGCG)) at atomic resolution based on a 1.1 Å data set at 100 K, which clearly illustrates that nucleic acids behave as dynamic, flexible molecules, fluctuating between different conformations. The improvement in resolution is accompanied by a clear improvement in the quality of the diffraction data and has resulted in a more precise structure, compared to that of other oligonucleotide−drug complexes in the database. The high data/parameter ratio made for the first time unrestrained block-diagonal least-squares refinement possible in which all atoms are treated anisotropically.

## Materials and Methods

**Crystallization and Data Collection.** Daunomycin was purchased from Rhône Poulenc. Synthesis and purification of the hexanucleotide d(CGCGCG) was performed at the Rega-Institute of the K. U. Leuven. Crystals were grown in sitting drops by vapor diffusion over a period of nine months from an initial crystallization mix containing 4.0 mM ss oligonucleotide, 6.4 mM daunomycin HCl, 10% MPD (4-methyl-2,4-pentanediol), 12 mM spermine tetrachloride, 80 mM NaCl, and 20 mM $BaCl_2$ in pH 7.0 sodium cacodylate buffer equilibrated against 500 $\mu$L 35% MPD at 16 °C. The daunomycin−d(CGCGCG) complex crystallized in space group $P4_12_12$ with unit cell dimensions $a = b =$ 27.81(8) Å and $c = 51.85(8)$ Å. This corresponds to a 2.59% decrease in unit cell volume compared to that at room temperature.[4] The asymmetric unit contains one strand of DNA and one drug molecule. One single crystal was used to collect a 97.3% complete data set from 15.0 to 1.1 Å, using cryocrystallographic techniques, at beamline 5.2*R* of the synchrotron Elettra at Trieste. The data were collected at a wavelength of 1.000 Å, using an 18 cm MAR-research image plate. Two data sets with different crystal-to-detector distances, exposure times, and oscillation ranges yielded 8581 reflections from 15 to 1.1 Å (96.3% above $2\sigma$), $R_{sym} = 4.1\%$. The $R_{sym}$-value did not increase abruptly for the highest angle data. Data were processed with *DENZO*[5]

* Author to whom correspondence should be addressed. Telephone: +32 16 327609. Fax: + 32 16 327990. E-mail: luc.vanmeervelt@chem.kuleuven.ac.be.

† Laboratorium voor Macromoleculaire Structuurchemie, Departement Scheikunde, Katholieke Universiteit Leuven.

‡ Laboratorium voor Analytische Chemie en Medicinale Fysicochemie, Faculteit Farmaceutische Wetenschappen, Katholieke Universiteit Leuven.

(1) Dickerson, R. E. *Nucleic Acids Res.* **1998**, *26*, 1906−1926.

(2) Babcock, M. S.; Olson, W. K. In *Computation of Biomolecular Structures: Achievements, Problems, and Perspectives*; Soumpasis, D. M., Jovin, T. M., Eds.; Springer-Verlag: Heidelberg, 1993; pp 65−85.

(3) Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1998**, *6*, 63−91.

(4) Wang, A. H.-J.; Gao, Y.-G.; Liam, Y.-C.; Li, Y. *Biochemistry* **1991**, *30*, 3812−3815.

(5) Otwinowski, Z. *DENZO. A Film Processing Program for Macromolecular Crystallography*; Yale University Press: New Haven, CT, 1991.

10.1021/ja992180m CCC: $19.00 
Published on Web 12/31/1999

**Table 1.** Summary of the Data Quality and Completeness

| resolution (Å) | unique reflections | mean $I/\sigma(I)$ | $I > 2\sigma(I)$ (%) | completeness (%) | $R_{sym}$[a] |
|---|---|---|---|---|---|
| 15.00−2.98 | 446 | 25.55 | 88.8 | 90.9 | 0.043 |
| 2.37 | 445 | 34.00 | 92.5 | 95.1 | 0.032 |
| 2.07 | 441 | 42.33 | 94.6 | 95.7 | 0.036 |
| 1.88 | 431 | 36.74 | 95.5 | 97.3 | 0.037 |
| 1.75 | 446 | 34.79 | 96.5 | 98.7 | 0.039 |
| 1.64 | 425 | 34.70 | 96.1 | 98.6 | 0.045 |
| 1.56 | 437 | 28.37 | 94.6 | 98.0 | 0.041 |
| 1.49 | 432 | 25.80 | 96.3 | 99.3 | 0.038 |
| 1.44 | 439 | 25.95 | 97.7 | 99.3 | 0.039 |
| 1.39 | 424 | 24.83 | 96.2 | 99.5 | 0.043 |
| 1.34 | 435 | 23.54 | 95.5 | 99.8 | 0.046 |
| 1.30 | 428 | 23.53 | 95.4 | 99.8 | 0.047 |
| 1.27 | 417 | 22.80 | 97.2 | 99.8 | 0.047 |
| 1.24 | 444 | 19.86 | 93.0 | 100.0 | 0.052 |
| 1.21 | 415 | 20.13 | 94.5 | 99.8 | 0.055 |
| 1.18 | 444 | 18.41 | 96.0 | 100.0 | 0.057 |
| 1.16 | 400 | 17.61 | 94.9 | 100.0 | 0.063 |
| 1.14 | 440 | 15.78 | 92.7 | 100.0 | 0.066 |
| 1.12 | 416 | 15.10 | 94.0 | 99.8 | 0.072 |
| 1.10 | 338 | 10.58 | 73.1 | 77.0 | 0.094 |
| overall | 8581 | 25.27 | 93.7 | 97.3 | 0.041 |

[a] $R_{sym} = \sum_{hkl} |I - \langle I \rangle| / \sum_{hkl} \langle I \rangle$.

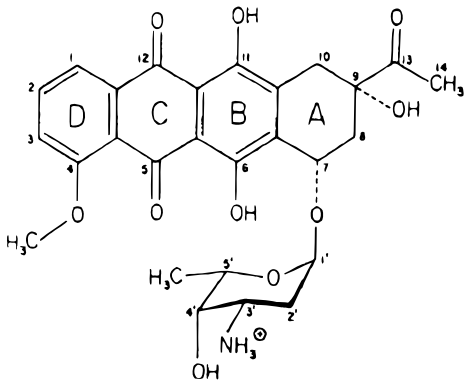



**Figure 1.** Molecular formula and nomenclature of daunomycin $C_{26}H_{30}NO_{10}$. The aglycon chromophore comprises four fused rings labeled *A*−*D*. The daunosamine sugar is attached to the only saturated ring *A*.

including overloads and reduced using *SCALEPACK*.[6] The quality of the data as a function of resolution is presented in Table 1.

Bases are labeled C1→G6 (5′-to-3′ direction) on strand 1, which represents the asymmetric unit. Strand 2, generated by symmetry operation $1 - y$, $1 - x$, $^1/_2 - z$, is labeled C1#→G6# (5′-to-3′). Double conformations are labeled *a* and *b*. The daunomycin molecule is labeled DAU, waters are O1 to O65. The nomenclature for daunomycin is given in Figure 1.

**Refinement.** As the program *SHELXL96* [7] showed up to be the best fitted[8] for oligonucleotide refinement, it was chosen to perform the refinement. In a first step the crystal structure was refined in the conventional way, making use of restraints, in a second step the restraints were removed, and in a third step full-matrix least-squares refinement was tried on both the restrained and the unrestrained model.

**Restrained Conjugate Gradient Refinement.** As the starting model for the 1.1 Å restrained least-squares refinement coordinates and *B*-factors of the 1.5 Å room-temperature daunomycin−d(CGCGCG) structure[4] were used. In the first instance, the geometries of DNA and drug were restrained. Bond lengths and angles within the bases were restrained to target values as specified by Taylor and Kennard;[9] all other chemically equivalent bond lengths and angles, including glycosidic linkages and the sugar−phosphate backbone as well as drug bond lengths and angles, were restrained by similar distance restraints (SADI) which do not require specification of an actual target value. To get an acceptable geometry for the chiral centers C1′ of sugars 5 and 3*b*, the planarity restraints (FLAT) of the attached bases had to be loosened and the chiral volume restraints (CHIV) had to be enforced. This suggests that the ordinarily used FLAT restraint is too strong relative to the CHIV restraint.

Refinement against $F^2$ in *SHELXL96* using all reflections reduced the $R_1$-factor[10] from 32.0% ($\omega R_2 = 67.7\%$)[11] to 21.3% ($\omega R_2 = 52.1\%$), after the addition of 41 water molecules. Double conformations for phosphates 3 and 4, with 0.5 as initial population parameters, were incorporated during the isotropic refinement, as well as a methylene bridge between atom N3″ of the daunosamine moiety and atom N2 of base G4.

After isotropic refinement, all hydrogen atoms were added at calculated positions, and anisotropic refinement was performed with the hydrogen atoms in riding mode with isotropic *U*-values of 1.2 times the $U_{eq}$-values of their parent atoms. During the anisotropic temperature factor refinement, two conformations for phosphate 5 were incorporated. From additional $F_o - F_c$ density double conformations were modeled for several water molecules. Both similarity (SIMU) and rigid bond (DELU) restraints on the $U_{ij}$ -values were applied during anisotropic refinement. Temperature factors of water molecules were restrained to be approximately isotropic (ISOR). This yielded a final $R_1$-value of 11.2% ($\omega R_2 = 30.3\%$) for all 8581 reflections in the 15.0−1.1 Å resolution range including 65 water molecules.

The criteria for identification of water sites applied were that (i) the density distribution around the peak in the $F_o - F_c$ map had to be spherical, (ii) the peak height needed to be greater than 3.5 standard deviations in the initial stages, and (iii) plausible hydrogen-bonding partners with reasonable geometry had to lie within 2.3 Å, and 3.3 Å had to be obtained upon subsequent refinement. If, after refinement, assigned water sites were not in at least $2\sigma$ (in the beginning) and $1.5\sigma$ $2F_o - F_c$ density (at the end), they were removed.

In a first stage, only solvent sites which were regarded as fully occupied, were included in the refinement. The electron density maps clearly showed double conformations for some of the waters surrounding the phosphates. Each of those partially occupied waters made possible hydrogen bonds to only one of the two phosphate conformations, and therefore it was logical to include them in the refinement. In a further stage, however, the maps showed double conformations for other waters, which were clearly present, but which could not be rationalized by looking at hydrogen bond distances. Those waters were split only if the distance between the oxygen atoms was at least 1.04 Å after refinement. This limit for double conformations to be distinguishable was the smallest distance observed between the two oxygens of a split water, for which a double conformation clearly could be rationalized. Each water that tended to a double conformation, but for which subsequent refinement yielded O···O stances less than 1.04 Å, was treated as a single anisotropic water.

No ions or spermine molecules could be identified.

**Unrestrained Conjugate Gradient Refinement.** Since the structure refined without problems and the maps were extremely clear, the removal of gradually more and more restraints was attempted (DELU, SIMU; ISOR for waters; FLAT, CHIV; SADI; DFIX; BUMP), starting from the refined restrained model. No new water molecules were assigned. This resulted in a completely unrestrained refinement with a final $R_1$-value for all 8581 reflections of 10.2% ($\omega R_2 = 27.5\%$).

**Full-Matrix Least-Squares Refinement.** In a next stage full-matrix least-squares refinement was attempted, as for small molecules. By

(10)

$$R_1 = \frac{\sum_{hkl} ||F_o| - |F_c||}{\sum_{hkl} |F_o|} \times 100\% \qquad \omega R_2 = \frac{\sum_{hkl} w_{hkl} ||F_o|^2 - |F_c|^2|}{\sum_{hkl} w_{hkl} |F_o|^2} \times 100\%$$

(6) Collaborative Computational Project, Number 4. *Acta Crystallogr*., Sect. *D* **1994**, *50*, 760−763.

(7) Sheldrick, G. M. *SHELXL96*. Program for Crystal Structure refinement. University of Göttingen: Germany, 1996.

(8) Schuerman, G. S.; Smith, C. K.; Turkenburg, J. P.; Dettmar, A.; Van Meervelt, L.; Moore M. H. *Acta Crystallogr*., Sect. *D* **1996**, *52*, 299−314.

(9) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 3209−3212.

(11) Dauter, Z.; Lamzin, V.; Wilson, K. S. *Curr. Opin. Struct. Biol.* **1997**, *7*, 681−688.

calculating the whole matrix, covariances between parameters were taken into account resulting in a more accurate final model. When the restrained model was used as the starting model, this produced no changes in model and *R*-values because the conjugate gradient least-squares method had been able to reach convergence. When the unrestrained model was used, the changes produced by block-diagonal least-squares refinement were also negligible. (All *x*, *y*, *z* and $U_{ij}$-parameters were refined in three blocks: the first block containing all DNA atoms, the second block containing the daunomycin molecule and residues 5 and 6 of the DNA, and the third block containing all water molecules. In a fourth cycle all *x*, *y*, *z*, but no $U_{ij}$-parameters were refined.)

To obtain estimated standard deviations (esd's) on geometric parameters, one cycle of full-matrix least-squares refinement without restraints, using zero shifts, was performed using either the restrained model or the unrestrained model as a starting model. No damping was applied during this cycle as this would lead to underestimated esd's. The initial data/parameter ratio during *x*, *y*, *z*, *B* refinement of the DNA−drug complex was 13.5, comparable to that of a small molecule. After the modeling of double conformations for phosphate groups 3 and 4, the insertion of the methylene bridge, and the addition of 41 waters, the ratio was still 10.0. Subsequent anisotropic refinement reduced the data/parameter ratio to as much as 5.0, illustrating the validity of our results at this stage. Further incorporation of 15 new water molecules and of double conformations for phosphate 5 and water molecules led to a final ratio of 3.7. This is still 2 times the usual ratio for macromolecules of 1.5−2. Hence, it was reasonable to employ anisotropic, individual temperature factors.[11] It was thanks to this favorable ratio that we were able to remove all restraints normally necessary in macromolecular refinement.

In order to maximize our data/parameter ratio we did not omit 10% of the data for calculation of $R_{free}$ as a cross-validation parameter against overfitting of the data.[12,13] Omitting even only a small percentage of the data to check whether overfitting has occurred, introduces on itself a degree of overfitting and limits the convergence of the refinement even at atomic resolution.[14] All maps were calculated using all of the data with *SHELXPRO*.[7] Examination of the maps was performed using *TURBO FRODO*[15] on a Silicon Graphics workstation. A summary of relevant information pertaining the refinement is given in Table 2.

The final coordinates from the unrestrained refinement have been deposited together with structure factors in the Nucleic Acid Database[16] (code DD0001). The average root mean square (rms) deviations for covalent bonds and covalent angles relative to the standard values for nucleotides[17,18] are 0.031 Å and 2.4°, respectively.

**Table 2.** Summary of Residuals, Data/Parameter Ratios and Temperature Factors[a]

| | |
|---|---|
| $R_{xyzb}$ (%) | 31.3 |
| $R_{xyzbw}$ (%) | 21.3 |
| $R_{anis}$ (%) | 14.8 |
| $R_{double\ conformations}$ | 11.2 |
| $R_{unrestrained}$ | 10.2 |
| no. waters | 65 (10 double conformations) |
| total no. non-H atoms | 182 (23 double conformations) |
| no. reflections | 8581 |
| data/parameters (isotropic) | 10 |
| data/parameters (anisotropic) | 5 |
| data/parameters (double conformations) | 3.7 |
| min $B_{complex}$ (Å$^2$) | 5.20 |
| max $B_{complex}$ (Å$^2$) | 21.03 |
| av $B_{complex}$ (Å$^2$) | 8.51 |
| av $B_{water}$ (Å$^2$) | 28.21 |

[a] $R_{xyzb}$ and $R_{xyzbw}$ are the respective *R*-factors calculated (i) after positional and temperature-factor refinement before the addition of solvent, and (ii) after positional and temperature-factor refinement including all possible water molecules; $R_{anis}$ is the *R*-value after anisotropic temperature-factor refinement. $R_{double\ conformations}$ is the *R*-value including the atoms of the double conformations where they apply. $R_{unrestrained}$ is the *R*-value after unrestrained refinement. Min $B_{complex}$, max $B_{complex}$, av $B_{complex}$, are the minimum, maximum ,and average temperature factors of the DNA−drug complex atoms excluding solvent, respectively; av $B_{water}$ is the average temperature factor for solvent alone.

## Results and Discussion

**Global DNA Conformation.** The self-complementary hexamer d(CGCGCG) adopts the double helical B-DNA conformation with a daunomycin molecule intercalated between the CpG steps at both ends of the duplex (Figure 2). The aglycon chromophore (rings *B* to *D*) of daunomycin skewers the double helix with ring *D* reaching the major groove and the daunosamine sugar lying in the minor groove. The position of the drug is stabilized by direct and solvent-mediated hydrogen bonds to the DNA. Furthermore, a covalent link exists between N2 of guanine G4 and N3″ of daunomycin. This was also seen in the same daunomycin−d(CGCGCG) structure[1] at room temperature and 1.5 Å resolution, but was not expected in our case. However, the MPD probably contained trace amounts of HCHO, causing a covalent cross-linking reaction. The values quoted below pertain to the unrestrained model.

Helical parameters as calculated by the program *RNA*[2] are as usual for this kind of complex. The C·G pairs above the chromophore moiety of the drug are buckled by 14.7°, while the G·C pairs below the chromophore have a larger buckle of 16.8° in the opposite direction. The twist angle at the intercalation step is 36.1°. Other steps have an average twist angle of 31.6°. The overall unwinding angle due to intercalation of the daunomycin drug is 4.5°.

Backbone torsion angles are illustrated in Table 3. Most of the torsion angles of conformation *a* (for a description of the double conformations, denoted *a* and *b*, see further) fall well within the ranges typical of B-DNA. The same is true for the torsion angles of conformation *b* of phosphate groups 4 and 5. The backbone conformation of phosphate 3*b* on the contrary, falls not at all in the ranges typical to B-DNA.

In general the B-DNA backbone conformations can be categorized in two conformations called $B_I$ and $B_{II}$. The $B_I$ conformation is defined[19] as having the torsion angle $\epsilon$ between 120° and 210°, and the angle $\zeta$ in the range 235°−295°; while the $B_{II}$ conformation has an $\epsilon$ of 210°− 300° and an $\zeta$ of 150°−210°. Both conformations of both phosphate groups 4 and 5 have the $B_I$ conformation. Phosphates 6 and 3*a* adopt the $B_{II}$ conformation. Phosphate 3*b* adopts almost the $B_{II}$ conformation, while phosphate 2 does not belong to either of the conformational groups.

**Thermal Parameters.** The mean *B*-values for the atoms in base, sugar, and phosphate groups of each nucleotide of d(CGCGCG) are shown in Figure 3. All *B*-values are very low, averaging 14.3 and 8.5 Å$^2$, with and without solvent molecules,

(12) Brünger, A. T. *Nature* **1992**, *355*, 472−474.
(13) Brünger, A. T. *Acta Crystallogr.*, Sect. *D* **1993**, *49*, 24−36.
(14) Sevcik, J.; Dauter, Z.; Lamzin, V. S.; Wilson, K. S. *Acta Crystallogr.*, Sect. *D* **1996**, *52*, 327−344.
(15) Jones, T. A.; Cambillau, C. *Silicon Graphics Geometry Partners Directory, 89.* Silicon Graphics Inc.: Mountain View, CA, 1989.
(16) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751−759.
(17) Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S. H.; Olson, W. K.; Berman, H. M. A. *J. Am. Chem. Soc.* **1996**, *118*, 519−529.
(18) Clowney, L.; Jain, S. C.; Srinivasan, A. R.; Westbrook, J.; Olson, W. K.; Berman, H. M. *J. Am. Chem. Soc*. **1996**, *118*, 509−518.
(19) Schneider, B.; Neidle, S.; Berman, H. M. *Biopolymers* **1997**, *42*, 113−124.

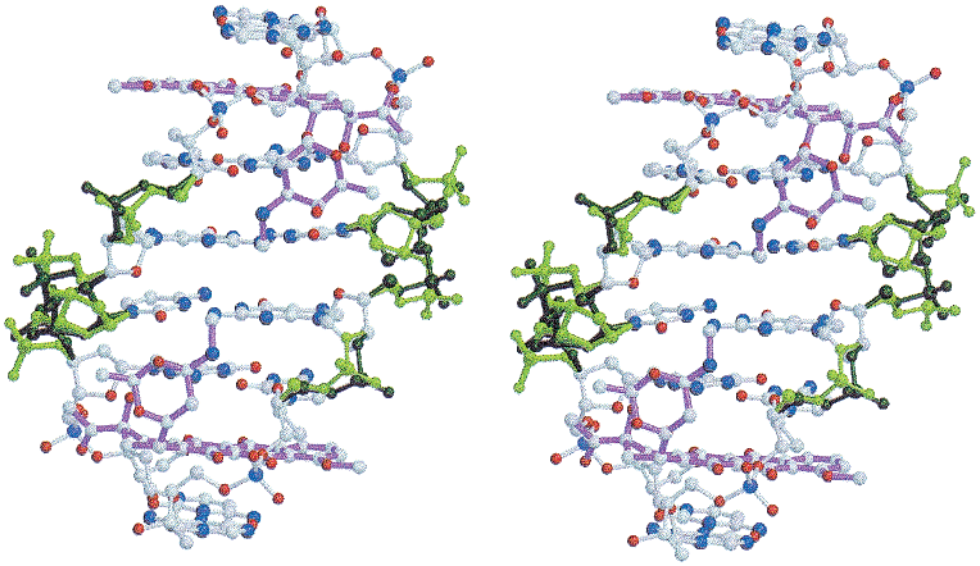


**Figure 2.** Stereoview of the daunomycin−d(CGCGCG) structure. All DNA atoms are colored by atom type: carbons (gray), nitrogens (blue), oxygens (red), phosphorus atoms (purple); the bonds of daunomycin are shown in pink. Conformations *a* and *b* are colored dark and light green, respectively. From top to bottom, the phosphates in the right backbone strand are running from P2 to P6.

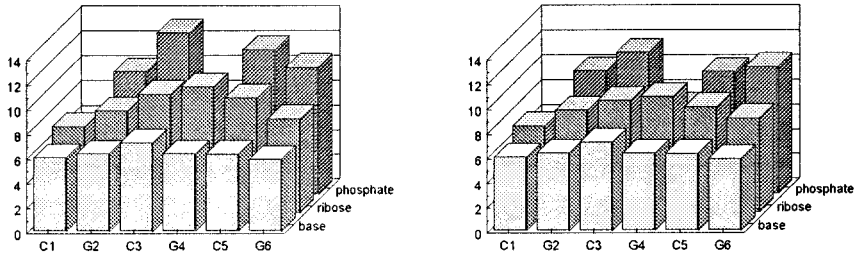



**Figure 3.** Bar diagram representing the thermal vibrations of (left) the *a* conformation and (right) the *b* conformation of the unrestrained daunomycin−d(CGCGCG) model. The thermal $B_{eq}$-value is averaged for atoms of the base, the sugar (excluding O3′ and O5′), and the phosphate group (including O3′ and O5′) of each residue.

**Table 3.** Backbone and Glycosidic Torsion Angles (deg)[a]

| | α | *β* | *γ* | *δ* | *ϵ* | *ζ* | *χ* |
|---|---|---|---|---|---|---|---|
| C1 | | | 53.2 | 141.7 | 232.1 | 286.2 | 207.3 |
| G2*a* | 289.9 | 172.6 | 45.0 | 146.5 | 219.3 | 170.3 | 268.5 |
| G2*b* | | | | 131.3 | 283.1 | 142.2 | |
| C3*a* | 301.5 | 137.4 | 55.1 | 124.5 | 186.9 | 272.5 | 216.1 |
| C3*b* | 61.6 | 224.0 | 186.8 | 89.5 | 199.1 | 286.1 | 199.7 |
| G4*a* | 269.6 | 197.3 | 66.2 | 122.5 | 183.8 | 275.2 | 257.2 |
| G4*b* | 304.8 | 181.5 | 54.7 | 152.5 | 193.2 | 267.7 | |
| C5*a* | 293.0 | 180.5 | 46.7 | 147.4 | 254.5 | 171.4 | 273.8 |
| C5*b* | 282.7 | 169.5 | 51.7 | 132.5 | | | |
| G6 | 283.2 | 178.7 | 51.3 | 145.3 | | | 279.2 |

[a] Pseudorotation angles were calculated using the program *NEWHELIX93* distributed by R. E. Dickerson[1].

respectively. One common striking feature appears in the *a* as well as the *b* conformations: the average *B*-factor of phosphate 4 is lower than that of all the other phosphates and lower than the sugar of base 4. This could be a consequence of the close contacts existing between phosphate 4 and atom O3′ of residue 6 of a symmetry related molecule. This atom forms direct hydrogen bonds with atoms O2P and O1P of phosphate 4*b* (respectively 2.63 and 3.28 Å) and with atom O1P of phosphate 4a (2.76 Å). All of the other phosphates are more solvent-exposed, consistent with their higher *B*-factors. Interestingly, phosphate 4, although in close contact with a symmetry equivalent DNA strand, can still adopt two distinct conformations.

Temperature factors of the waters range from 8.6 to 85.2 Å$^2$ and average at ∼28.2 Å$^2$.

**Hydration Network of DNA.** The present structure displays an extensive hydration network connecting symmetry-related molecules. Between the coaxially packed DNA fragments, water molecules are observed to mediate indirect contacts between closely spaced molecules by bridging backbone oxygen atoms of the phosphodiester linkage. Table 4 describes the water-mediated connections of oxygen atoms of all six phosphate groups to phosphate oxygens, base nitrogen atoms, or drug oxygen atoms of symmetry-related molecules. Also, the terminal 5′- and 3′-hydroxyl groups of the DNA strand are connected to symmetry-related molecules, the latter one even making a direct hydrogen bond to a neighboring phosphate 4 group.

One asymmetric unit contains 65 water molecules, indicating that the double helix complexed with two daunomycin molecules is hydrated by 130 solvent molecules. Of the total number of water molecules in the asymmetric unit, 36.5 are located in the first coordination-shell, 24.5 are second-shell waters, and only 4 belong to the third hydration-shell. While the minor groove is almost completely occupied by the amino sugar of the daunomycin drug, ring *D* of the aglycon moiety protrudes into the major groove. This may be the reason almost the same amount of water is found in both grooves (21 in the minor and 24 in the major groove). Twenty water molecules are attached to phosphate groups and are situated in none of the grooves.

**Hydrogen Atoms: Quality of the Maps.** In general, X-ray crystallography does not provide accurate structural information

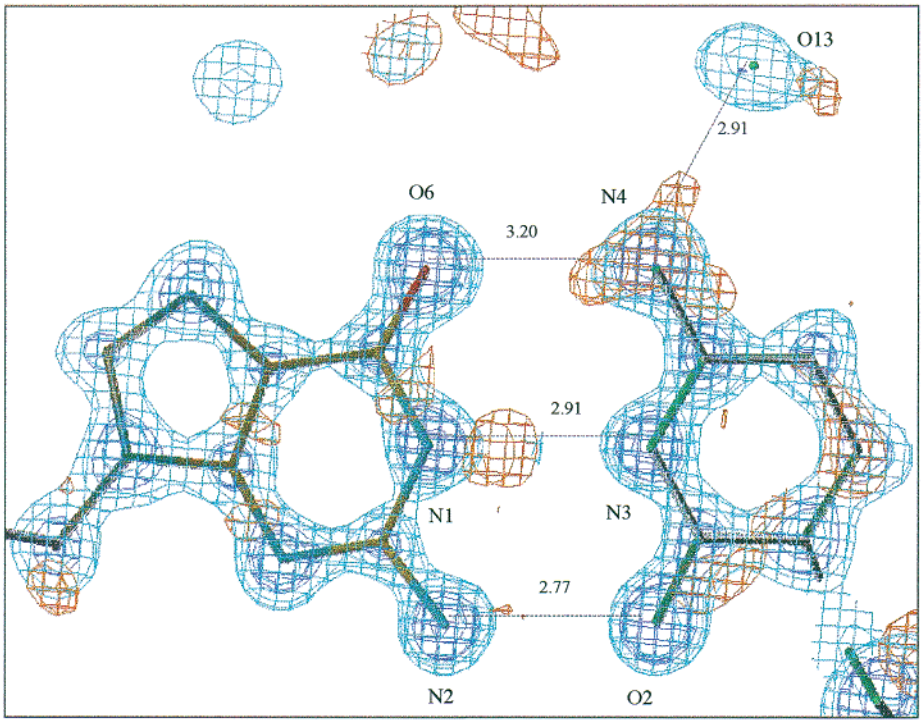



**Figure 4.** $F_o - F_c$ map (contouring level 0.6e$^-$/Å$^3$, shown in brown) on top of $2F_o - F_c$ map (contouring levels 3.0e$^-$/Å$^3$ and 1.0e$^-$/Å$^3$ respectively shown in middle blue and light blue), showing hydrogen atoms in base pair G2·C5#.

**Table 4.** Hydrogen Bonding Network Connecting Symmetry Related Molecules

| hydrogen bond acceptor | hydrogen bond donor (in symmetry-related molecule) | number of mediating water molecules | name of mediating water molecule |
|---|---|---|---|
| C1O5′ | C3O1P*b* | 1 | O17*a* |
| G2O1P | G4O2P*a* | 2 | O46, O24 |
| | G4O2P*b* | 2 | O46, O24 |
| G2O2P | G4O2P*a* | 1 | O10*a* |
| | G4O2P*b* | 1 | O10*b* |
| C3O1P*b* | C1O5′ | 1 | O17*a* |
| | D31O12 | 2 | O17, O15 |
| C3O2P*a* | G2N7 | 3 | O45*a*, O20, O11 |
| C3O2P*b* | G2N7 | 2 | O20, O11 |
| G4O1P*b* | G6O3′ | 0 | |
| G4O1P*b* | G6O3′ | 0 | |
| G4O2P*b* | G6O3′ | 0 | |
| C5O1P*a* | G6O3′ | 2 | O29*a*, O42 |
| C5O1P*b* | G6O3′ | 2 | O29*b*, O42 |
| C5O2P*a* | C1N4 | 3 | O28, O43, O38 |
| C5O2P*b* | C1N4 | 3 | O28, O43, O38 |
| G6O2P | G4O1P*a* | 2 | O36, O27*a* |
| | G4O1P*b* | 2 | O36, O27*b* |
| G6O3′ | G4O1P | 0 | |
| | G4O2P | 0 | |

on hydrogen atoms. Even for small molecules it is often difficult to locate them in difference electron density maps, unless they are involved in a hydrogen-bonding pattern. The high quality of present electron density maps, however, is proven by the fact that they reveal the presence of hydrogen atoms. As illustrated in Figure 4, the $F_o - F_c$ map clearly shows the presence of three hydrogen atoms in base pair G2·C5#: one connecting nitrogen N1 of guanine G2 and nitrogen N3 of cytosine C5# and the two hydrogen atoms of amino group N4 of cytosine C5# of which one makes a hydrogen bond with oxygen O6 of G2 and one forms a bond with water molecule O13. Other base pairs as well as some sugar moieties contain similar information in their $F_o - F_c$ maps. In the daunomycin molecule the hydrogen atom of hydroxyl 11 which is involved in an intra-ring hydrogen bridge with atom O12 is also visible in the $F_o - F_c$ map prior to anisotropic refinement. Having density maps of a quality comparable to those of small molecules, we may expect to be able to distill extra features out of this daunomycin−d(CGCGCG) structure. These features, recalling small molecule structures, are treated in the next two sections.

**Static and Dynamic Disorder: Conformational Flexibility.** The high quality electron density maps emerging contain distinct lobes of electron density with short separations for all phosphates not situated at an intercalation step, indicating that double conformations exist for those phosphates. It is only the two phosphates at both ends of the hexamer that have one unique conformation, probably due to the fact that both of them are situated at the intercalation step where the helix is unwound, resulting in a decreased flexibility.

Figure 5a shows how two distinct lobes of electron density fit two completely different conformations for phosphate group 3, which are generated by a rotation over ∼180° around the O5′−C5′ and C3′−O3′ bonds. The phosphorus atoms of the two conformations lie 1.61 Å apart. The different α, $\beta$, $\gamma$, and $\delta$ angles thus generated probably cause sugar 3 to adopt two distinct conformations as well (Figure 5b). Figure 5c illustrates how phosphate 4 also shows two clearly distinct conformations with the phosphorus atoms displaced over 1.04 Å, causing less pronounced changes in torsion angles. At phosphate 5 the map shows one long-drawn lobe of electron density, indicating the multiple conformations for this phosphate group. The whole phosphate group can move from left to right over a distance of 0.80 Å, as can be seen in Figure 5d. This time, however, no separate electron clouds are visible so that it seems that the phosphate group can adopt all intermediate conformations as well. Both phosphates at the ends of the double helix adopt only one conformation. The first one, phosphate 2, is very well

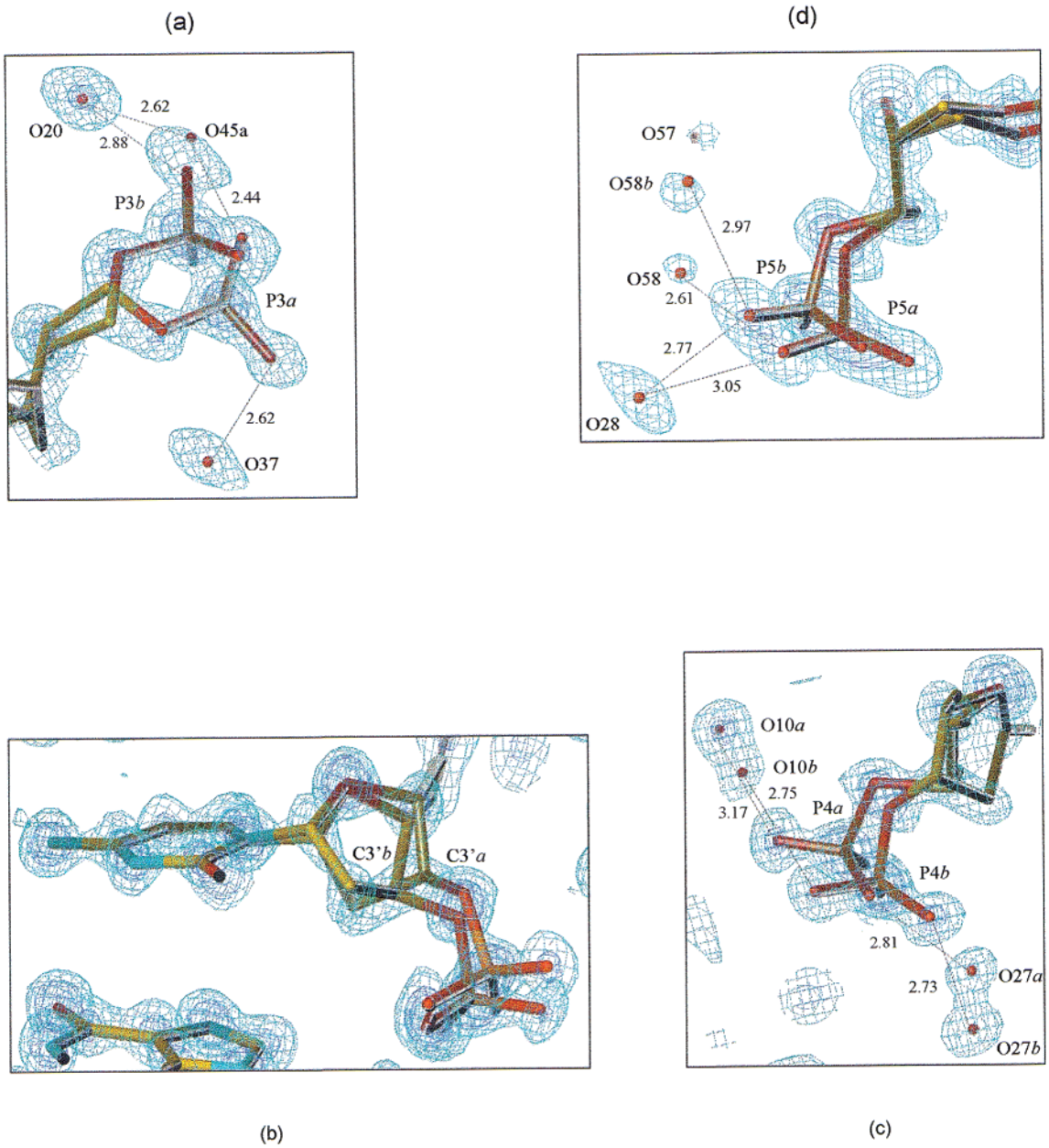



**Figure 5.** Double conformations for phosphate groups and sugar moiety with surrounding waters superimposed on sections of the $2F_o - F_c$ electron density map (contouring levels 4.0e$^-$/Å$^3$ and 1.5e$^-$/Å$^3$ shown in respectively dark- and light blue). (a) phosphate group 3, (b) sugar moiety 3, (c) phosphate group 4**,** and (d) phosphate group 5. The water molecules O17, O29, and O48 are not visible in the figures.

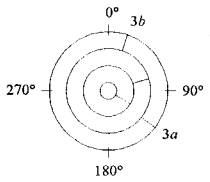
determined, with small, almost isotropic B-factors. The last one, phosphate 6, shows rather dynamic disorder, with anisotropic B-factors indicating motion in mainly one direction. However it was not sensible to model two distinct conformations for this phosphate, since they would be separated over only 0.58 Å. The alternate conformations occur each time in nearly equal populations, implicating that in this DNA-drug complex the DNA backbone presents itself as flexible base-linker able to adopt different, probably equi-energetic, conformations.

The quality of the maps allows us to model double conformations not only for the backbone but also for the water structure which follows this movement. Partially occupied water sites have been identified in those areas where flexibility in the backbone induces concomitant changes in the local solvent structure, of which some are shown in Figure 5. The conformational flexibility of phosphate group 3 is reflected in a double conformation for water molecule O48. This has in its turn the consequence that water O17 can only exist with conformation *b* of the DNA. In the minor groove, for the *a* conformation of atom O2P, the alternative (nonfilled O2P*b*) site is occupied by water molecule O45*a* (visible in Figure 5a).

The shift of phosphate group 4 is in the minor groove accompanied by a double conformation for water O10, and in the major groove by a double conformation for water O27, as shown in Figure 1c. The fact that phosphate group 5 can adopt all conformations between *a* and *b* leads to double conformations for O29 and for O58 in the minor groove. Besides double conformations, several waters show anisotropy. In the major groove O28 (Figure 5d) and O33 are examples of such waters, they form a bridge between the continuously moving phosphate group 5 and the anisotropic phosphate 6.

**Sugar Puckering: Comparison with Previous Structures.** Both previously reported similar structures, DAU−d(CGCGCG)[4] at 1.5 Å resolution and DAU−d(CG$^{ara}$CGCG),[20] display one

(20) Zhang, H.; Gao, Y.-G.; Van der Marel G. A.; Van Boom J. H.; Wang, A. H.-J. *J. Biol. Chem.* **1993**, *268*, 10095−10101.




**Pseudorotation angles *P* (°) and sugar conformations for the unrestrained model**

| residue name | *P* | sugar conformation |
|---|---|---|
| C1 | 165.6 | C2′-*endo* |
| G2*a* | 141.6 | ~C1′-*exo* |
| G2*b* | | |
| C3*a* | 126.9 | C1′-*exo* |
| C3*b* | 19.1 | C3′-*endo* |
| G4*a* | 165.2 | C2′-*endo* |
| G4*b* | | |
| C5*a* | 144.7 | ~C2′-*endo* |
| C5*b* | | |
| G6 | 183.7 | ~C3′-*exo* |

Pseudorotation angles were calculated using the program *NEWHELIX93* distributed by R. E. Dickerson[1].

**Figure 6.** Wheel representing the pseudorotation angle *P* for sugar 3 as found in DAU−d(CG$^{ara}$CGCG) (inner circle), in DAU−d(CGCGCG) at 1.5 Å, room temperature (middle circle), and in DAU−d(CGCGCG) at 1.1 Å, 100 K (outer circle).

single conformation for the sugar of residue 3. The 1.5 Å resolution room-temperature structure reports an O4′-*endo* sugar puckering for the C3 deoxyribose (pseudorotation angle $P$ = 73°), while the structure of DAU−d(CG$^{ara}$CGCG) reports a higher $P$ of 120° for the arabinosyl sugar of cytosine 3. This high $P$-value was ascribed to the extra O2′ in the 'up' position at C2′ of the arabinosyl sugar. Contrarily, the present 1.1 Å structure at 100 K gives two possible conformations for sugar 3 (Figure 5b). Sugar 3*a* has $P$ = 127°, conform to the pseudorotation angle in DAU−d(CG$^{ara}$CGCG), which therefore does not seem to be caused by the O2′ at C2′. Sugar 3*b* has $P$ = 19°. It is striking that the pseudorotation angle of 73°, found in the 1.5 Å DAU−d(CGCGCG) structure, is the average of both sugar conformations reported here (Figure 6). This finding may point out a more general characteristic of room-temperature structures: while room-temperature data collections may provide us with an average view, cryocooling enables us to get a far more detailed picture.

Furthermore, in the 1.5 Å resolution room-temperature structure all DNA sugar puckers remain in the C2′-*endo* family ($P$ ranging from 150° to 176°) except for the cited sugar of cytosine 3. The present structure reveals more variations in the pseudorotation angle, ranging from 19.1° to 183.7° in the unrestrained model. Thus, the previous observation that all sugars except that of residue 3 remain in the C2′-*endo* family does not hold any more. Besides the C2′-*endo*, the C1′-*exo* for C3*a* and G2 and the C3′-*exo* for C6 are also observed, while C3*b* tends to the C3′-*endo* conformation. Thus, although the idea of a flexible sugar−phosphate backbone was supported already by its higher temperature factors in crystal structures, this crystal structures illustrates for the first time a double conformation for a longer backbone region including not only three phosphate groups but also one sugar moiety.

Root mean square differences between conformation *a* and *b* of the present 1.1 Å and the 1.5 Å resolution structure DAU−d(CGCGCG) are respectively 0.223 and 0.354 Å. This difference is large ($>3\sigma$) compared to the mean esd on positions found in our structure. The main differences between both structures are in the phosphate backbone, exactly at those phosphate groups for which now distinct conformations are observed. The phosphate conformations observed in the 1.5 Å study are intermediate between the *a* and *b* conformations reported here for phosphate groups 4 and 5. Only phosphate 3 was reported to be most close to our *a* conformation.

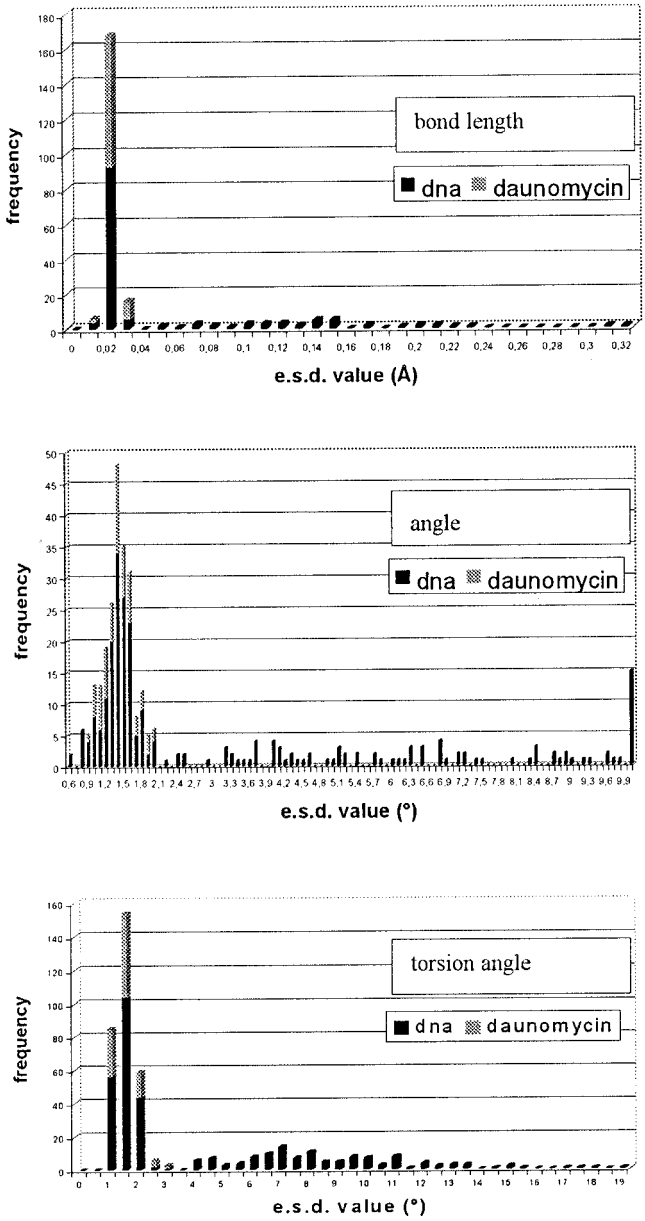



**Figure 7.** Histogram of esd's on (top) bond lengths, (middle) angles, and (bottom) torsion angles in the unrestrained daunomycin−d(CGCGCG) model.

**Precision.** At the end of the unrestrained least-squares refinement one cycle of full-matrix least-squares without restraints (zero damping and zero shifts) was performed, to take into account all covariances between parameters. From the inversion of the full matrix good estimates of the standard deviations on all atomic positional and thermal parameters are calculated. From the former, esd's on all geometrical parameters can be derived. Estimated standard deviations on geometrical parameters after restrained and unrestrained refinement are similar; therefore, only the latter ones are quoted (Figure 7).

The esd's on bond lengths between fully occupied atoms are very low, with a mode (i.e., the most frequently occurring value) at 0.02 Å, but partially occupied atoms have much larger esd's, ranging up to 0.32 Å. Estimated standard deviations on angles in the complex show a mode at 1.4° and average at 3.3°. Fifteen double conformation angles have esd's larger than 10°. The most frequently occurring value of the esd's on torsion angles in the unrestrained DNA model is 2°. The average torsion angle esd is 4.2°, but some outliers with esd's up to 18.8° exist. In the

daunomycin molecule the esd's on torsion angles are lower, averaging at a value as small as 1.5°.

The present low-temperature, high-resolution structure offers the nice possibility of getting estimated standard deviations on the hydrogen bonds in the Watson−Crick base pairs. These esd's all fall well within a narrow range from 0.014 to 0.018 Å and average at the very low value of 0.017 Å. Remarkably, this value is even lower than most of the esd's on covalent bonds. This is a consequence of the positional esd's in the bases being the lowest of the whole structure, which is logical as the complementary base pairs are definitely the most accurate known regions of the double helix.

While it is important to establish the general precision of the molecular structure, from the point of view of a biochemist, it is not the esd's, but the geometric parameters themselves that are more interesting, especially now that they are known with such a high precision and that after unrestrained refinement. For example, it is interesting to look at hydrogen bonds in Watson−Crick pairs and to compare them with the generally accepted values given by Saenger.[21] The values we find are 2.955, 2.933 and 2.764 Å (with an esd of 0.017 Å) for respectively the N4−O6, N3−N1, and O2−N2 hydrogen bonds in a C·G pair. The distances given by Saenger are respectively 2.91, 2.95, and 2.86 Å. Hence, the C·G pairs in our structure seem to have a little larger opening (definition, see ref 22) than those given by Saenger.

The angles between donor atom−hydrogen atom−acceptor atom in the Watson−Crick pairs average 170.9(5)°, close to the linear arrangement observed for a good hydrogen bond by small-molecule crystallography.

These values, however, are based on only a few bond lengths and angles and must be treated with caution. More atomic resolution structures are necessary to study in great detail the hydrogen-bonding geometry in oligonucleotides, which was until now only performed by small-molecule crystallographic studies on the nucleotide building blocks.

Estimated standard deviations on hydrogen bonds between respectively base and phosphate atoms, on the one side, and water molecules, on the other side, average at 0.025 Å and 0.06 Å (with modes at respectively 0.02 and 0.04 Å). These values reflect the very different mobility of bases and phosphates in the DNA helices. While the former are stacked very compactly, forming a rigid core, the phosphate backbone, on the contrary, is very flexible. When we add to this the esd's on hydrogen bonds between only water molecules in the hydration network, averaging 0.06 Å (with a mode at 0.06 Å), we see no difference between esd's on water−water hydrogen bonds and those on phosphate−water hydrogen bonds. This confirms the general idea that the mobility behavior of the phosphates resembles that of the hydration-shell waters more than that of the bases.

When comparing the esd's on hydrogen bonds in general with those on covalent bonds, we find no significant differences. This is a very nice illustration of the importance of the hydration network around biological macromolecules. It proves that a molecule in the biological fluid must not be seen on its own but rather forms a larger entity with its tightly attached water structure.

## Conclusions

Today more than 20 DNA−anthracycline complexes have been crystallized and solved by X-ray diffraction techniques.[17] All of these data collections were carried out at room temperature, with the exception of one at 4 °C. The highest resolution reached for anthracycline complexes with natural sequences is 1.5 Å; with modified sequences it is 1.2 Å. Although the limit of resolution for these complexes was already high, none of these structures reports the presence of alternate conformations.

In contrast, the structure of the daunomycin−d(CGCGCG) complex presented here reveals, at first, additional details of the molecular structure and the solvent network. Moreover, the extremely high resolution of 1.1 Å of the present structure (increasing the number of data), together with the fact that the asymmetric unit contains only one strand of DNA and one drug molecule (reducing the amount of parameters to be determined) produces a data/parameter ratio comparable to that of small molecules. Therefore, for the first time unrestrained least-squares refinement of an oligonucleotide became possible.

**Bridging the Gap between X-ray and NMR Results.** NMR specialists prefer to report an assembly of final structures (resulting from a structure determination) rather than one mean global structure, which they argue would result in a loss of information. For oligonucleotides, for example, such an assembly can consist of 15 structures. Progress in NMR techniques, however, reduces this number. On the other hand, the high resolution of the present X-ray structure also allows us to distinguish in the electron density maps different conformations for the major part of the DNA backbone. This is a general trend in X-ray crystallography: atomic resolution structures start to reveal multiple conformations. Thus, while NMR spectroscopic studies develop toward the reporting of a smaller number of conformations, X-ray crystallographic studies start to report a larger number of conformations. Probably in the future both methods will be able to report the same, limited number of possible conformations and hence to bridge the gap between X-ray and NMR results.

**Crystal Packing Effects.** Interesting to note at this point is that the crystal lattice seems not as tyrannical as is sometimes claimed.[23] Crystal packing forces, at least in this structure in $P4_12_12$, are not able to impose the fine structural features upon the DNA crystal structure but, on the contrary, allow for considerable flexibility. This is illustrated by phosphate 4, which, although in close contact with a symmetry equivalent DNA strand, still can adopt two distinct conformations. While the formation of hydrogen bridges in this close contact area is able to reduce the dynamic disorder (*B*-factors), crystal packing forces are not able to lower the static disorder of these phosphates. The only factor that is able to reduce the conformational flexibility of the backbone is the unwinding of the helix at the intercalation steps.

Even stronger, the observed flexibility of the sugar−phosphate backbone in this structure could just be a consequence of the crystal packing. Individual helices are stacked on top of each other, forming long pseudo-continuous helices that are ordered coaxially in the unit cell, creating large water channels between them. These parallel channels run through the whole crystal, thus allowing for considerable solvent movement and, hence, also backbone movement. It was even found that water molecules with higher temperature factors (O46, O51, O61, O62, O63, being the best examples) were located at the borders of those long solvent channels. They form the transition region between the well-defined solvent network and the diffuse solvent in the channels (less than 0.40 $e^-/Å^3$ in the last $F_o - F_c$ map).

**Conformational Fluctuations.** The structure determination of DAU−d(CGCGCG) at atomic resolution and at 100 K

(21) Saenger, W. In *Principles of Nucleic Acid Structure;* Canter, C. R., Ed.; Springer-Verlag: New York, 1984.

(22) EMBO Cambridge Workshop. *EMBO J.* **1989**, *8*, 1−4.

(23) Dickerson, R. E.; Goodsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3579−3583.

illustrates that lowering the temperature results in lowering the dynamic motion of atom groups and in freezing out the different static conformers. Hence, the $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, $\xi$ torsion angles and puckering parameters, although used to describe static sugar−phosphate backbone conformations for room-temperature structures, could be the resulting average of a range of plausible conformers. Furthermore, one should take into account their esd's as well. In most room-temperature structures these esd's are rather large, allowing for considerable deviation from the calculated parameters. Thus, for getting insight in the three-dimensional structure of macromolecules in living cells the use of cryogenic techniques, although this might seem contradictory, together with high energy synchrotron radiation, are of great help and should become standard techniques.

Conformational fluctuations might be important in protein−DNA interactions. Interactions between biological molecules require a sufficient flexibility from both partners in order to allow for optimal interactions. The interactions of protein side chains with DNA, for example, are only possible after DNA unwinding or kinking in the case of side chain intercalation or changing the DNA backbone conformation and hydration in the case of groove binding. This DAU−d(CGCGCG) structure shows that even after the tight complexation of the daunomycin drug the DNA b ackbone surrounding the interaction site keeps this flexibility. This could suggest constant local adaptions of the interaction site conformation and hydration to occur in order to optimize the (dynamic) complexation of biological macromolecules.

**Acknowledgment.** This work was partly supported by the Human Capital and Mobility Program of the European Community and by the Fund for Scientific Research (Flanders). We thank the staff of the Elettra X-ray diffraction beamline at Trieste for assistance, Professor A. Van Aerschot and Guy Schepers (Rega-Institute, K.U. Leuven) for synthesizing the oligonucleotide and Professor H. Reynaers (K.U. Leuven) for continuous support.

JA992180M