# Molecular conformation of 2′-deoxy-3′,5′-di-*O*-acetyl guanosine. Crystal structure and high resolution proton nuclear magnetic resonance investigations

LEO H. KOOLE[1] AND HENK M. BUCK

*Department of Organic Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

AND

JAN A. KANTERS AND ARIE SCHOUTEN

*Laboratory of Crystal and Structural Chemistry, State University of Utrecht, Utrecht, The Netherlands*

Received November 6, 1987[2]

LEO H. KOOLE, HENK M. BUCK, JAN A. KANTERS, and ARIE SCHOUTEN. Can. J. Chem. **66**, 2634 (1988).

2′-Deoxy-3′,5′-di-*O*-acetyl guanosine crystallizes in the triclinic space group $P1$ ($C_1^1$, No. 1), and the cell dimensions are $a = 8.643(1)$ Å, $b = 10.122(1)$ Å, $c = 10.391(1)$ Å, $\alpha = 87.04(1)°$, $\beta = 73.58(1)°$, $\gamma = 72.37(1)°$, $V = 830.4(2)$ Å$^3$; $Z = 2$ molecules per cell. Least-squares refinement converged at $R = 0.031$ for 3450 observed reflections. The asymmetric unit consists of two independent molecules. The guanine bases are linked via N(1)—H...N(7) and N(2)—H...O(6) hydrogen bonds to form a virtually planar system. Moreover, the conformational preferences of the title compound in DMSO-$d_6$ solution have been determined with 300 MHz $^1$H NMR. It is found that the X-ray structure and the solution conformation are essentially similar. The C(4′)—C(5′) linkage resides in the *gauche*$^+$ ($g^+$) or *trans* ($t$) rotamers in the solid state, while an approximately equal distribution over $g^+$, $t$, and *gauche*$^-$ is found in solution.

LEO H. KOOLE, HENK M. BUCK, JAN A. KANTERS et ARIE SCHOUTEN. Can. J. Chem. **66**, 2634 (1988).

La désoxy-2′ di-*O*-acétyl-3′,5′ guanosine cristallise dans le groupe d'espace triclinique $P1$ ($C_1^1$, No. 1) avec $a = 8{,}643(1)$, $b = 10{,}122(1)$ et $c = 10{,}391(1)$ Å, $\alpha = 87{,}04(1)$, $\beta = 73{,}58(1)$ et $\tau = 72{,}37(1)°$, $V = 830{,}4(2)$ Å$^3$ et $Z = 2$. La structure a été résolue par la méthode des moindres carrés (matrice entière) jusqu'à une valeur de $R$ de 0,031 pour 3450 réflexions observées. L'unité asymétrique comporte deux molécules indépendantes. Les bases guanines sont liées par le biais de liaisons hydrogènes N(1)—H...N(7) et N(2)—H...O(6) et elles forment un système virtuellement plan. De plus, opérant en solution dans le DMSO-$d_6$ et faisant appel à la RMN du $^1$H à 300 MHz, on a déterminé les préférences conformationnelles du composé mentionné dans le titre. On a trouvé que la structure déterminée par diffraction des rayons-X possède essentiellement la même conformation qu'en solution. A l'état solide, la liaison C(4′)—C(5′) se retrouve dans les rotamères *gauche*$^-$ ($g^+$) ou *trans* ($t$); en solution, on retrouve une distribution approximativement égale des rotamères $g^+$, $t$ et *gauche*$^-$.

[Traduit par la revue]

## Introduction[3]

Recent literature covers a substantial number of X-ray crystallographic studies on acetylated DNA or RNA nucleosides (1–9). The interest in these compounds stems primarily from their remarkable structural features. It has been found that acetylated nucleosides can serve as unique models to study different modes of hydrogen-bonded base pairing, which must be considered as alternatives for the conventional antiparallel coupling of complementary bases according to the Watson and Crick (10) model (adenine (A) – thymine (T), or guanine (G) – cytosine (C)). For example, it was found that 3′,5′-di-*O*-acetyl thymidine crystallizes in a *parallel* base-paired conformation, in which the T bases are linked via two N(3)—H...O(4) hydrogen bonds (6, 11, 12). We have recently shown that the possibility of parallel T–T base-pair formation may have a more general significance. Exclusive methylation of the phosphate groups in the mononucleotide d(pTp), and in the oligonucleotides d($T_2$), d($T_3$), d($T_4$), d($T_6$), and d($T_8$) results in the formation of a remarkably slim parallel double helical structure in aqueous solution (11–13). The parallel strands in this structure are linked via T–T base pairs as first observed in the crystal structure of 3′,5′-di-*O*-acetyl thymidine. Very recent UV hyperchromicity experiments with unmodified d($T_8$) have provided preliminary evidence for the formation of an analogous double helix, if one equivalent of the cationic protein polylysine is present (13). Most likely, complexation of polylysine with the negatively charged phosphodiester groups in d($T_8$) leads to a substantial reduction of the effective negative charge on the phosphate groups. This is essential for the formation of T–T base pairs (11–13).

To investigate self association of the other DNA bases, we have previously studied the structure of 2′-deoxy-3′,5′-di-*O*-acetyl adenosine (1). It was observed that the A base does not form parallel A–A base pairs but, instead, any array of one-dimensional ribbons in which the bases are linked via two crystallographically-independent hydrogen bonds: N(6)—H...N(1) and N(6)—H...N(7). A highly comparable hydrogen bonding scheme was found by Wilson *et al.* for the crystal structure of the acetylated RNA counterpart structure, 2′,3′,5′-tri-*O*-acetyl adenosine (3).

In this paper we wish to report a structural study on the acetylated DNA nucleoside 2′-deoxy-3′,5′-di-*O*-acetyl guanosine (Fig. 1). The X-ray structure reveals that the asymmetric unit consists of two independent molecules, **I** and **II**. The planar guanine fragments of **I** and **II** are connected via N(1)—H...N(7) and N(2)—H...O(6) hydrogen bonds to form a nearly planar system with the two guanine rings at an angle of 5.0(1)°. Furthermore, a comparison is made of the X-ray data of the title compound, and the solution conformation as determined with 300 MHz $^1$H NMR spectroscopy.

## Experimental section

*Preparation*

Acetic anhydride (1.83 g, 17.92 mmol) was added to a stirred suspension of dry 2′-deoxyguanosine (2.00 g, 7.49 mmol) and

[1]To whom all correspondence should be addressed.

[2]Revision received May 31, 1988.

[3]Nomenclature in this work follows the IUPAC-IUB recommendations on nucleotide conformational nomenclature. See: "Abbreviations and symbols for the description of conformation of polynucleotide chains", Recommendations 1982. Eur. J. Biochem. **131**, 9 (1983).

Printed in Canada / Imprimé au Canada

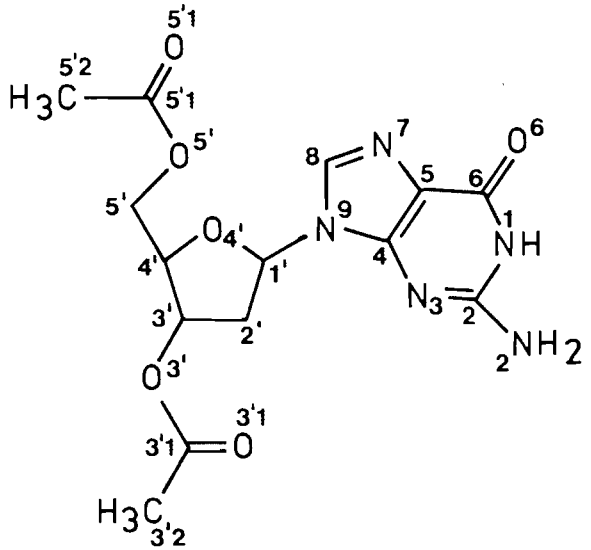

FIG. 1. Structural formula and atomic labelling of the 2′-deoxy-3′,5′-di-*O*-acetyl guanosine molecule.

4-*N*,*N*′-dimethylaminopyridine (0.15 g, 1.23 mmol) in a mixture of dry acetonitrile (100 mL) and 2.00 g (19.71 mmol) of triethylamine (14). After 2 h, methanol (2 mL) was added, and stirring was continued for 5 min. The reaction mixture was then filtered. 2′-Deoxy-3′,5′-di-*O*-acetyl guanosine was obtained as a white powder, which was purified by washing with ethanol and diethylether. Yield: 2.55 g (96%). Recrystallization from ethanol yielded plate-shaped colorless crystals. $^1$H NMR (DMSO-$d_6$): δ 1.94 (3H, s, $CH_3$), 1.97 (3H, s, $CH_3$), 2.33 (1H, m, H(2″)), 2.81 (1H, m, H(2′)), 4.11 (2H, m, H(4′/5″), 4.15 (1H, m, H(5′)), 5.19 (1H, m, H(3′)), 6.03 (1H, dd, H(1′)), 6.40 (2H, bs, $NH_2$), 7.80 (1H, s, H(8)).

*Crystallographic measurements and structure resolution*

$C_{14}H_{17}N_5O_6$ fw = 351.32

Triclinic, $a$ = 8.643(1) Å, $b$ = 10.122(1) Å, $c$ = 10.391(1) Å, α = 87.04(1)°, β = 73.58(1)°, γ = 72.37(1)°, $V$ = 830.4(2) Å$^3$, $Z$ = 2, $D_c$ = 1.405 g cm$^{-3}$, λ(MoKα) = 0.71073 Å, $t$ = 20°C, μ(MoKα) = 1.21 cm$^{-1}$, $F$(000) = 368; space group $P\bar{1}$ ($C_i^1$, No. 1).

A plate-shaped colorless crystal with dimensions (mm) 1.3 × 0.4 × 0.3 was mounted on a Nonius CAD4 diffractometer. Unit-cell dimensions were determined by least-squares refinement of the setting angles of 25 reflections in the θ range 17.1–22.0°. Intensities of 3795 reflections were measured within the sphere of reflection limited by 2θ = 55° and $-11 \leq h \leq 11$, $-12 \leq k \leq 12$, and $0 \leq l \leq 13$, using Zr-filtered MoKα radiation. The ω/2θ scan technique was applied with ω = (0.60 + 0.35 tan θ)°. The intensities of three reference reflections (014, 104, and 140), measured every 60 reflections, showed an average deviation less than 1% during the duration of the data collection. The intensities were corrected for Lorentz polarization effects, but not for absorption. A total of 3450 reflections that satisfied the criterion $I \geq 2.5\sigma(I)$ were retained for the structure determination and refinement. σ($I$) was derived from counting statistics. The structure was solved with the direct methods program of SHELXS84 (15), the coordinates of all non-hydrogen atoms being determined from an $E$-map. The hydrogen atoms were located on subsequent difference maps. The overall thermal parameter of the non-methyl hydrogen atoms refined to 0.054(4) Å$^2$. The hydrogen atoms of the methyl groups were placed at idealized positions (C—H 1.00 Å, H—C—H 109.5°). In the refinement the hydrogen atoms of the methyl groups were riding on their bonded atoms. The overall thermal parameter of the hydrogen atoms of the methyl groups refined to 0.14(1) Å$^2$. Anisotropic blocked full-matrix refinement (8 blocks) on $F$ of 552 parameters converged at $R$ = 0.031 and $R_w$ = 0.034 with $w = 2.6552/\sigma^2(F_o) + 0.00012F_o^2$. The quantity minimized was $\Sigma w\Delta F^2$, the goodness of fit $s = \Sigma w\Delta F^2/(m - n)^{1/2}$, was 2.21 and the average and maximal shift to error ratio's 0.01 and 0.08, respectively. A final $\Delta F$ showed maximal fluctuations of +0.14 and −0.14 e Å$^{-3}$. Scattering factors were those of Cromer and Mann (16), and for hydrogen of Stewart, Davidson, and Simpson (17). Programs used include SHELX76 (18) (refinement) and EUCLID (19) (geometry and illustrations).

*Nuclear magnetic resonance spectroscopy*

One-dimensional spectra were recorded in the FT mode at 300.1 MHz on a Bruker CXP 300 NMR spectrometer, interfaced with an ASPECT 2000 computer (NMR facility at the Eindhoven University of Technology). An internal $(CD_3)_2SO$ field-frequency lock, and a spectral window of 3000 Hz (16 K data points) were used. Prior to Fourier transformation, the spectra were zero-filled to 32 K, and resolution enhanced by an appropriate Gaussian multiplication. A computer simulation-iteration routine was used in order to accurately extract chemical shifts and $J$-coupling constants. This proved to be of great value, especially for the non-first order pattern that corresponds to the nearly isochronous protons H(4′), H(5′), and H(5″) (*vide infra*).

## Results and discussion

*Crystal structure*[4]

The fractional coordinates, along with equivalent isotropic temperature factors are summarized in Table 1. Bond distances between non-hydrogen atoms are given in Table 2. Table 3 lists bond angles between non-hydrogen atoms for the molecules **I** and **II**. Table 4 contains a selected set of torsion angles, describing the geometry of the furanose ring, the acetylated backbone, as well as the conformation around the glycosidic (C(1′)—N(9)) bond.

*Furanose and backbone conformation*

The 2′-deoxyribofuranose rings in the molecules **I** and **II** are found in the C(2′)-*endo* puckered form. A more precise description of the sugar ring conformations can be given with the pseudorotation concept (20), which relates the five endocyclic torsion angles mathematically to a phase angle of pseudorotation ($P$), which actually indicates which part of the ring is bent, and a puckering amplitude ($\nu_{max}$), which identifies the deviation from planarity of the ring. For the title compound, $P$ values of 167° and 164° are found for **I** and **II**, respectively. The puckering amplitudes amount to 30.6° and 35.6° for **I** and **II**, respectively. The bond distances and angles within the furanose unit are in agreement with other known C(2′)-*endo* puckered nucleosides (21, 23). The ring C—O bond distances are unequal, in such a way that C(4′)—O(4′) is approximately 0.03 Å longer than C(1′)—O(4′) (molecule **I**: C(1′)—O(4′) = 1.408(3) Å, C(4′)—O(4′) = 1.432(3) Å; molecule **II**: C(1′)—O(4′) = 1.415(2) Å, C(4′)—O(4′) = 1.442(2) Å). The difference in observed bond length is due to the anomeric effect, which is known to cause a slight shortening of C(1′)—O(4′) in nucleosides (22). The anomeric effect essentially stems from the proclivity of the heterocyclic base ring, as transmitted through N(9). The conformation around the exocyclic C(4′)—C(5′) bond shows a marked difference for molecules **I** and **II**. The *trans* (*t*) conformation (*trans* location of O(5′) and C(3′)) with γ(O(5′)—C(5′)—C(4′)—C(3′)) = 165.7(2)° is found for **I**. The C(4′)—C(5′) conformation is *gauche*$^+$ ($g^+$) (O(5′) *gauche* with respect to O(4′) and C(3′)), for **II**, γ = 57.8(2)°. The acetyl geometries are as previously noted (1–9), showing short C=O bonds. For molecule **I**, it is found that

[4]Tables of structure factor amplitudes, anisotropic temperature factors of non-hydrogen atoms, hydrogen atom coordinates, and the internal geometry may be purchased from the Depository of Unpublished Data, CISTI, National Research Council of Canada, Ottawa, Ont., Canada K1A 0S2.

TABLE 1. Refined fractional coordinates of non-hydrogen atoms, and equivalent isotropic temperature factors $U_{eq}$ (Å$^2$)[a]; estimated standard deviations are given in parentheses

| | $x$ | | $y$ | | $z$ | | $U_{eq}$ | |
|---|---|---|---|---|---|---|---|---|
| Atom | Molecule **I** | Molecule **II** | Molecule **I** | Molecule **II** | Molecule **I** | Molecule **II** | Molecule **I** | Molecule **II** |
| C(1′) | −0.0117(2) | −0.0123(2) | 0.2796(2) | 1.2572(2) | 0.1771(2) | −0.3847(2) | 0.0397(5) | 0.0406(5) |
| C(2′) | −0.0332(2) | 0.0570(3) | 0.1892(2) | 1.3408(2) | 0.0758(2) | −0.3109(2) | 0.0431(6) | 0.0461(6) |
| C(3′) | 0.1291(2) | 0.0473(3) | 0.0705(2) | 1.4705(2) | 0.0452(2) | −0.3908(2) | 0.0412(5) | 0.0421(5) |
| C(4′) | 0.2583(2) | 0.0806(2) | 0.1359(2) | 1.4166(2) | 0.0652(2) | −0.5335(2) | 0.0392(5) | 0.0403(5) |
| C(5′) | 0.3578(3) | 0.2590(3) | 0.1761(3) | 1.3930(2) | −0.0665(2) | −0.6212(2) | 0.0509(6) | 0.0475(6) |
| C(3′1) | 0.2241(4) | −0.1546(3) | −0.1606(3) | 1.6931(2) | 0.1082(3) | −0.3809(3) | 0.0700(10) | 0.0624(8) |
| C(3′2) | 0.1880(6) | −0.3294(4) | −0.2618(4) | 1.7809(3) | 0.2138(5) | −0.3063(4) | 0.1050(20) | 0.0850(10) |
| C(5′1) | 0.5085(3) | 0.5247(3) | 0.3380(2) | 1.2257(2) | −0.1361(3) | −0.6383(2) | 0.0625(6) | 0.0562(6) |
| C(5′2) | 0.5734(4) | 0.6310(4) | 0.4391(3) | 1.1301(3) | −0.0881(4) | −0.5634(4) | 0.0910(10) | 0.0800(10) |
| C(2) | −0.4383(2) | −0.1644(2) | 0.5694(2) | 0.9757(2) | 0.4505(2) | −0.0832(2) | 0.0368(5) | 0.0360(5) |
| C(4) | −0.2302(2) | −0.0011(2) | 0.5095(2) | 1.0294(2) | 0.2626(2) | −0.2716(2) | 0.0381(5) | 0.0363(5) |
| C(5) | −0.2622(2) | 0.1125(2) | 0.6401(2) | 0.8982(2) | 0.2107(2) | −0.2983(2) | 0.0438(5) | 0.0428(5) |
| C(6) | −0.3996(3) | 0.0848(3) | 0.7495(2) | 0.7915(2) | 0.2872(2) | −0.2091(2) | 0.0459(6) | 0.0423(5) |
| C(8) | −0.0487(3) | 0.2013(3) | 0.5126(2) | 1.0206(2) | 0.0662(2) | −0.4582(2) | 0.0512(6) | 0.0530(6) |
| O(3′) | 0.1063(2) | −0.1232(2) | −0.0345(2) | 1.5618(1) | 0.1411(2) | −0.3387(2) | 0.0555(5) | 0.0524(4) |
| O(4′) | 0.1635(2) | 0.0393(2) | 0.2581(2) | 1.2879(1) | 0.1488(2) | −0.5217(1) | 0.0560(5) | 0.0428(4) |
| O(5′) | 0.4429(2) | 0.3729(2) | 0.2666(2) | 1.2956(2) | −0.0347(2) | −0.5571(1) | 0.0562(5) | 0.0500(4) |
| O(3′1) | 0.3399(3) | −0.0530(3) | −0.1823(2) | 1.7301(2) | 0.0092(2) | −0.4654(3) | 0.0970(10) | 0.0889(8) |
| O(5′1) | 0.5115(3) | 0.5655(2) | 0.3178(2) | 1.2417(2) | −0.2503(2) | −0.7570(2) | 0.0845(6) | 0.0808(6) |
| O(6) | −0.4487(2) | 0.1687(2) | 0.8717(1) | 0.6685(1) | 0.2597(2) | −0.2161(2) | 0.0654(5) | 0.0549(5) |
| N(1) | −0.4823(2) | −0.0613(2) | 0.7020(2) | 0.8416(2) | 0.4075(2) | −0.1023(2) | 0.0438(5) | 0.0406(4) |
| N(2) | −0.5308(2) | −0.2966(2) | 0.5451(2) | 1.0038(2) | 0.5704(2) | 0.0266(2) | 0.0497(5) | 0.0452(5) |
| N(3) | −0.3113(2) | −0.1396(2) | 0.4672(2) | 1.0760(2) | 0.3803(1) | −0.1665(2) | 0.0384(4) | 0.0374(4) |
| N(7) | −0.1460(2) | 0.2387(2) | 0.6402(2) | 0.8945(2) | 0.0873(2) | −0.4170(2) | 0.0538(5) | 0.0532(5) |
| N(9) | −0.0930(2) | 0.0571(2) | 0.4277(2) | 1.1077(2) | 0.1682(2) | −0.3755(2) | 0.0419(5) | 0.0443(5) |

[a] $U_{eq}$ = (1/3)(trace of orthogonalized $U_{ij}$ tensor).

TABLE 2. Bond distances (Å)

| | Length | | | Length | |
|---|---|---|---|---|---|
| Bonds | Molecule **I** | Molecule **II** | Bonds | Molecule **I** | Molecule **II** |
| O(6)—C(6) | 1.227(2) | 1.230(2) | N(7)—C(5) | 1.388(3) | 1.390(3) |
| O(3′)—C(3′) | 1.438(3) | 1.447(3) | N(7)—C(8) | 1.300(3) | 1.300(3) |
| O(3′)—C(3′1) | 1.358(4) | 1.349(2) | N(9)—C(4) | 1.373(3) | 1.376(3) |
| O(4′)—C(1′) | 1.408(3) | 1.415(2) | N(9)—C(8) | 1.373(3) | 1.370(3) |
| O(4′)—C(4′) | 1.432(3) | 1.442(2) | N(9)—C(1′) | 1.461(3) | 1.458(3) |
| O(5′)—C(5′) | 1.437(3) | 1.449(3) | C(4)—C(5) | 1.380(3) | 1.378(3) |
| O(5′)—C(5′1) | 1.336(3) | 1.344(3) | C(5)—C(6) | 1.424(3) | 1.418(3) |
| O(3′1)—C(3′1) | 1.190(4) | 1.188(4) | C(1′)—C(2′) | 1.514(3) | 1.515(3) |
| O(5′1)—C(5′1) | 1.206(4) | 1.200(3) | C(2′)—C(3′) | 1.510(3) | 1.510(3) |
| N(1)—C(6) | 1.395(3) | 1.400(3) | C(3′)—C(4′) | 1.524(3) | 1.525(3) |
| N(1)—C(2) | 1.368(3) | 1.368(3) | C(4′)—C(5′) | 1.506(3) | 1.507(3) |
| N(2)—C(2) | 1.333(3) | 1.337(3) | C(3′1)—C(3′2) | 1.495(5) | 1.506(4) |
| N(3)—C(2) | 1.321(3) | 1.327(3) | C(5′1)—C(5′2) | 1.473(4) | 1.475(4) |
| N(3)—C(4) | 1.347(2) | 1.349(3) | | | |

C(3′1)—O(3′1) = 1.190(4) Å, C(5′1)—O(5′1) = 1.206(4) Å, and for **II**, C(3′1)—O(3′1) = 1.188(4) Å, C(5′1)—O(5′1) = 1.200(4) Å.

*Geometry and orientation of the base*

The glycosyl bond length C(1′)—N(9) and bond angle O(4′)—C(1′)—N(9) are 1.461(3) Å, 107.9(2)° for **I**, and 1.458(3) Å, 107.0(2)° for **II**. The guanine rings are approximately planar, the r.m.s. deviation of the 9 atoms from the least-squares plane is 0.010(4) Å, for **I** and **II**. The orientation of the base falls in the *anti* range, with values of χ(O(4′)—C(1′)—N(9)—C(4)) of −130.3(2)° for **I**, and −159.4(2)° for **II**. These data are indicated as spokes *1* and *2* in the conformational wheel of Fig. 2, showing the distribution of the glycosidic torsion angle χ in the crystal structures of acetylated nucleosides. Figure 2 clearly indicates that the χ values for acetylated nucleosides have a bimodal distribution. Nine *anti* conformations occur in the range χ = −148 ± 33°, and three *syn* conformations have been found in the range χ = 63 ± 11°. The favored *anti* and *syn* domains of acetylated nucleosides

TABLE 3. Bond angles (deg)

| Bonds | Angle Molecule I | Angle Molecule II | Bonds | Angle Molecule I | Angle Molecule II |
|---|---|---|---|---|---|
| C(3′)—O(3′)—C(3′1) | 114.9(2) | 116.7(2) | C(2′)—C(1′)—O(4′) | 106.1(2) | 105.8(2) |
| C(1′)—O(4′)—C(4′) | 111.3(2) | 109.8(1) | C(2′)—C(1′)—N(9) | 113.9(2) | 113.4(2) |
| C(5′)—O(5′)—C(5′1) | 116.1(2) | 116.0(1) | O(4′)—C(1′)—N(9) | 107.9(2) | 107.0(2) |
| C(2)—N(1)—C(6) | 125.7(2) | 125.4(2) | C(3′)—C(2′)—C(1′) | 103.5(2) | 102.6(2) |
| C(4)—N(3)—C(2) | 112.0(2) | 111.8(2) | O(3′)—C(3′)—C(2′) | 108.0(2) | 106.2(2) |
| C(5)—N(7)—C(8) | 104.5(2) | 104.5(2) | O(3′)—C(3′)—C(4′) | 111.1(2) | 112.1(2) |
| C(1′)—N(9)—C(4) | 125.9(2) | 125.6(2) | C(2′)—C(3′)—C(4′) | 103.5(2) | 102.8(2) |
| C(1′)—N(9)—C(8) | 127.6(2) | 127.6(2) | O(4′)—C(4′)—C(3′) | 106.2(2) | 106.6(2) |
| C(4)—N(9)—C(8) | 106.5(2) | 106.4(2) | O(4′)—C(4′)—C(5′) | 109.2(2) | 109.6(2) |
| N(3)—C(2)—N(1) | 123.7(2) | 123.8(2) | C(3′)—C(4′)—C(5′) | 111.2(2) | 115.1(2) |
| N(9)—C(4)—N(3) | 125.8(2) | 125.7(2) | C(4′)—C(5′)—O(5′) | 106.7(2) | 108.4(2) |
| N(9)—C(4)—C(5) | 105.2(2) | 105.4(2) | O(3′)—C(3′1)—O(3′1) | 122.8(3) | 122.3(2) |
| N(3)—C(4)—C(5) | 129.0(2) | 128.9(2) | O(3′)—C(3′1)—C(3′2) | 110.5(3) | 111.1(2) |
| C(4)—C(5)—C(6) | 118.5(2) | 118.9(2) | O(3′1)—C(3′1)—C(3′2) | 126.7(3) | 126.6(2) |
| C(4)—C(5)—N(7) | 110.7(2) | 110.5(2) | O(5′1)—C(5′1)—O(5′1) | 122.1(2) | 122.8(2) |
| C(6)—C(5)—N(7) | 130.8(2) | 130.6(2) | O(5′)—C(5′1)—C(5′2) | 110.0(2) | 111.8(2) |
| N(1)—C(6)—C(5) | 111.1(2) | 111.1(2) | O(5′1)—C(5′1)—C(5′2) | 126.9(3) | 125.4(3) |
| N(1)—C(6)—O(6) | 120.2(2) | 119.9(2) | N(1)—C(2)—N(2) | 116.7(2) | 116.5(2) |
| C(5)—C(6)—O(6) | 128.8(2) | 128.9(2) | N(3)—C(2)—N(2) | 119.6(2) | 119.8(2) |
| N(7)—C(8)—N(9) | 113.1(2) | 113.2(2) | | | |

TABLE 4. Selected torsion angles (deg)

| Bonds | Torsion angle Molecule I | Torsion angle Molecule II |
|---|---|---|
| C(4′)—O(4′)—C(1′)—C(2′) | −16.2(2) | −20.3(2) |
| O(4′)—C(1′)—C(2′)—C(3′) | 28.7(2) | 33.7(2) |
| C(1′)—C(2′)—C(3′)—C(4′) | −29.8(2) | −33.4(2) |
| C(2′)—C(3′)—C(4′)—O(4′) | 21.0(2) | 22.3(2) |
| C(2′)—C(3′)—C(4′)—C(5′) | −97.8(2) | −99.4(2) |
| C(1′)—O(4′)—C(4′)—C(3′) | −3.1(2) | −1.4(2) |
| C(3′1)—O(3′)—C(3′)—C(2′) | 164.8(2) | 170.2(2) |
| C(3′)—O(3′)—C(3′1)—C(3′2) | −179.6(3) | −172.9(2) |
| C(3′)—O(3′)—C(3′1)—O(3′1) | 1.7(4) | 6.4(4) |
| C(3′)—C(4′)—C(5′)—O(5′) | 165.7(2) | 57.8(2) |
| C(5′1)—O(5′)—C(5′)—C(4′) | −166.0(2) | 157.3(2) |
| C(5′)—O(5′)—C(5′1)—C(5′2) | 173.0(2) | 179.7(2) |
| C(5′)—O(5′)—C(5′1)—O(5′1) | −7.2(4) | 0.0(3) |
| C(8)—N(9)—C(1′)—O(4′) | 49.4(3) | 28.5(3) |
| C(1′)—N(9)—C(8)—N(7) | 179.9(2) | 173.9(2) |
| C(5)—N(7)—C(8)—N(9) | −0.1(3) | −0.8(3) |
| C(8)—N(7)—C(5)—C(4) | 0.6(2) | 0.7(2) |
| N(9)—C(4)—C(5)—N(7) | −0.8(2) | −0.3(2) |
| C(8)—N(9)—C(4)—C(5) | 0.7(2) | −0.1(2) |
| C(2)—N(3)—C(4)—C(5) | 0.5(3) | −1.7(3) |
| C(4)—N(3)—C(2)—N(1) | 0.5(3) | 0.1(3) |
| C(6)—N(1)—C(2)—N(3) | −1.5(3) | 1.3(3) |
| C(2)—N(1)—C(6)—C(5) | 1.3(3) | −0.9(3) |
| C(4)—C(5)—C(6)—N(1) | −0.3(3) | −0.6(3) |
| C(2)—N(1)—C(6)—O(6) | −178.9(2) | 179.7(2) |
| C(4)—N(3)—C(2)—N(2) | −179.4(2) | −179.9(2) |
| C(6)—N(1)—C(2)—N(2) | 178.4(2) | −178.8(2) |
| C(4)—C(5)—C(6)—O(6) | 179.9(2) | 178.7(2) |

are in close agreement with the structural data set of Altona *et al.* (21).

*Hydrogen bonding and molecular stacking*

The N—H and $NH_2$ groups of the purine fragment of molecules **I** and **II** are hydrogen-bond donors in seven intermolecular hydrogen bonds (Table 5). Two hydrogen bonds between molecules **I** and **II** form a dimer-type ribbon typical for guanosine derivatives. These bonds, N(1)...N(7) and N(2)...O(6), connect molecules **II** and **I**, and the same interaction exists between molecule **I** and a translationally related molecule **II** at $(-1 + x, y, 1 + z)$. This repeated pattern gives rise to an infinite chain of paired bases running in the $\bar{1}\,0\,1$ direction. Each molecule has an additional hydrogen bond between the $NH_2$ group and O(5′1) of the acetyl group at C(5′) of a translationally related molecule, also at $(-1 + x, y, 1 + z)$. The seventh hydrogen bond is between the $NH_2$ group of **II** and the ester oxygen $O_{5'}$ of the acetyl group at $C_{5'}$ of **I** at $(-1 + x, 1 + y, z)$. This implies that one N—H bond of the $NH_2$ group of **II** is involved in a bifurcate interaction, which is planar as follows from the sum of angles (359.5°), having the H atom as apex. The bifurcation results in weak hydrogen bonds as may be seen from their geometries (Table 5). Like in other acetylated pyrimidine and purine nucleosides (1–9), the acetyl groups play an important role in the stacking, in that the polar carbonyl group is oriented toward the base ring of neighbouring molecules. As Table 4 shows, the two carbonyl oxygen atoms O(3′1) and O(5′1) of both **I** and **II** have close contacts with atoms of the guanosine ring. The greater part of these contacts and also the shortest ones involve the carbonyl groups of molecule **II**.

*Conformation in solution*

The set of experimental *J*-coupling constants, determined at 298 K, are summarized in Table 6. The data refer to the solvent DMSO-$d_6$, since 2′-deoxy-3′,5′-di-*O*-acetyl guanosine was found to be virtually insoluble in $D_2O$. An iterative computer simulation procedure was used in order to obtain accurate NMR spectral data (Fig. 4). The structure of the 2′-deoxyribose ring can be described in terms of a rapid C(2′)-*endo* ⇄ C(3′)-*endo* conformational equilibrium (25, 26):

$$x(\text{C(2′)-}endo) = (17.8 - J(1'2'') - J(2''3''))/10.9 = 0.88$$

A similar preference for the C(2′)-*endo* conformation is

TABLE 5. Intermolecular hydrogen bond geometry and structural details of acetyl-base-stacking[a]

| Donor | Acceptor | *D*....*A* (Å) | H....*A* (Å) | *D*—H....*A* (deg) |
|---|---|---|---|---|
| N(1,I)—H(1) | N(7,IIA) | 2.823(3) | 2.06(3) | 178(3) |
| N(2,I)—H(2) | O(6,IIA) | 2.879(3) | 2.06(3) | 147(3) |
| N(2,I)—H(3) | O(5'1,IA) | 2.886(3) | 2.05(3) | 155(3) |
| N(1,II)—H(1) | N(7,I) | 2.872(3) | 1.93(3) | 176(3) |
| N(2,II)—H(2) | O(6,I) | 2.883(3) | 2.09(3) | 156(3) |
| N(2,II)—H(3) | O(5'1,IIA) | 3.095(3) | 2.42(2) | 139(2) |
| N(2,II)—H(3) | O(5',IB) | 3.092(3) | 2.45(3) | 134(2) |
| Acetyl-base stacking (Å) | | | | |
| O(3'1,I) — | C(6,IIC) | 3.650(4) | O(3'1,II) | —C(6,IIE) 3.363(4) |
| O(5'1,I) — | N(3,IID) | 3.555(3) | O(3'1,II) | —C(5,IIE) 3.338(3) |
| O(3'1,II) — | O(6,IIE) | 3.562(3) | O(5'1,II) | —N(3,IF) 3.324(3) |
| O(3'1,II) — | N(7,IIE) | 3.567(3) | O(5'1,II) | —C(4,IF) 3.697(3) |

[a]I and II refer to molecules **I** and **II**, respectively. Symmetry code: (A) $-1 + x, y, 1 + z$; (B) $-1 + x, 1 + y, z$; (C) $x, -1 + y, z$; (D) $1 + x, -1 + y, z$; (E) $x, 1 + y, z$; (F) $1 + x, 1 + y, -1 + z$.

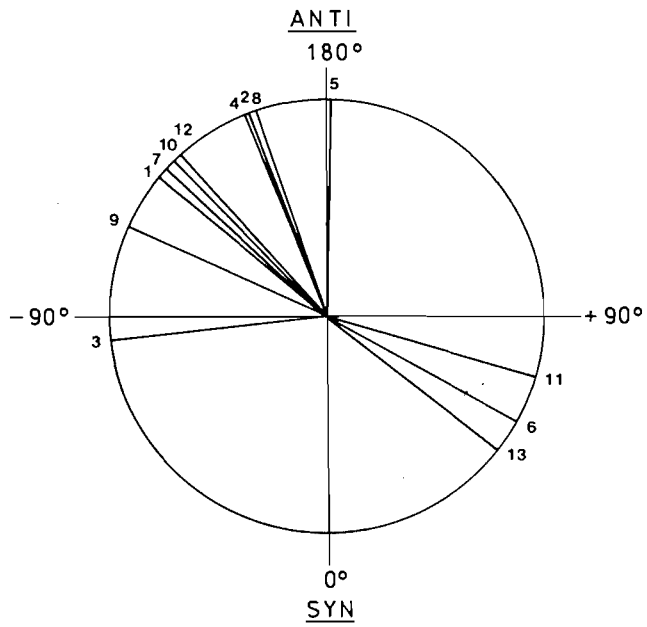


FIG. 2. Conformational wheel representing the glycosidic torsion angles for backbone-acetylated nucleosides (χ(O(4')—C(1')—N(9)—C(4)) for purine, and χ(O(4')—C(1')—N(9)—C(2) for pyrimidine structures). *1*, *2*: χ = −130° and −159°, 2'-deoxy-3',5'-di-*O*-acetyl guanosine (this work); *3*: χ = −84°, 2'-deoxy-3',5'-di-*O*-acetyl adenosine (ref. 1); *4*, *5*: χ = −158° and 179°, 2',3',5'-tri-*O*-acetyl guanosine (orthorhombic form, ref. 2); *6*: χ = 61°, 2',3',5'-tri-*O*-acetyl adenosine (ref. 3); *7*: χ = −133°, 1/1 complex of 2'-deoxy-3',5'-di-*O*-acetyl-5-bromo uridine and 2'-deoxy-3',5'-di-*O*-acetyl-5-iodo uridine (ref. 4); *8*: χ = −161°, 2',3',5'-tri-*O*-acetyl guanosine (ref. 5); *9*, *10*: χ = −115° and −136°, 3',5'-di-*O*-acetyl thymidine (refs. 6, 11); *11*: χ = 74°, 2',3',5'-tri-*O*-acetyl uridine (ref. 7); *12*: χ = −138°, 2'-deoxy-3',5'-di-*O*-acetyl-5-iodo uridine (ref. 8); *13*: χ = 52°, 2'-deoxy-3'-*O*-acetyl adenosine (ref. 9).

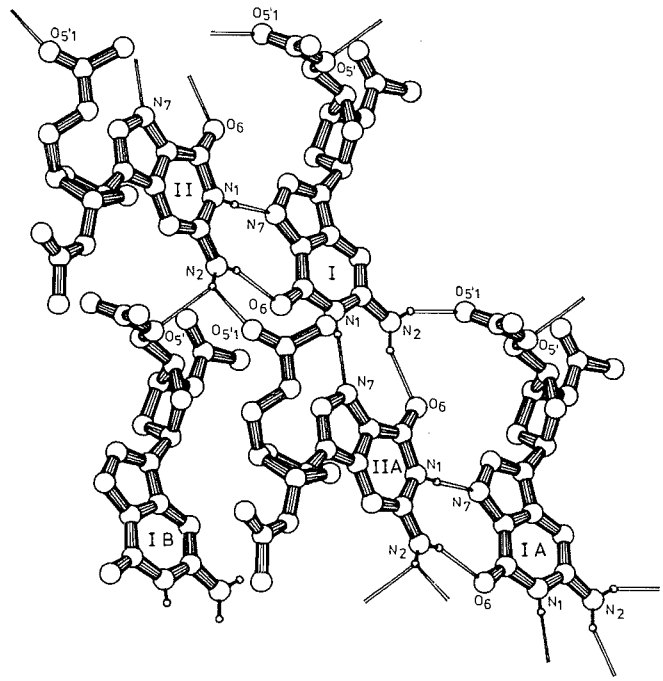


FIG. 3. Projection of a part of the structure showing the hydrogen bond network as seen approximately perpendicular the (1 2 0) plane. Only N and O atoms involved in hydrogen bonds have been labelled. A and B denote molecules **I** and **II** at positions $-1 + x, y, 1 + z$ and $-1 + x, 1 + y, z$, respectively. Hydrogen atoms not involved in hydrogen bonds have been omitted for clarity.

commonly encountered for 2'-deoxyribo nucleotides, including 3',5'-di-*O*-acetyl thymidine and 2'-deoxy-3',5'-di-*O*-acetyl adenosine (1, 12, 24).

The conformation around the exocyclic C(4')—C(5') bond was characterized on the basis of $J(4'5')$ and $J(4'5'')$ (1, 25, 26). From the experimental data one calculates: $x(g^+) = 0.31$, $x(t) = 0.40$, $x(gauche^-) = 0.29$, i.e., only a slight preference

TABLE 6. Vicinal coupling constants (in Hz) between the protons of the backbone and the 2'-deoxyribofuranose ring, measured in DMSO-$d_6$ at 20°C

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $J(1'2')$ | 8.4 | $J(3'4')$ | 2.0 |
| $J(1'2'')$ | 5.8 | $J(4'5')$ | 4.9 |
| $J(2'2'')$ | −14.0 | $J(4'5'')$ | 5.7 |
| $J(2'3')$ | 6.4 | $J(5'5'')$ | −11.9 |
| $J(2''3')$ | 2.4 | | |

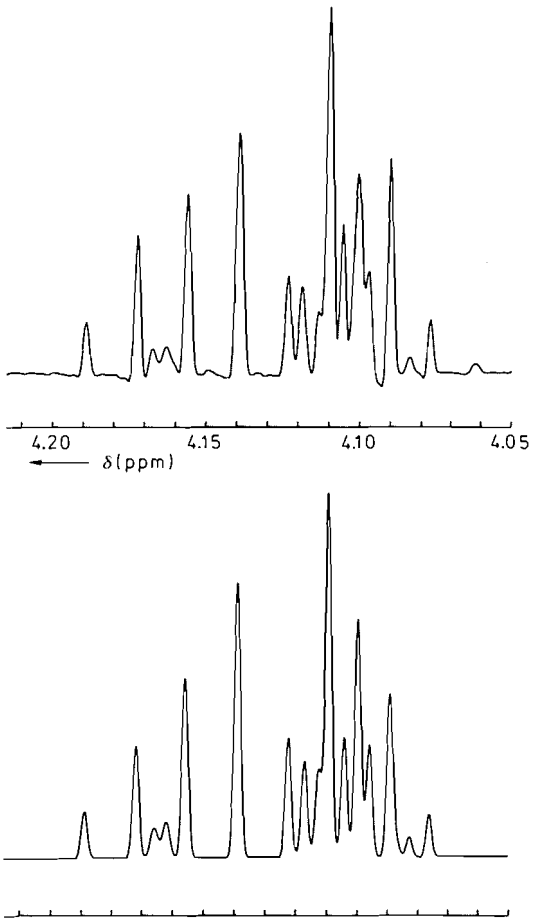


FIG. 4. Upper trace: expansion of the H(4′)/H(5′)/H(5″) pattern (non-first order) in the 300 MHz $^1$H NMR spectrum. Lower trace: computer simulation.

for *t* over $g^+$ and *gauche*$^-$ ($g^-$) is observed. It should be noted that 3′,5′-di-*O*-acetyl thymidine and 2′-deoxy-3′,5′-di-*O*-acetyl adenosine exhibit a clear preference for $g^+$ in solution.

## Conclusion

The crystal structure of 2′-deoxy-3′,5′-di-*O*-acetyl guanosine consists of two independent molecules **I** and **II** in the asymmetric unit. The guanine bases are connected via two N(1)—H...N(7), and two N(2)—H...O(6) hydrogen bonds. The hydrogen donating base in this structure in fact mimics the C base in the Hoogsteen C—G base pair (27). Two hydrogen bonds N(2)—H...O(5′1) connect pairs of molecule **I** and pairs of molecule **II**. An additional N(2)—H...O(5′) interaction between **I** and **II** gives rise to a bifurcate hydrogen bond. Highly comparable hydrogen bonding patterns (except for the bifurcated hydrogen bond) have been found for the acetylated RNA counterpart structure 2′,3′,5′-tri-*O*-acetyl guanosine, both in the monoclinic (5) and orthorhombic (2) form. The conformational properties of the furanose unit and the acetylated backbone show a close correspondence between the solid state and the DMSO solution. The molecules **I** and **II** both have C(2′)-*endo* type furanose rings in the solid state, while a clear preference for C(2′)-*endo* is also observed in solution. With respect to the C(4′)—C(5′) conformation, it is found that an approximately equal distribution over the rotamers $g^+$, *t*, and $g^-$ exists in solution. Both *t* (molecule **I**) and $g^+$ (molecule **II**) are encountered in the crystal structure. The crystal structure of 2′-deoxy-3′,5′-di-*O*-acetyl guanosine differs substantially from the structures of other acetylated DNA nucleosides, i.e., 2′-deoxy-3′,5′-di-*O*-acetyl adenosine (in which the A bases are linked in one-dimensional chains via N(6)—H...N(1) and N(6)—H...N(7) hydrogen bonds), and 3′,5′-di-*O*-acetyl thymidine, for which a parallel T–T base pair was found. Work is in progress to solve the X-ray structure of the fourth acetylated DNA nucleoside, 2′-deoxy-3′,5′-di-*O*-acetyl cytidine.


1. L. H. KOOLE, H. M. BUCK, J. A. KANTERS, and A. SCHOUTEN. Can. J. Chem. **65**, 326 (1987).
2. J. N. LOW, P. TOLLIN, C. C. WILSON, and S. H. SCRIMGEOUR. Acta Crystallogr. Sect. C, **42**, 700 (1986).
3. C. C. WILSON, P. TOLLIN, and R. A. HOWIE. Acta Crystallogr. Sect. C, **42**, 697 (1986).
4. C. C. WILSON and P. TOLLIN. Acta Crystallogr. Sect. C, **41**, 465 (1985).
5. C. C. WILSON, J. N. LOW, and P. TOLLIN. Acta Crystallogr. Sect. C, **41**, 1123 (1985).
6. C. C. WILSON, J. N. LOW, P. TOLLIN, and H. R. WILSON. Acta Crystallogr. Sect. C, **40**, 1712 (1984).
7. J. N. LOW and C. C. WILSON. Acta Crystallogr. Sect. C, **40**, 1030 (1984).
8. C. C. WILSON, J. N. LOW, and D. W. YOUNG. Acta Crystallogr. Sect. C, **39**, 1103 (1983).
9. S. T. RAO and M. SUNDARALINGAM. J. Am. Chem. Soc. **92**, 4963 (1970).
10. J. D. WATSON and F. H. C. CRICK. Nature (London), **171**, 173 (1953).
11. L. H. KOOLE, M. H. P. VAN GENDEREN, H. FRANKENA, H. J. M. KOCKEN, J. A. KANTERS, and H. M. BUCK. Proc. Kon. Ned. Akad. Wet. Ser. B, **89**, 51 (1986); communicated by H. M. Buck at the meeting of November 25, 1985.
12. L. H. KOOLE, M. H. P. VAN GENDEREN, and H. M. BUCK. J. Am. Chem. Soc. **109**, 3916 (1987).
13. M. H. P. VAN GENDEREN, L. H. KOOLE, and H. M. BUCK. Proc. Kon. Ned. Akad. Wet. Ser. B, 181 (1987); communicated by H. M. Buck at the meeting of June 22, 1987.
14. A. MATSUDA, M. SHINOZAKI, M. SUZUKI, K. WATANABE, and T. MIYASAKA. Synthesis, 385 (1986).
15. G. M. SHELDRICK. SHELXS 84. Program for crystal structure solution. Univ. of Göttingen, Federal Republic of Germany.
16. D. T. CROMER and J. B. MANN. Acta Crystallogr. Sect. A, **24**, 321 (1968).
17. R. F. STEWART, E. R. DAVIDSON, and W. T. SIMPSON. J. Chem. Phys. **42**, 3175 (1965).
18. G. M. SHELDRICK. SHELXS 76. Program for crystal structure determination. Univ. of Cambridge, England.
19. A. L. SPEK. EUCLID. Computational crystallography. *Edited by* D. Sayre. Clarendon Press, Oxford, England. 1982. p. 528.
20. C. ALTONA and M. SUNDARALINGAM. J. Am. Chem. Soc. **94**, 8205 (1972).
21. H. P. M. DE LEEUW, C. A. G. HAASNOOT, and C. ALTONA. Isr. J. Chem. **20**, 108 (1980).
22. A. J. KIRBY. The anomeric effect and related stereo-electronic effects at oxygen. Springer Verlag, Berlin. 1983.
23. U. THEWALT, C. E. BUGG, and R. E. MARSH. Acta Crystallogr. Sect. B, **26**, 1089 (1970).
24. D. B. DAVIES. *In* Progress in nuclear magnetic resonance spectroscopy. *Edited by* J. W. Emsley, J. Feeney, and L. H. Sutcliffe. Pergamon Press, Oxford, England. 1978.
25. C. A. G. HAASNOOT, F. A. A. M. DE LEEUW, H. P. M. DE LEEUW, and C. ALTONA. Recl. Trav. Chim. Pays-Bas, **98**, 576 (1979).
26. C. ALTONA. Recl. Trav. Chim. Pays-Bas, **101**, 413 (1982).
27. K. HOOGSTEEN. Acta Crystallogr. **16**, 907 (1963).