

The Crystal and Molecular Structure of a Calcium Salt of Guanylyl-3',5'-Cytidine (GpC)

BY B. HINGERTY,* E. SUBRAMANIAN,† S. D. STELLMAN,‡ T. SATO,§ S. B. BROYDE|| AND R. LANGRIDGE

Department of Biochemical Sciences, Frick Chemistry Laboratory, Princeton University,
Princeton, New Jersey 08540, U.S.A.

(Received 26 March 1976; accepted 1 May 1976)

The calcium salt, $\text{Ca}(\text{C}_{19}\text{H}_{24}\text{N}_8\text{O}_{12}\text{P})_2 \cdot 18\text{H}_2\text{O}$, of guanylyl-3',5'-cytidine (GpC) has been refined to an R of 8.2% for 2918 observed reflections (11% for 4237 reflections, including unobserved). The molecule crystallized in space group $P2_1$ with $a = 21.224$, $b = 34.207$, $c = 9.327$ Å, $\beta = 90.527^\circ$, $Z = 4$. The asymmetric unit contains four GpC, 36 waters and two Ca^{2+} ions, for a total of 198 non-hydrogen atoms. The four GpC occur as two dimers related by a pseudo C -face-centering. Each dimer consists of two crystallographically independent GpC as Watson-Crick base-pairs, and possesses a pseudo twofold axis broken by a Ca^{2+} ion and associated solvent. The structure was solved by an unusual series of steps including semi-empirical potential-energy methods, packing analysis, rigid-body refinement, least-squares and difference Fourier techniques, and direct-methods tangent-formula phase refinement. The four GpC have conformational angles in the range of helical RNA, but are not identical. The different crystallographic environments perturb the GpC from exact symmetry and demonstrate the range of the basic helical conformations. All eight bases are *anti*, sugars are all $C(3')$ *endo*, the $C(4')\text{--}C(5')$ bond rotations are *gauche-gauche*, and the ω', ω angle pair about the O-P bonds is *gauche-gauche*.

Introduction

The dinucleoside phosphates represent the smallest subunit of DNA and RNA which possess all the conformational angles characteristic of the larger polymeric nucleic acids. Structural studies on dinucleoside phosphates are thus useful in augmenting the information derived from fiber diffraction patterns of nucleic acids. Here we describe the X-ray analysis of the structure of a Ca^{2+} salt of guanylyl-3',5'-cytidine (GpC) (Fig. 1). Preliminary reports on this work have been published (Hingerty, Subramanian, Stellman, Broyde, Sato & Langridge, 1975). X-ray analysis of a Na^+ salt of GpC, crystallizing in space group $C2$, with one molecule per asymmetric unit, has been reported (Day, Seeman, Rosenberg & Rich, 1973), who also report obtaining a $P2_1$ form with $Z = 2$. In our work, the Ca^{2+} salt of GpC crystallizes in space group $P2_1$ with four crystallographically independent GpC per asymmetric unit, and is morphologically different from the other forms, except for similarities noted later.

Experimental

Crystal parameters

Crystallization: 4.9 mg of GpC (Miles Laboratories) were suspended in 2.0 ml of 10 mM tris buffer, pH 8.7,

containing 5 mM CaCl_2 . The vessel containing the suspension was sealed and placed in a water bath. The solid dissolved at 50°C with stirring. The temperature of the bath was gradually decreased at the rate of 1 or 2°C per day, down to 22°C . The bath was placed against a cool surface which acted as a temperature gradient. In this arrangement CaGpC began to crystallize at about 35°C . The crystals were unstable in air when separated from the mother liquor. To avoid this problem a crystal was slowly dried over a period of days. Such a crystal had slightly smaller cell dimensions but was then stable in the X-ray beam without mother liquor.

Intensities were collected on a Hilger & Watts Y290 diffractometer with Ni-filtered $\text{Cu } K\alpha$ radiation ($\mu = 20.5 \text{ cm}^{-1}$). The data were recorded out to 1.1 Å in the θ - 2θ step scan mode in 80 steps of 0.01° , giving a scan range of 0.8° in θ . Counts were accumulated for 3 s/step, except for the background measurements which were recorded for 30 s. The crystal was mounted about c and a monitor reflection (040) used to correct the data for radiation damage, which ranged from 10–15%. 4237 reflections were recorded; of these, 2918 exceeded background by one standard deviation, and were considered to be observed. A low cut-off (to choose the threshold) on intensity was used because the presence of a high degree of pseudosymmetry caused a large number of systematically weak reflections. The data were corrected for Lorentz, polarization, and background effects, but not for absorption.

Crystal data: $\text{Ca}(\text{C}_{19}\text{H}_{24}\text{N}_8\text{O}_{12}\text{P})_2 \cdot 18\text{H}_2\text{O}$; M.W. 1539.2; space group $P2_1$ (based upon the systematic absences $0k0$, k odd); $a = 21.224$, $b = 34.207$, $c = 9.372$ Å, $\beta = 90.527^\circ$; $V = 6803.9$ Å³; $Z = 4$; $D_c = 1.503$, D_m (density gradient) = 1.512 g cm^{-3} ; crystal size: $0.03 \times 0.06 \times 0.45$ mm.

Pseudosymmetry: A striking aspect of unit-cell sym-

* To whom correspondence should be addressed at MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, England.

† Present address: Laboratory of Molecular Biology, National Institutes of Health, Bethesda, Maryland 20014, U.S.A.

‡ Present address: American Health Foundation, 1370 Avenue of the Americas, New York, New York 10019, U.S.A.

§ Present address: Shionogi Research Laboratory, Fukushima-ku, Osaka, Japan.

|| Present address: School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, U.S.A.

metry was apparent from the intensity distribution observed in the three zonal photographs: $h0l$, $0kl$ and $hk0$. Among $h0l$ reflections, $h=2n+1$ are mostly absent or weak though there are exceptions. Among $0kl$ reflections, $k=4n$ are quite strong, while $k=4n+2$ are nearly absent (or weak) and $k=2n+1$ are intermediate in strength. A similar pattern was observed among $hk0$ reflections. A statistical analysis of the E 's indicated a centrosymmetric distribution (Table 1); the E 's also displayed a peculiar rational dependence (Table 2) (Stout & Jensen, 1968). All these observations can be reconciled by assuming that there is a pseudotranslational repetition along \mathbf{b} , equal to $\sim b/2$, and that the pseudo unit cell has a space group symmetry of $C2$. It is interesting to note that in NaGpC (Rosenberg, Seeman, Day & Rich, 1973) this pseudo unit cell itself becomes the true cell.

Table 1. Expected and observed statistical parameters on the normalized structure factor amplitudes ($P2_1$)

	Centro-symmetric	Noncentro-symmetric	GpC
Average $E^2 - 1$	0.968	0.736	1.020
Average E	0.798	0.886	0.772
$E1$, %	32.0	36.8	27.8
$E2$, %	5.0	1.8	5.1
$E3$, %	0.3	0.1	0.7

Table 2. Rational dependence of structure factor amplitudes in GpC ($P2_1$)

	Average E^2	Expected E^2
$k = \text{even}$		
$h + (k/2) = \text{even}$	2.25	1.0
$h + (k/2) = \text{odd}$	0.25	1.0
$k = \text{odd}$	0.75	1.0

Monoclinic space group	Real cell $P2_1$	Pseudo cell $C2$	Day <i>et al.</i> (1973) $C2$
a	21.224 Å	21.224 Å	21.460 Å
b	34.207	17.104	16.927
c	9.372	9.372	9.332
β	90.527(°)	90.527(°)	90.54(°)
Z	8	4	4

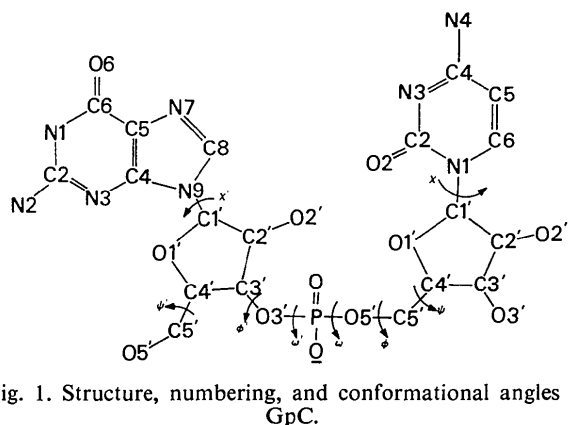


Fig. 1. Structure, numbering, and conformational angles for GpC.

Structure analysis

A Patterson synthesis was calculated with all reflections ($P2_1$ symmetry) to locate the P atoms (Sussman, 1971). From the distribution of peaks in sections at $y=0$ and $\frac{1}{2}$, the presence of pseudosymmetry in the unit cell was obvious. Though there were four P atoms in the asymmetric unit, the distribution of interatomic vectors suggested that they were very closely related to each other by $C2$ symmetry.

An attempt was therefore made to carry out a preliminary analysis of the structure in the pseudosymmetric space group $C2$. For this all reflections of the type $k=2n+1$ were ignored, and the rest were re-indexed with respect to the pseudo unit cell. Since the only difference between the pseudo and true unit cell is that the former has $b' \simeq b/2$, the re-indexing amounts to dividing all k indices by 2. For this subset of reflections $hk'l$, $h+k'=2n+1$ are generally weak and were also ignored, leaving only $h+k'=2n$, thus accounting for the $C2$ symmetry.

A Fourier calculation based on P phases yielded several peaks. Fourier and isotropic structure factor least-squares calculations with the newly appearing peaks were carried out. However, the result was either a distorted base or unrecognizable sugar group or unacceptable bond distances or angles. This analysis is obviously vitiated by the fact that the deviation from pseudosymmetry is very significant. In addition, the ratio of number of parameters to number of observations was only of the order of 1:4.

It was clear that a meaningful analysis in the pseudosymmetric space group $C2$ (and for that matter in the true space group $P2_1$) could only be carried out by constrained refinement in which the bases and sugar rings are held fixed close to their standard dimensions, and in which two GpC related by the pseudo twofold axis are paired together through hydrogen bonding in a complementary way (Watson-Crick type). Furthermore, it was also desirable to use suitable starting values for the conformational angles about various single bonds. In particular, it was necessary to use reasonable values for (a) the conformation about glycosidic bonds, (b) the conformation of the sugar ring (pucker) and (c) the conformations around the phosphate linkage. The values in the literature for these angles were largely for mononucleosides and mononucleotides. It was therefore decided to compute the conformational energy of the molecule as a function of the various dihedral angles. Then the parameters corresponding to the minimum energy together with the interactive computer graphics system were used in a packing search over orientation space. Details of these computations have been published (Stellman, Hingerty, Broyde, Subramanian, Sato & Langridge, 1973). The packing search used constraints outlined earlier: (a) the unit cell was required to have pseudo $C2$ symmetry, (b) the location of the P atom was taken to be that deduced from the Patterson synthesis, (c) the orientation

of the bases was fixed with respect to the pseudo two-fold axis, and (*d*) a dimeric association of molecules through complementary Watson-Crick base-pairing was assumed. The last constraint was incorporated by a geometric function. The interactive computer graphics system was used to visualize the calculated structures. From these calculations a trial structure was obtained which had an *R* of 39.3% referred to 3 Å data only, which included just 68 reflections of the sub-lattice *C*2. Recalculation of the structure factors for all 768 observed reflections at 1.1 Å resolution gave *R* = 49% for these pseudocell reflections.

Structural refinement

Refinement in pseudocell *C*2

In order to obtain a better trial structure, a constrained refinement procedure was followed: all bond lengths and angles were held constant, and *R* was minimized *via* the Powell (1964) algorithm, in a manner analogous to the prior energy minimization. Since the Powell algorithm uses numerical methods, the complex procedure used to take derivatives of internal parameters was avoided.

The following 14 parameters were used as variables in the linked-atom refinement: (1) the eight dihedral angles χ' , ψ' , ϕ' , ω' , ω , ϕ , ψ , χ (Table 3), (2) two coordinates of the P atom [equivalent to the center of mass $x(p)$ and $z(p)$ with $y(p)$ fixed at 0, since *b* is unique], (3) the three Eulerian angles for the orientation of the molecule as a whole, *th*, *Q'* and *Q*, (4) the pseudorotation parameter, *P*, for the sugar puckering (Altona & Sundaralingam, 1972), which was assumed identical for both sugars.

The minimization procedure described here is analogous to linked-atom least squares in that the dependent variable to be minimized is *R*. It differs

substantially, however, in the type of minimization procedure used (gradient rather than matrix inversion), and in the geometrical parameters used as independent variables. This procedure reduced *R* to 36.2% for the 768 reflections, with the results summarized in Table 4.

Least-squares treatment was then resumed with the block-diagonal least-squares routine *LSSQR* (Trotter, 1965) in which the individual atoms were allowed to vary with isotropic temperature factors. This reduced *R* to 25%. A series of difference syntheses were then calculated in which the $\frac{1}{2}\text{Ca}^{2+}$ ion and most of the water sites were found. The Ca^{2+} was originally identified as a water molecule, but its temperature factor refined to -2.3 \AA^2 . Because no other peaks as strong as this were seen and since the peak was near the phosphate group, it was assumed to be a Ca^{2+} ion. It was assigned an occupancy of one half for electrical neutrality. With the cation plus water, *R* was reduced to 14.4% for 768 observed *C*2 reflections when some disordered water sites were also included.

Refinement in real space group *P*2₁

In order to refine the structure in the correct space group *P*2₁, it was evident that some knowledge of the deviations from the subcell *C*2 would be needed. The only way to refine the structure was to calculate the phases of the unknown reflections ($\frac{3}{4}$ of the data) by an unbiased direct method.

Direct methods: On theoretical grounds, since $\frac{1}{4}$ of the phases were known from the subcell *C*2, the enantiomorph was already determined. Because so many phases were already known, it was then possible to expand the phase set by the weighted tangent formula if the origin were properly fixed. To fix the origin, a reflection of index *k* odd was necessary. Most of these reflections were systematically weak due to pseudosymmetry. To remedy this problem the normalized structure amplitudes (*E*'s) were renormalized (Hauptman, Karle & Karle, 1960) by classes of reflections and shells of $(\sin \theta/\lambda)^2$.

With such a set of *E*'s the program *MULTAN* (Germain, Main & Woolfson, 1970) was used in an expanded version that allowed 1000 reflections and 10 000 relationships for $E > 1.0$. All the phases of *C*2 were used that calculated at least 60% of their observed values. These were fixed during the first cycles of refinement. The reflection 412 was arbitrarily given a phase of 45° to define the origin. The absolute figure of merit was 1.25, a good value indicating a reasonable result.

Table 3. Definition of dihedral angles for *GpC*

Angle	Bonds
χ'	O(1')-C(1')-N(9)-C(8)
ψ'	C(3')-C(4')-C(5')-C(8)
ϕ'	P-O(3')-C(3')-C(4')
ω'	O(5')-P-O(3')-C(3')
ω	C(5')-O(5')-P-O(3')
ϕ	C(4')-C(5')-O(5')-P
ψ	C(3')-C(4')-C(5')-O(5')
χ	C(6)-N(1)-C(1')-O(1')

All angles *A-B-C-D* are measured clockwise from *A* to *D* when viewed along *B-C* with *A* eclipsing *D* as 0°.

Table 4. Linked-atom refinement results

X and *Z* are of the P atom, *P* is the pseudorotation parameter, and *th*, *Q'* and *Q* are Euler angles.

Parameters (initial)	Reflections	χ'	ψ'	ϕ'	ω'	ω	ϕ	ψ	χ	<i>X</i>	<i>Z</i>	<i>th</i>	<i>Q'</i>	<i>Q</i>	<i>P</i>	<i>R</i>
13	125	-1	62	200	296	271	194	57	21	2.75	-1.25	233	0	280	108	0.39
		22	11	184	299	267	196	65	37	2.86	-1.09	209	0	275	110	0.317
14	171	23	8	185	300	265	196	65	37	2.86	-1.09	209	0.5	275	110	0.34
14	768	21	0	190	296	265	194	65	37	2.85	-1.05	209	4	275	110	0.362

Fourier interpretation: Once the phases had been calculated, an E map was produced. It was apparent that considerable difficulty in interpretation could be expected because of the high noise level from phase errors, even though the map was weighted by the phase probability. The electron densities of the E map and average $C2$ map were compared and the known atoms in one map aided in the determination of the equivalent atoms in the other. In addition, the atoms were always selected on the basis of correct stereochemistry and not necessarily on peak strength.

Isotropic refinement: On the basis of the above procedure all 160 non-hydrogen atoms for the four GpC were selected from the E map, although some later had to be corrected. Only one Ca^{2+} and some clearly defined waters could be identified at this time. Since the linked-atom method yielded an ambiguous ψ' value (Fig. 1), site disorder was included for three $\text{O}(5')$ atoms.

The first structure factor calculation contained four GpC, one Ca^{2+} ion, and 24 waters, for 188 atoms. The initial R with 1839 3σ reflections in $P2_1$ was 49.3%. With the block-diagonal program ($LSSQR$), refinement of coordinates followed by temperature factors reduced R to 25.6%. A second Ca^{2+} ion was identified by refinement to a negative temperature factor and was consistent with an average Ca^{2+} site from $C2$. With several incorrect water and $\text{O}(5')$ positions deleted and another difference map, R dropped to 21.7% for four GpC, two Ca^{2+} ions, and 17 waters or 179 atoms.

From this point on, the library system $CRYM$ (kindly provided by G. N. Reeke) was used, which allowed great flexibility in matrix manipulations. The structure was refined in block-matrix form with four atoms related by pseudosymmetry in each block. To improve the refinement, more data were included, down to 1σ (2918 reflections), and the unobserveds were included if $F_c > F_o$. The errors in the structure were then eliminated by successive cycles of difference syntheses and least squares, with careful attention to the geometry of the molecules. R dropped to 14.3% for four GpC, two Ca^{2+} ions and 33 fully occupied waters (195 atoms), and when three disordered water molecules were included R was reduced to 13.1% (198 atoms). The clearly defined disordered water positions with closer than normal contact distances were refined with populations of p and $1-p$ with variable temperature factors.

Anisotropic refinement: Since the structure had only 4239 reflections for 198 non-hydrogen atoms, the system was not sufficiently overdetermined for full anisotropic refinement. Because much work is being currently expended on refining macromolecules (Sieker & Jensen, 1974), it was decided to see how far a structure as large as this could be refined with limited data.

The four P atoms and two Ca^{2+} ions were first refined anisotropically and R for the observed reflections reduced to 12.6%. With Hamilton's (1964) R^*

factor ratio test for all 4237 reflections (R about 3% higher), the significance level was better than 0.5% with 3399 degrees of freedom and 30 new parameters [ratio = 1.032 and $R(30, 3399, 0.005) = 1.008$] (Rogers, 1965).

We then asked whether a significant improvement in the residuals could be obtained if all non-disordered atoms were anisotropic. The result of this calculation was a reduction in R to 9.5%. With the R factor ratio test for all data points once more, the ratio was again significant at better than 0.5%, since the R factor ratio for all reflections was 1.25 and $R(940, 2459, 0.005) = 1.20$. Good refinement was not possible unless all the data were used with the proper weights, because only limited data were available. To improve the anisotropic values it was decided to correct the data for absorption since $\mu = 20 \text{ cm}^{-1}$ and the crystal was $0.03 \times 0.06 \times 0.45 \text{ mm}$. To do this the program $ACACA$ (Wuensch & Prewitt, 1965) was adapted to our computer (IBM 360/91). A Gaussian quadrature routine was appended to the program to use an $8 \times 8 \times 8$ gridwork. When these new results were included and the refinement continued, R reduced to 9.1%. The vibration tensors improved and any unusually low or slightly negative values were reset positive. Considering the low resolution of the data this was a reasonable result, and was better than initially expected.

Two successive difference syntheses were calculated and yielded 61 H atoms. These could be found with difficulty, with heights between 0.2 and 0.4 $e \text{ \AA}^{-3}$, although the map had a high noise level. No attempt was made to find hydroxyl or water H atoms. H atoms could only be refined when all other atoms were held fixed. This reduced R to 8.6%. Those reflections that suffered from secondary extinction (040 and 120) were then weighted zero and further refinement produced a final R of 8.2% (8.5% weighted) for 2918 reflections (11% including unobserved). No corrections were made for anomalous dispersion. The goodness-of-fit ratio* was 1.49 and final shifts were less than the estimated standard deviations. The final difference map showed residual noise at about $\pm 0.4 e \text{ \AA}^{-3}$.

Results and discussion

The four GpC are similar but not identical

The GpC, numbered one through four, obey the approximate relations given in Table 5, which are exact in space group $C2$. The magnitude of the deviation is shown in Table 6. Base-pairing is between GpC related by an approximate twofold axis, not to be confused with the helix axis which is nearly perpendicular to it (Fig. 2). Thus, GpC(1) is base-paired to a unit-cell-translated GpC(2) and similarly for (3) and (4). One Ca(1) of the asymmetric unit is associated with the phosphate group of GpC(1) and GpC(2) and four Ca(1) with GpC(3) and GpC(4). Hence, there is one Ca^{2+} ion

* $R = \frac{\sum |F_o| - |F_c|}{\sum |F_o|}$.

* $[\sum w\Delta^2/(m-n)]^{1/2}$, where $m-n$ = number of degrees of freedom.

for every two GpC in the crystal, and therefore two in the asymmetric unit. The Ca^{2+} ion is unsymmetrically located since it is not on the pseudo twofold axis [Fig. 3(a) and (b)]. With this known result, that only two unsymmetrical Ca^{2+} ions were present, the Patterson map was looked at again to see if these peaks could be located. The peaks were not easily seen since they were masked by pseudosymmetry. However, a Ca-P peak of higher multiplicity was seen, especially when the map was recalculated with the C2 reflections omitted, because this emphasized the Ca^{2+} contributions.

Table 5. Symmetry relations among the four GpC (exact in C2, approximate in P2)

(1)	$x,$	$y,$	z
(2)	$-x,$	$y,$	$-z$
(3)	$\frac{1}{2}+x,$	$\frac{1}{2}+y,$	z
(4)	$\frac{1}{2}-x,$	$\frac{1}{2}+y,$	$-z$

The $\frac{1}{2}$ translation for y in the C2 cell is doubled.

The GpC are paired by six Watson-Crick hydrogen bonds [Fig. 4(a) and (b)]. Fig. 4(c) shows the overall packing and Fig. 4(d) the four GpC in the same orientation, indicating their similarity. As shown in Fig. 5, the molecules pack so as to maximize the overlap of six-membered aromatic rings. Viewed down an axis perpendicular to the plane of the bases, one encounters $-(\text{G}-\text{C})-(\text{C}-\text{G})-(\text{G}-\text{C})-$ where parentheses enclose bases belonging to the same molecule.

Of interest is the $\text{P}\cdots\text{P}$ distance in the dimers, as this gives some measure of the relation of the base-paired GpC to each other. We have two different values for this quantity: 17.66 and 17.65 Å. Another informative quantity is the interglycosidic $\text{C}(1')\cdots\text{C}(1')$ distance, since it gives a measure of the distances between the paired bases. For this there are four values: 10.59, 10.42, 10.53 and 10.66 Å. On the basis of observed conformational angles, estimates of the helical parameters are calculated in Table 7. The data appear to indicate an eightfold helix, in contrast to NaGpC which is tenfold (Rosenberg, Seeman, Day & Rich, 1973). The numbers are not exact integers since a dinucleoside monophosphate has large end effects, as it is only a fragment of a helix.

Conformational angles

In Table 7, we show the values of eight conformational angles (defined in Table 3) for each of the four GpC of the asymmetric unit. All nucleosides are in the *anti* conformation. Although all the observations can

be classified as helical RNA-like, there are significant differences in some angles. These differences are not attributable to the uncertainty in the coordinates which produces an estimated standard error in any dihedral angle of about 2 to 3°. If all four observations are averaged, the estimated standard deviation is reduced to 1–2°. Of the seven angles reported by Day *et al.* (1973) for NaGpC, χ' , ϕ' and ϕ do not fall within the range of our observations for those angles. The differences may be related to the different solvation scheme for the Na^+ and Ca^{2+} salts.

To analyze any significant differences among the four GpC a half-normal probability plot (Abrahams & Keve, 1971) was carried out with a program kindly supplied by De Camp. With only bond lengths, bond angles and dihedral angles, all six different pairs of GpC were compared. Because of the high standard deviations, the largest Δp was only 6 and the slope varied from 1.52 to 2.11, indicating a somewhat underestimated standard deviation (Δp measures the number of estimated standard deviations separating two observations of the same quantity). This was expected, since the structure was too large to refine by full-matrix methods. As shown by De Camp (1973), a more sensitive test of conformational differences is found by comparing all interatomic distances less than 4.65 Å. This was done for all six pairs of GpC, as before. This yielded the result that GpC(2) and GpC(3) and GpC(1) and GpC(4) are approximately the same, with only marginal maximum Δp 's of 7.5 and 7.7 respectively. For the following pairs, (3,4), (2,4), (1,2) and (1,3), the

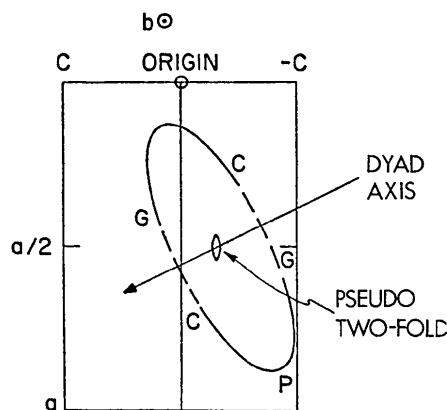


Fig. 2. Watson-Crick base-paired dimer showing the difference between the dyad axis, which is in the direction of the arrow, and the crystallographic pseudo twofold axis which is perpendicular to the plane of the paper.

Table 6. Fractional coordinates of P atoms observed in GpC, and expected values from pseudosymmetry relations of Table 3

Molecule	Observed			Expected			Deviation (Å)
	x	y	z	x	y	z	
(1)	+0.1098	-0.0074	-0.1261				
(2)	+0.8630	+0.0018	+0.1046	+0.8902	-0.0074	+0.1261	0.66
(3)	+0.6236	+0.2524	-0.1312	+0.6098	+0.2426	-0.1261	0.46
(4)	+0.3806	+0.2440	+0.1072	+0.3902	+0.2426	+0.1261	0.25

Table 7. Conformational angles observed for four reported GpC, NaGpC and two recent models of RNA (standard deviations in parentheses)

Species	χ'	ψ'	ϕ'	ω'	ω	ϕ	ψ	χ	Stacking distance (Å)	Repeat	Base parallelism deviation (°)
GpC(1)	8 (3)	53 (3)	222 (2)	294 (2)	291 (2)	181 (2)	47 (3)	33 (2)	3.58	7.8	8.5
GpC(2)	4 (3)	56 (3)	217 (2)	291 (2)	293 (2)	172 (2)	57 (2)	21 (3)	3.73	8.1	7.2
GpC(3)	1 (3)	54 (3)	224 (2)	290 (2)	286 (2)	167 (2)	63 (3)	25 (3)	3.66	7.3	6.5
GpC(4)	7 (3)	53 (3)	216 (2)	288 (2)	283 (2)	181 (2)	52 (3)	28 (3)	3.62	8.3	5.4
NaGpC	13		209	291	284	186	51	25	3.55	10.0	11.6
RNA-11			209	286	298	180	48	12	3.45	11.0	
RNA-12			193	200	295	193	44	17	3.47	12.0	

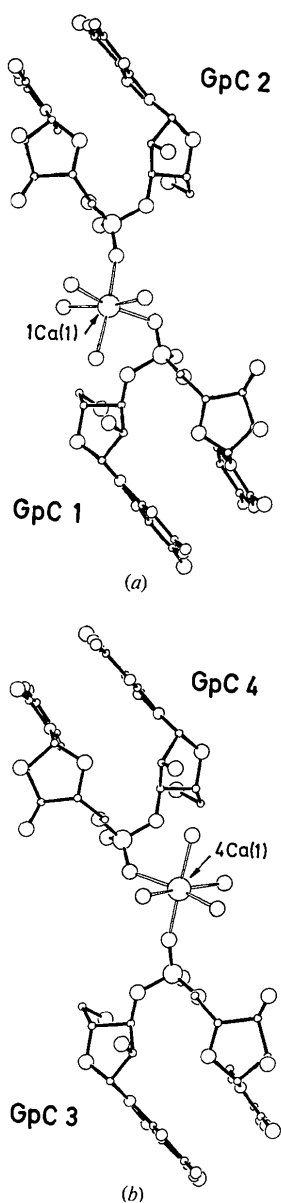


Fig. 3. Calcium coordination for (a) GpC(1,2) and (b) GpC(3,4) showing different calcium orientations about pseudo twofold axis.

maximum Δp 's were 11.3, 11.8, 13.1 and 13.4 respectively, corresponding to differences of up to 0.5 Å. This is explained by the Ca^{2+} ions being associated in the same way with (1) and (4), and in the opposite sense with (2) and (3) (Fig. 3). The modified solvation scheme caused by the Ca^{2+} ions apparently perturbs the GpC into slightly different conformations. The conformational angles χ' , ϕ , ψ , χ are more nearly equal for (2,3) and (1,4) than for the other combinations. NaGpC (Day *et al.*, 1973) is probably more similar to the (1,4) pair since the Na^+ ion is in the same vicinity as the Ca^{2+} ion in our structure. There may also be some differences due to charge and ionic size, particularly for ϕ .

Ribose conformations

All ribose moieties are C(3')-endo but apparently have slightly different pseudorotation parameters (Table 8). Also listed are the τ conformational angles, which agree well with those previously observed for C(3')-endo sugars (Sundaralingam, 1973), within our estimated standard error of 2°. In addition all O(5') atoms are *gauche* to C(3') and *gauche* to O(1') as expected from previous studies. The deviations of C(2'), C(3') and C(5') from the planes defined by O(1'), C(1') and C(4') are given in Table 9 for all eight ribose moieties.

These sugars would be classified by the pseudo-

 Table 8. τ conformational angles (°) and pseudorotation parameter **P** for ribose moieties (standard deviations in parentheses)

	τ_0	τ_1	τ_2	τ_3	τ_4	P
G-1*	8 (2)	-31 (2)	41 (2)	-33 (2)	13 (2)	3
C-1	-4 (2)	-21 (2)	37 (2)	-40 (2)	28 (2)	24
G-2	8 (3)	-30 (2)	38 (2)	-32 (2)	15 (2)	4
C-2	-1 (2)	-27 (2)	40 (2)	-40 (2)	28 (2)	19
G-3	5 (3)	-25 (2)	38 (2)	-33 (2)	17 (2)	10
C-3	6 (3)	-34 (3)	46 (2)	-40 (2)	23 (2)	10
G-4	9 (3)	-29 (2)	40 (2)	-34 (2)	15 (2)	5
C-4	-2 (2)	-30 (2)	44 (2)	-43 (2)	31 (2)	19

Observed ranges of τ angles [C(3')-endo]

τ_0	-11 to 10	τ_3	-40 to -26
τ_1	-31 to -14	τ_4	11 to 33
τ_2	31 to 42		

* Notation: guanine ribose of GpC(1).

rotation scheme of Altona & Sundaralingam (1972) as type *N*. Both sugars in NaGpC and in each molecule of UpA (Sussman, Seeman, Kim & Berman, 1972; Rubin, Brennan & Sundaralingam, 1972) fall into this category, despite great differences in other conformational angles for UpA. The range of in-ring sugar conforma-

tional angles tabulated by Altona & Sundaralingam (1972) for a variety of type *N* structures is given in Table 8.

All but three of the forty values lie within the range of previously observed torsion angles for nucleosides and nucleotides. These three are only 3, 4 and 2° beyond the observed range and, with a 2 to 3° standard deviation, most likely represent new, valid observations.

Table 9. Deviation (Å) of C(2'), C(3') and C(5') from the plane of C(1'), C(4'), O(1') for ribose moieties

	C(2')	C(3')	C(5')		C(2')	C(3')	C(5')
G-1	-0.21	+0.36	+0.78	G-3	-0.13	+0.45	+0.77
C-1	+0.11	+0.69	+0.51	C-3	-0.14	+0.63	+0.69
G-2	-0.22	+0.40	+0.77	G-4	-0.22	+0.43	+0.94
C-2	+0.04	+0.74	+0.75	C-4	+0.06	+0.83	+0.81

Atomic coordinates and data

A list of atomic coordinates is given in Table 10 with their estimated standard deviations from the blocked matrices, where each of four atoms from corresponding GpC are included in the same matrix. The temperature

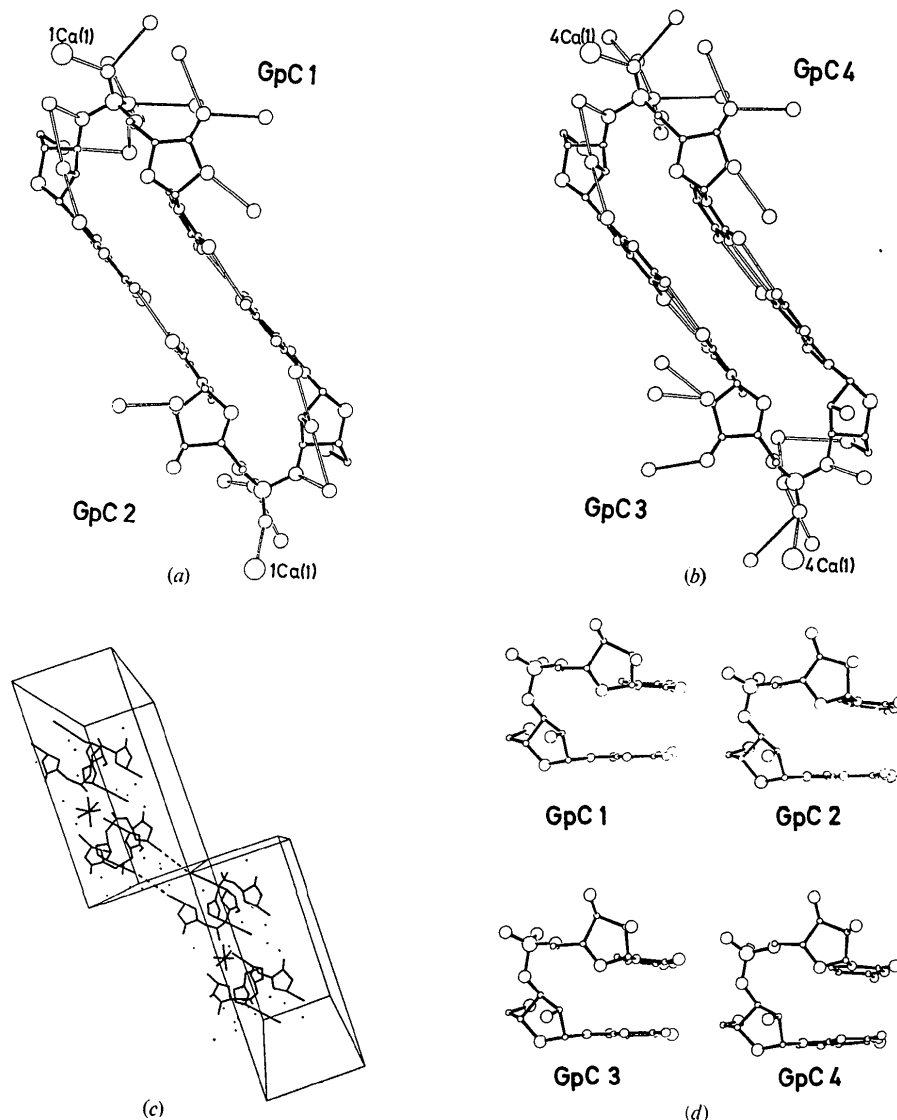


Fig. 4. (a) GpC(1,2) and (b) GpC(3,4) Watson-Crick base-paired in the crystal, with water and calcium interactions shown. Water interactions with base moieties are omitted for clarity. (c) Overall packing scheme. (d) The four GpC shown in the same orientation indicating basic similarity.

Table 10. *Coordinates and temperature factors*

The coordinates are in fractions of the unit cell ($\times 10^4$, for H $\times 10^3$) and the temperature factors are in the following form:

$$\exp [-(B_{11}h^2 + B_{22}k^2 + B_{33}l^2 + 2B_{12}hk + 2B_{13}hl + 2B_{23}kl)].$$

B_{11} , B_{22} and B_{12} are multiplied by 10^4 ; B_{33} , B_{13} and B_{23} are multiplied by 10^3 . The estimated standard deviations are given in parentheses and correspond to the last position. Isotropic temperature factors are given as $\exp(-B \sin^2 \theta/\lambda^2)$. Hydrogen atoms and disordered waters are given isotropic B 's. The population factors for disordered waters were constrained to p and $1-p$ for closely contacting atoms and are as follows:

$2W(8)A$	0.36 (4)	$3W(5)A$	0.42 (5)
$2W(8)B$	0.64 (4)	$3W(5)B$	0.58 (5)
$1W(6)A$	0.59 (4)	$3W(6)A$	0.31 (4)
$1W(6)B$	0.41 (4)	$3W(6)B$	0.69 (4).

Notation: (N9)1G is atom N(9) of molecule (1) in the guanine residue.

	<i>x</i>	<i>y</i>	<i>z</i>	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
(N9)1G	3512 (8)	-429 (6)	161 (24)	21 (6)	11 (3)	14 (4)	4 (8)	3 (3)	4 (2)
(N9)2G	6219 (8)	-247 (6)	-655 (23)	17 (5)	8 (3)	15 (4)	-18 (7)	3 (3)	-2 (2)
(N9)3G	8633 (8)	2105 (6)	124 (24)	25 (6)	3 (2)	17 (5)	2 (7)	2 (3)	3 (2)
(N9)4G	1378 (8)	2142 (6)	-543 (22)	17 (6)	11 (3)	7 (4)	-14 (7)	5 (2)	3 (2)
(C8)1G	3468 (10)	-805 (8)	167 (28)	16 (7)	14 (4)	8 (5)	-14 (9)	-1 (3)	-5 (2)
(C8)2G	6223 (9)	-653 (7)	-774 (31)	7 (6)	7 (3)	21 (6)	-1 (8)	-4 (3)	-2 (2)
(C8)3G	8580 (12)	1699 (8)	49 (31)	42 (9)	7 (3)	14 (6)	-16 (10)	3 (4)	3 (2)
(C8)4G	1391 (13)	1733 (8)	-636 (39)	51 (11)	7 (4)	33 (8)	-39 (10)	0 (5)	2 (3)
(N7)1G	3864 (10)	-1014 (6)	-717 (24)	58 (9)	8 (3)	11 (4)	-7 (8)	-4 (3)	4 (2)
(N7)2G	5804 (8)	-822 (6)	89 (26)	15 (6)	10 (3)	25 (5)	-10 (7)	2 (3)	0 (2)
(N7)3G	8956 (8)	1537 (8)	-875 (27)	13 (6)	9 (3)	27 (5)	-21 (6)	1 (3)	-2 (2)
(N7)4G	922 (10)	1581 (8)	221 (27)	46 (8)	32 (5)	16 (5)	-15 (11)	14 (4)	-3 (3)
(C5)1G	4205 (12)	682 (9)	-1313 (31)	39 (10)	18 (5)	8 (5)	15 (11)	-7 (4)	-4 (3)
(C5)2G	5517 (10)	-470 (7)	762 (31)	14 (7)	7 (3)	20 (6)	-16 (8)	0 (4)	-6 (3)
(C5)3G	9298 (9)	1880 (7)	-1271 (28)	9 (7)	9 (3)	11 (5)	4 (8)	5 (3)	4 (2)
(C5)4G	635 (10)	1948 (8)	766 (25)	16 (7)	14 (4)	3 (4)	-11 (8)	6 (3)	-6 (2)
(C6)1G	4695 (9)	-613 (9)	-2444 (27)	1 (6)	23 (5)	5 (4)	3 (9)	7 (3)	-1 (3)
(C6)2G	5061 (10)	-462 (9)	1860 (26)	18 (8)	16 (4)	10 (5)	-2 (10)	-3 (3)	6 (3)
(C6)3G	9772 (10)	1921 (8)	-2380 (31)	19 (8)	13 (4)	13 (6)	10 (9)	2 (4)	-3 (3)
(C6)4G	157 (10)	1938 (8)	1837 (31)	14 (8)	14 (4)	16 (6)	20 (9)	-5 (4)	-4 (3)
(N1)1G	4827 (9)	-259 (6)	-2741 (25)	27 (6)	3 (2)	18 (5)	-14 (7)	3 (3)	5 (2)
(N1)2G	4882 (7)	-119 (6)	2185 (20)	6 (5)	12 (3)	6 (3)	14 (6)	4 (2)	-5 (2)
(N1)3G	9966 (8)	2288 (6)	-2657 (24)	24 (7)	6 (3)	11 (4)	-9 (7)	-3 (3)	1 (2)
(N1)4G	61 (8)	2301 (6)	2326 (24)	14 (6)	8 (3)	12 (4)	-3 (7)	-1 (3)	1 (2)
(C2)1G	4592 (10)	76 (7)	-2171 (30)	9 (7)	7 (3)	17 (5)	1 (7)	-6 (3)	-3 (2)
(C2)2G	5140 (11)	246 (10)	1738 (26)	29 (8)	23 (6)	1 (4)	15 (12)	-3 (3)	3 (3)
(C2)3G	9774 (9)	2648 (9)	-1972 (29)	1 (6)	21(5)	12 (5)	-6 (9)	5 (3)	-5 (3)
(C2)4G	326 (11)	2665 (8)	1918 (27)	22 (8)	15 (4)	3 (4)	2 (10)	1 (3)	2 (2)
(N2)1G	4748 (10)	422 (6)	-2611 (26)	49 (8)	1 (2)	16 (5)	-17 (8)	-1 (3)	1 (2)
(N2)2G	4933 (9)	551 (6)	2242 (28)	28 (7)	5 (3)	22 (5)	5 (7)	5 (3)	-2 (2)
(N2)3G	9986 (8)	2974 (7)	-2386 (26)	17 (6)	11 (3)	19 (5)	2 (7)	5 (3)	4 (2)
(N2)4G	140 (10)	2989 (6)	2417 (28)	54 (9)	3 (3)	23 (6)	8 (8)	15 (4)	3 (2)
(C4)1G	3972 (10)	-305 (7)	-787 (25)	21 (7)	9 (3)	3 (4)	-11 (8)	4 (3)	-6 (2)
(C4)2G	5775 (10)	-108 (8)	363 (25)	23 (8)	19 (4)	3 (4)	-30 (10)	0 (3)	0 (3)
(C4)3G	9088 (10)	2192 (7)	-795 (26)	20 (7)	8 (3)	3 (4)	12 (8)	4 (3)	0 (2)
(C4)4G	937 (11)	2304 (9)	445 (31)	20 (8)	17 (4)	14 (6)	-13 (10)	-2 (4)	7 (3)
(O6)1G	4944 (7)	-916 (6)	-3043 (20)	20 (5)	13 (3)	15 (4)	19 (6)	5 (2)	0 (2)
(O6)2G	4834 (7)	-768 (5)	2378 (20)	22 (5)	4 (2)	19 (4)	9 (5)	7 (2)	0 (2)
(O6)3G	10002 (8)	1629 (5)	-3044 (20)	37 (6)	7 (2)	16 (4)	-7 (6)	6 (3)	2 (2)
(O6)4G	-184 (7)	1652 (6)	2379 (22)	23 (6)	14 (3)	23 (4)	-19 (7)	7 (3)	-3 (2)
(N3)1G	4118 (7)	53 (6)	-1118 (23)	13 (5)	8 (3)	14 (4)	3 (6)	7 (2)	3 (2)
(N3)2G	5612 (8)	253 (6)	750 (21)	34 (6)	6 (2)	6 (3)	5 (7)	3 (2)	-4 (2)
(N3)3G	9287 (8)	2604 (7)	-1042 (25)	17 (6)	16 (3)	13 (4)	-17 (7)	-4 (3)	-3 (2)
(N3)4G	810 (9)	2650 (6)	892 (25)	23 (6)	10 (3)	16 (4)	3 (7)	-3 (3)	0 (2)
(N4)1C	3859 (8)	-751 (6)	-5479 (24)	23 (6)	4 (2)	15 (4)	-13 (6)	4 (3)	2 (2)
(N4)2C	5946 (11)	-810 (7)	5015 (27)	61 (9)	11 (3)	13 (5)	1 (9)	5 (4)	0 (2)
(N4)3C	8887 (8)	1666 (7)	-5541 (27)	12 (6)	14 (3)	23 (5)	5 (7)	0 (3)	3 (2)
(N4)4C	1042 (9)	1695 (7)	4859 (24)	28 (7)	14 (3)	10 (4)	-13 (8)	7 (3)	2 (2)
(C4)1C	3627 (9)	-388 (6)	-5110 (28)	12 (7)	1 (3)	14 (5)	6 (7)	-1 (3)	0 (2)
(C4)2C	6187 (12)	-438 (8)	4691 (26)	55 (10)	8 (4)	1 (4)	-14 (11)	2 (4)	3 (2)
(C4)3C	8733 (9)	2044 (9)	-5206 (27)	2 (6)	24 (5)	5 (4)	-13 (10)	1 (3)	-3 (2)
(C4)4C	1251 (9)	2035 (8)	4679 (29)	8 (7)	14 (4)	11 (5)	-13 (9)	-7 (3)	-3 (3)
(C5)1C	3065 (9)	-387 (9)	-4196 (27)	6 (6)	17 (4)	8 (5)	4 (9)	3 (3)	-1 (3)
(C5)2C	6697 (10)	-412 (9)	3691 (26)	15 (7)	21 (5)	2 (4)	-15 (10)	6 (3)	-1 (3)
(C5)3C	8169 (12)	2103 (9)	-4161 (27)	44 (10)	14 (4)	1 (4)	9 (11)	-3 (4)	3 (2)
(C5)4C	1777 (10)	2119 (9)	3799 (29)	13 (7)	14 (4)	8 (5)	9 (9)	-6 (3)	-5 (3)
(C6)1C	2829 (12)	10 (9)	-4016 (29)	33 (9)	16 (4)	8 (5)	-5 (4)	-5 (4)	-7 (3)

Table 10 (cont.)

	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> ₁₁	<i>B</i> ₂₂	<i>B</i> ₃₃	<i>B</i> ₁₂	<i>B</i> ₁₃	<i>B</i> ₂₃
(C6)2C	6880 (10)	-24 (7)	3592 (26)	18 (7)	6 (3)	7 (4)	-4 (3)	-4 (3)	-1 (2)
(C6)3C	7977 (10)	2469 (7)	-3835 (28)	18 (8)	8 (3)	7 (5)	-4 (3)	-4 (3)	0 (2)
(C6)4C	2085 (10)	2463 (8)	3645 (31)	19 (8)	14 (4)	11 (5)	-5 (4)	-5 (4)	3 (3)
(N1)1C	3151 (7)	322 (6)	-4495 (23)	12 (5)	7 (3)	16 (4)	1 (3)	1 (3)	2 (2)
(N1)2C	6543 (8)	303 (6)	4073 (23)	23 (6)	9 (3)	11 (4)	-1 (3)	-1 (3)	-4 (2)
(N1)3C	8346 (8)	2776 (6)	-4392 (22)	20 (6)	7 (3)	6 (4)	0 (3)	0 (3)	-1 (2)
(N1)4C	1761 (9)	2783 (7)	4259 (25)	26 (7)	13 (3)	13 (4)	3 (3)	3 (3)	4 (2)
(N3)1C	3937 (8)	-99 (5)	-5638 (22)	24 (6)	2 (2)	9 (4)	1 (3)	1 (3)	1 (2)
(N3)2C	5849 (8)	-138 (6)	5203 (23)	23 (6)	7 (3)	14 (4)	0 (3)	0 (3)	-2 (2)
(N3)3C	9023 (8)	2377 (6)	-5622 (21)	21 (6)	7 (3)	6 (4)	3 (3)	3 (3)	2 (2)
(N3)4C	996 (8)	2406 (7)	5316 (22)	23 (6)	10 (3)	7 (4)	0 (3)	0 (3)	2 (2)
(C2)1C	3668 (9)	233 (7)	-5423 (25)	8 (6)	6 (3)	5 (4)	-3 (3)	-3 (3)	-3 (2)
(C2)2C	6039 (11)	238 (7)	4994 (23)	30 (8)	4 (3)	9 (5)	-2 (3)	-2 (3)	8 (2)
(C2)3C	8859 (13)	2713 (9)	-5299 (33)	42 (10)	18 (5)	7 (5)	1 (4)	1 (4)	5 (3)
(C2)4C	1216 (11)	2774 (9)	5152 (32)	23 (8)	20 (5)	11 (5)	-2 (4)	-2 (4)	1 (3)
(O2)1C	3867 (7)	566 (5)	-5933 (18)	32 (5)	7 (2)	8 (3)	4 (2)	4 (2)	1 (1)
(O2)2C	5750 (8)	512 (5)	5519 (23)	45 (7)	7 (2)	22 (4)	5 (3)	5 (3)	-1 (2)
(O2)3C	9141 (7)	3023 (5)	-5772 (24)	30 (8)	11 (3)	27 (5)	14 (3)	14 (3)	-1 (2)
(O2)4C	1051 (7)	3055 (5)	5794 (20)	31 (5)	7 (2)	16 (4)	5 (3)	5 (3)	2 (2)
(C1')1C	2880 (11)	703 (8)	-4289 (29)	38 (9)	9 (3)	6 (5)	13 (10)	4 (3)	-5 (2)
(C1')2C	6665 (15)	697 (8)	3934 (28)	82 (14)	10 (4)	1 (4)	1 (13)	9 (4)	2 (2)
(C1')3C	8164 (10)	3150 (8)	-4121 (26)	22 (8)	11 (4)	3 (4)	14 (8)	7 (3)	1 (2)
(C1')4C	1993 (10)	3190 (7)	4131 (25)	22 (7)	8 (3)	1 (4)	-6 (8)	8 (3)	-3 (2)
(C2')1C	2382 (10)	790 (7)	-5476 (28)	24 (8)	9 (3)	4 (4)	16 (8)	0 (3)	-2 (2)
(C2')2C	7140 (14)	847 (7)	5207 (27)	75 (13)	5 (3)	1 (4)	20 (11)	7 (4)	-1 (2)
(C2')3C	7759 (14)	3328 (9)	-5321 (32)	58 (12)	15 (5)	11 (6)	32 (12)	3 (5)	-1 (3)
(C2')4C	2508 (11)	3286 (7)	5227 (28)	50 (11)	2 (3)	7 (5)	-17 (9)	-2 (4)	-3 (2)
(C3')1C	1765 (11)	664 (8)	-4792 (25)	25 (8)	11 (4)	1 (4)	3 (9)	-2 (3)	0 (2)
(C3')2C	7848 (11)	746 (9)	4747 (28)	29 (9)	16 (4)	7 (5)	-12 (10)	-5 (4)	3 (3)
(C3')3C	7079 (10)	3212 (8)	-4775 (31)	12 (7)	9 (3)	17 (6)	9 (9)	-1 (4)	1 (3)
(C3')4C	3162 (10)	3146 (7)	4651 (31)	14 (7)	7 (3)	16 (6)	-4 (8)	-3 (3)	1 (2)
(C4')1C	1865 (10)	791 (6)	-3225 (28)	28 (8)	1 (3)	9 (5)	16 (8)	-4 (3)	0 (2)
(C4')2C	7733 (11)	867 (7)	3039 (29)	35 (9)	8 (4)	7 (5)	-22 (9)	3 (3)	2 (2)
(C4')3C	7152 (13)	3321 (8)	-3101 (27)	56 (11)	10 (4)	1 (4)	0 (11)	-5 (4)	0 (2)
(C4')4C	2999 (11)	3327 (8)	3065 (29)	26 (8)	5 (3)	9 (5)	19 (6)	-2 (3)	3 (2)
(C5')1C	1476 (11)	635 (8)	-2026 (23)	47 (10)	13 (4)	3 (4)	50 (11)	10 (4)	6 (2)
(C5')2C	8234 (10)	681 (7)	2039 (29)	23 (8)	7 (3)	9 (5)	17 (9)	2 (3)	-2 (2)
(C5')3C	6613 (10)	3182 (8)	-2223 (28)	18 (7)	10 (4)	5 (5)	-15 (8)	5 (3)	-1 (2)
(C5')4C	3452 (10)	3125 (8)	1908 (33)	15 (7)	12 (4)	17 (6)	16 (9)	2 (4)	7 (3)
(O5')1C	1474 (7)	198 (5)	-2290 (20)	29 (5)	10 (2)	10 (3)	11 (6)	1 (2)	-3 (2)
(O5')2C	8216 (7)	266 (5)	2175 (20)	27 (5)	9 (2)	11 (3)	7 (6)	2 (2)	1 (2)
(O5')3C	6638 (7)	2773 (4)	-2363 (19)	33 (5)	6 (2)	7 (3)	15 (5)	4 (2)	0 (1)
(O5')4C	3402 (7)	2715 (5)	2055 (19)	36 (6)	6 (2)	6 (3)	-7 (6)	0 (2)	-1 (1)
P(1)	1098 (3)	-74 (2)	-1261 (9)	23 (2)	10 (1)	7 (1)	9 (3)	3 (1)	0 (1)
P(2)	8630 (3)	18 (2)	1046 (9)	18 (2)	7 (1)	10 (1)	3 (2)	3 (1)	0 (1)
P(3)	6236 (3)	2524 (3)	-1312 (9)	23 (2)	11 (1)	6 (1)	0 (3)	0 (1)	0 (1)
P(4)	3806 (4)	2440 (3)	1072 (9)	25 (2)	12 (1)	8 (1)	-7 (3)	3 (1)	2 (1)
(O2')1C	2403 (8)	1206 (5)	-5705 (22)	59 (7)	8 (2)	20 (4)	-1 (7)	17 (3)	-3 (2)
(O2')2C	6999 (9)	1233 (5)	5214 (20)	34 (8)	6 (2)	12 (3)	-6 (7)	12 (3)	-4 (2)
(O2')3C	7840 (10)	3761 (5)	-5489 (23)	82 (9)	8 (3)	22 (4)	-9 (8)	11 (4)	2 (2)
(O2')4C	2471 (8)	3720 (5)	5501 (20)	43 (7)	9 (2)	14 (4)	7 (6)	3 (3)	-1 (2)
(O3')1C	1238 (7)	865 (5)	-5392 (19)	26 (5)	13 (3)	12 (3)	-4 (6)	7 (2)	1 (2)
(O3')2C	8277 (9)	962 (6)	5191 (21)	55 (7)	18 (3)	11 (4)	-46 (8)	-1 (3)	5 (2)
(O3')3C	6608 (9)	3460 (5)	-5422 (19)	64 (8)	12 (3)	7 (3)	-9 (7)	10 (3)	-2 (2)
(O3')4C	3639 (7)	3356 (5)	5268 (19)	32 (6)	14 (3)	11 (3)	-17 (6)	-4 (2)	2 (2)
(O1')1C	2540 (6)	688 (5)	-2960 (17)	11 (4)	13 (2)	5 (3)	-13 (5)	0 (2)	-3 (2)
(O1')2C	7107 (7)	721 (5)	2665 (18)	35 (6)	11 (2)	4 (3)	-10 (5)	4 (2)	2 (2)
(O1')3C	7798 (8)	3173 (6)	-2863 (18)	31 (5)	18 (3)	3 (3)	9 (7)	0 (2)	1 (2)
(O1')4C	2346 (7)	3213 (6)	2813 (19)	31 (6)	14 (3)	6 (3)	-8 (6)	2 (2)	-3 (2)
O1(P1)	504 (7)	83 (5)	-902 (19)	23 (5)	12 (3)	11 (3)	3 (6)	5 (2)	-3 (2)
O1(P2)	9258 (6)	193 (5)	846 (20)	11 (4)	9 (2)	15 (3)	8 (6)	-1 (2)	0 (2)
O1(P3)	5640 (7)	2721 (6)	-1064 (22)	21 (5)	16 (3)	18 (4)	5 (6)	7 (3)	-2 (2)
O1(P4)	4424 (7)	2594 (5)	719 (19)	22 (5)	10 (2)	7 (3)	-4 (6)	-1 (2)	-1 (2)
O2(P1)	1126 (7)	-452 (6)	-1922 (20)	27 (5)	13 (3)	13 (4)	-3 (6)	-3 (2)	-1 (2)
O2(P2)	8554 (7)	-387 (5)	1454 (20)	33 (5)	6 (2)	16 (4)	-18 (6)	4 (2)	2 (2)
O2(P3)	6202 (7)	2113 (5)	-1837 (19)	32 (6)	9 (2)	9 (3)	-14 (6)	-1 (2)	0 (2)
O2(P4)	3784 (7)	2026 (5)	1660 (19)	24 (5)	11 (3)	9 (3)	-1 (6)	-2 (2)	2 (2)
(O3')1G	1481 (6)	-49 (5)	168 (17)	18 (5)	8 (2)	9 (3)	6 (5)	2 (2)	1 (2)
(O3')2G	8259 (6)	96 (5)	-446 (18)	21 (5)	11 (2)	9 (3)	-7 (6)	5 (2)	-1 (2)
(O3')3G	6015 (7)	2555 (5)	147 (21)	24 (5)	12 (3)	18 (4)	1 (6)	10 (2)	2 (2)

Table 10 (cont.)

	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> ₁₁	<i>B</i> ₂₂	<i>B</i> ₃₃	<i>B</i> ₁₂	<i>B</i> ₁₃	<i>B</i> ₂₃
(O3')4G	3400 (6)	2467 (5)	-396 (18)	18 (5)	6 (2)	9 (3)	-1 (5)	0 (2)	2 (1)
(C3')1G	2097 (11)	-197 (7)	323 (27)	49 (9)	4 (3)	4 (5)	11 (9)	3 (3)	2 (2)
(C3')2G	7670 (11)	-57 (7)	-702 (27)	21 (7)	11 (4)	8 (4)	-8 (9)	-2 (3)	1 (2)
(C3')3G	7227 (9)	2359 (7)	338 (23)	3 (6)	6 (3)	10 (5)	-6 (7)	-1 (3)	-1 (2)
(C3')4G	2312 (9)	2294 (6)	-635 (23)	2 (6)	1 (3)	15 (5)	-6 (6)	3 (3)	-2 (2)
(C4')1G	2120 (9)	-411 (7)	1832 (29)	6 (6)	10 (4)	13 (5)	8 (9)	-4 (3)	-5 (3)
(C4')2G	7597 (10)	-187 (8)	-2361 (30)	16 (7)	16 (4)	13 (5)	14 (9)	-1 (3)	-9 (3)
(C4')3G	7205 (10)	2166 (9)	1866 (32)	13 (7)	16 (4)	18 (6)	16 (10)	0 (4)	1 (3)
(C4')4G	2771 (9)	2138 (9)	-2251 (29)	6 (6)	23 (5)	13 (5)	18 (10)	-12 (3)	-8 (3)
(C5')1G	1843 (15)	-790 (8)	2059 (32)	75 (13)	8 (4)	9 (5)	-25 (12)	-7 (5)	1 (2)
(C5')2G	7959 (14)	-561 (9)	-2803 (37)	72 (12)	16 (5)	25 (8)	48 (13)	-9 (5)	-10 (3)
(C5')3G	6888 (13)	1813 (8)	2040 (35)	53 (11)	11 (4)	29 (7)	-2 (11)	28 (5)	-3 (3)
(C5')4G	3019 (15)	1733 (8)	-2460 (36)	81 (14)	12 (4)	23 (7)	10 (13)	24 (5)	-4 (3)
(O5')1G	2084 (9)	-1060 (6)	1002 (22)	64 (8)	14 (3)	12 (4)	-7 (8)	-2 (3)	1 (2)
(O5')2G	7732 (9)	-887 (7)	-1789 (34)	39 (6)	18 (3)	71 (9)	14 (7)	19 (4)	-11 (3)
(O5')3G	7092 (8)	1502 (6)	1076 (23)	40 (6)	12 (3)	23 (4)	4 (7)	-3 (3)	5 (2)
(O5')4G	2727 (9)	1459 (6)	-1500 (23)	51 (7)	19 (3)	19 (4)	-7 (8)	5 (3)	-7 (2)
(O1')1G	2821 (8)	-401 (5)	2055 (18)	43 (6)	14 (3)	5 (3)	7 (7)	1 (2)	2 (2)
(O1')2G	6923 (7)	-159 (5)	-2583 (17)	29 (5)	19 (3)	3 (3)	4 (6)	-2 (2)	-1 (2)
(O1')3G	7919 (7)	2140 (6)	2076 (20)	40 (6)	16 (3)	12 (4)	1 (7)	5 (3)	4 (2)
(O1')4G	2078 (7)	2170 (6)	-2465 (19)	24 (5)	21 (3)	10 (3)	-2 (7)	5 (2)	-3 (2)
(C1')1G	3149 (10)	-137 (8)	1057 (25)	25 (8)	14 (4)	1 (4)	-15 (9)	11 (3)	-4 (2)
(C1')2G	6586 (12)	53 (8)	-1512 (29)	43 (10)	11 (4)	8 (5)	-10 (10)	11 (4)	-6 (2)
(C1')3G	8305 (12)	2380 (8)	1134 (30)	30 (9)	12 (4)	11 (5)	-3 (10)	1 (4)	6 (3)
(C1')4G	1744 (12)	2413 (7)	-1374 (32)	40 (9)	7 (3)	13 (6)	-12 (9)	-2 (4)	-3 (2)
(C2')1G	2584 (10)	63 (7)	344 (28)	17 (7)	7 (3)	14 (5)	-11 (8)	-1 (3)	9 (2)
(C2')2G	7124 (10)	232 (9)	-587 (27)	10 (7)	26 (5)	4 (4)	-9 (10)	-2 (3)	-5 (3)
(C2')3G	7802 (11)	2624 (8)	269 (26)	20 (8)	10 (4)	9 (5)	7 (9)	-3 (3)	2 (2)
(C2')4G	2287 (11)	2610 (7)	-469 (26)	26 (8)	11 (3)	3 (4)	-34 (9)	6 (2)	-1 (2)
(O2')1G	2489 (7)	435 (5)	946 (20)	35 (6)	8 (2)	11 (3)	2 (6)	-1 (2)	-3 (2)
(O2')2G	7285 (7)	612 (4)	-1065 (18)	36 (5)	7 (2)	8 (3)	-12 (6)	2 (2)	2 (1)
(O2')3G	7666 (7)	2970 (5)	1027 (9)	32 (5)	10 (2)	9 (3)	0 (6)	4 (2)	0 (2)
(O2')4G	2462 (6)	2953 (5)	-1134 (19)	17 (5)	17 (3)	10 (3)	-24 (6)	-2 (2)	2 (2)
1Ca(1)	209 (2)	426 (0)	1227 (6)	22 (3)	8 (2)	11 (1)	-1 (1)	2 (1)	-1 (1)
4Ca(1)	4680 (2)	2967 (2)	-1312 (6)	23 (3)	11 (2)	7 (1)	1 (2)	3 (1)	0 (1)
1W(3)	54 (6)	988 (6)	-316 (19)	32 (6)	7 (2)	11 (3)	4 (6)	1 (2)	2 (2)
3W(3)	4868 (8)	3506 (6)	235 (17)	4 (6)	7 (2)	4 (3)	-1 (6)	1 (2)	0 (1)
1W(4)	47 (7)	862 (4)	-4143 (20)	31 (5)	3 (2)	14 (3)	-2 (5)	1 (2)	0 (2)
2W(4)	-130 (7)	879 (5)	-7111 (20)	16 (4)	9 (2)	17 (4)	12 (5)	6 (2)	-1 (2)
3W(4)	5055 (8)	3424 (5)	-2979 (18)	48 (6)	8 (3)	7 (3)	-6 (6)	1 (2)	0 (2)
4W(4)	4863 (7)	3312 (6)	4009 (21)	21 (5)	16 (3)	18 (4)	-2 (7)	4 (2)	3 (2)
1W(10)	1085 (11)	-436 (7)	-5048 (26)	107 (11)	23 (4)	21 (5)	43 (11)	13 (4)	1 (3)
2W(10)	9222 (10)	-190 (9)	4680 (35)	57 (8)	48 (6)	49 (8)	-20 (13)	2 (5)	-4 (4)
3W(10)	5705 (13)	2253 (9)	-4906 (34)	119 (15)	39 (6)	39 (7)	56 (15)	-2 (6)	0 (4)
4W(10)	3878 (10)	2124 (9)	4855 (26)	61 (8)	41 (5)	18 (4)	-15 (11)	-1 (3)	4 (3)
1W(1)	352 (8)	-11 (6)	3167 (22)	28 (6)	21 (3)	23 (4)	-2 (7)	-3 (3)	7 (2)
2W(1)	9478 (8)	224 (6)	-2746 (22)	37 (6)	20 (3)	18 (4)	5 (8)	1 (3)	4 (2)
3W(1)	5435 (9)	2723 (6)	2630 (21)	44 (7)	21 (3)	14 (4)	3 (8)	-3 (3)	-4 (2)
4W(1)	4619 (10)	2502 (7)	-3280 (24)	59 (8)	32 (5)	14 (4)	5 (10)	1 (3)	-4 (2)
1W(2)	1225 (7)	682 (5)	1724 (20)	32 (6)	16 (3)	11 (4)	10 (7)	-2 (2)	2 (2)
2W(2)	8563 (7)	815 (5)	-2045 (18)	34 (6)	12 (3)	5 (3)	-3 (6)	-2 (2)	-1 (2)
3W(2)	6376 (7)	3260 (6)	1876 (19)	36 (6)	13 (2)	5 (3)	22 (7)	-3 (2)	-2 (2)
4W(2)	3682 (6)	3326 (6)	-1795 (18)	8 (4)	21 (3)	5 (3)	-1 (6)	-2 (2)	4 (2)
1W(9)	3535 (10)	797 (7)	74 (32)	62 (8)	21 (4)	44 (7)	17 (9)	-9 (4)	0 (3)
2W(9)	6208 (8)	945 (6)	-205 (25)	39 (6)	16 (3)	24 (5)	-6 (8)	3 (3)	2 (2)
3W(9)	8810 (10)	3335 (7)	187 (25)	60 (8)	19 (3)	20 (4)	-11 (9)	4 (3)	1 (2)
4W(9)	1354 (10)	3387 (8)	-325 (30)	5 (8)	29 (5)	37 (6)	-3 (10)	11 (4)	6 (3)
1W(7)	4336 (12)	1299 (8)	-4771 (28)	97 (11)	19 (4)	29 (6)	-43 (11)	9 (4)	-5 (3)
2W(7)	5483 (12)	1464 (12)	4605 (47)	52 (9)	63 (9)	93 (14)	31 (15)	9 (6)	-23 (6)
3W(7)	9563 (15)	3822 (10)	-4707 (33)	123 (15)	34 (5)	37 (6)	-26 (14)	6 (6)	-6 (4)
4W(7)	648 (13)	3944 (13)	4948 (37)	72 (11)	72 (9)	45 (8)	64 (17)	1 (5)	3 (5)
1W(8)	4358 (12)	1224 (9)	-1887 (37)	92 (12)	28 (5)	54 (9)	33 (12)	-13 (4)	-2 (3)
3W(8)	9667 (16)	3752 (11)	-1666 (57)	100 (14)	40 (7)	111 (16)	21 (15)	17 (8)	-14 (6)
4W(8)	831 (13)	3789 (8)	2075 (31)	107 (13)	23 (4)	35 (6)	13 (12)	15 (5)	-1 (3)
1W(5)	3102 (16)	1397 (8)	2061 (28)	195 (19)	26 (4)	25 (5)	-70 (16)	41 (6)	-6 (3)
2W(5)	7111 (11)	1546 (6)	-2153 (21)	100 (10)	14 (3)	10 (3)	6 (8)	14 (3)	-3 (2)
4W(5)	1507 (13)	3927 (8)	-2811 (38)	112 (13)	21 (4)	71 (10)	17 (13)	35 (6)	-7 (3)

Table 10 (cont.)

	x	y	z	B		x	y	z	B
2W(8)A	5407 (40)	1370 (29)	2019 (90)	17 (4)	H(2')C2	722 (5)	74 (3)	593 (17)	7 (4)
2W(8)B	4532 (18)	1414 (12)	874 (47)	13 (2)	H(1')C2	611 (6)	90 (4)	356 (20)	13 (6)
1W(6)A	-378 (18)	-704 (12)	77 (45)	11 (2)	H(5')G2	770 (4)	-60 (2)	-305 (12)	1 (3)
1W(6)B	350 (26)	-962 (13)	-888 (68)	11 (3)	H(3')G2	763 (6)	-29 (5)	-48 (19)	12 (6)
3W(5)A	7898 (20)	3978 (13)	2085 (49)	4 (2)	H(N2)G3	962 (4)	331 (3)	-209 (14)	11 (5)
3W(5)B	8541 (34)	3847 (19)	2365 (78)	22 (3)	H(N2)G3	1024 (4)	282 (3)	-244 (12)	1 (3)
3W(6)A	4999 (19)	1849 (13)	-451 (50)	3 (2)	H(N1)G3	1012 (5)	206 (3)	-247 (16)	5 (4)
3W(6)B	5422 (19)	1511 (13)	-1138 (47)	14 (2)	H(C8)G3	835 (4)	164 (3)	45 (13)	2 (3)
H(N2)G1	472 (5)	80 (3)	-221 (15)	3 (4)	H(N2)C3	862 (5)	153 (3)	-525 (17)	7 (4)
H(N2)G1	512 (5)	38 (4)	-244 (17)	8 (5)	H(C5)C3	826 (5)	195 (3)	-362 (16)	5 (4)
H(N1)G1	515 (7)	-34 (5)	-308 (22)	12 (7)	H(C6)C3	737 (5)	261 (4)	-316 (17)	6 (5)
H(C8)G1	307 (7)	-103 (5)	92 (12)	11 (6)	H(4')C3	707 (5)	367 (3)	-281 (17)	6 (5)
H(N2)C1	404 (6)	-118 (4)	-502 (18)	9 (5)	H(3')C3	696 (7)	279 (4)	-501 (21)	11 (6)
H(N2)C1	427 (5)	-81 (3)	-627 (16)	7 (7)	H(2')C3	756 (7)	319 (5)	-636 (21)	14 (6)
H(C5)C1	292 (8)	-51 (6)	-383 (26)	17 (8)	H(1')C3	862 (6)	326 (4)	-355 (18)	9 (5)
H(C6)C1	237 (4)	4 (3)	-309 (12)	1 (3)	H(4')G3	697 (6)	237 (4)	223 (11)	12 (6)
H(5')C1	111 (4)	87 (3)	-209 (14)	4 (4)	H(5')G3	661 (8)	181 (5)	179 (25)	16 (7)
H(3')C1	143 (6)	36 (4)	-497 (18)	9 (5)	H(5')G3	694 (5)	181 (4)	292 (18)	7 (5)
H(1')C1	358 (6)	86 (4)	-448 (21)	12 (6)	H(3')G3	719 (4)	215 (3)	-28 (12)	1 (3)
H(1')G1	315 (5)	19 (4)	169 (17)	9 (4)	H(N2)G4	61 (6)	334 (4)	219 (19)	11 (5)
H(2')G1	270 (4)	10 (2)	-62 (11)	1 (3)	H(N2)G4	-29 (4)	325 (3)	286 (14)	4 (4)
H(4')G1	204 (4)	-16 (3)	250 (14)	4 (4)	H(N1)G4	-26 (7)	227 (5)	272 (22)	13 (6)
H(5')G1	179 (5)	-84 (3)	289 (16)	7 (4)	H(C8)G4	184 (6)	169 (4)	-158 (18)	8 (5)
H(3')G1	194 (5)	-39 (4)	-15 (17)	7 (5)	H(N2)C4	131 (6)	151 (4)	385 (18)	8 (5)
H(N2)G2	515 (7)	92 (5)	164 (21)	14 (6)	H(N2)C4	87 (5)	168 (3)	554 (16)	5 (4)
H(N2)G2	475 (6)	39 (5)	382 (20)	11 (6)	H(C5)C4	207 (6)	195 (4)	347 (18)	8 (5)
H(N1)G2	449 (5)	-3 (4)	305 (16)	8 (5)	H(C6)C4	264 (5)	230 (4)	317 (17)	8 (5)
H(C8)G2	646 (6)	-89 (4)	-86 (17)	9 (5)	H(5')C4	323 (5)	343 (4)	92 (18)	9 (5)
H(N2)C2	608 (5)	-112 (4)	430 (16)	6 (4)	H(3')C5	326 (4)	276 (2)	464 (12)	1 (3)
H(N2)C2	568 (6)	-88 (4)	544 (18)	9 (5)	H(2')C4	239 (6)	325 (4)	614 (20)	11 (5)
H(C5)C2	714 (5)	-69 (4)	297 (16)	6 (4)	H(1')G4	149 (6)	268 (4)	-196 (19)	9 (5)
H(C6)C2	731 (4)	2 (3)	325 (13)	3 (3)	H(2')G4	218 (4)	272 (2)	95 (11)	1 (3)
H(5')C2	866 (4)	80 (2)	243 (12)	1 (3)	H(5')G4	306 (4)	171 (2)	-358 (11)	1 (3)
H(5')C2	808 (4)	72 (3)	95 (15)	4 (4)	H(3')G4	278 (6)	198 (4)	-6 (19)	11 (6)
H(3')C2	823 (4)	47 (3)	509 (14)	4 (4)					

factors are also included with their estimated standard deviations.*

Bond distances and angles

To discuss the enormous amount of data would involve investigating every bond length and angle in each GpC for completeness. In view of the 1.1 Å resolution and the fact that such a lengthy analysis would add little to the discussion of the results, a more compact synopsis of these results is given in Figs. 6 and 7. In the figures each bond length and bond angle is shown as the weighted average of the four observations for that parameter, and the number in parentheses is the standard deviation of the mean. These results can then be compared to more accurate determinations made on smaller structures. Within the standard deviations (about 0.02–0.04 Å and 1–2° for bond lengths and angles respectively) the great majority of these observations fall within standard values such as those cited by Arnott, Dover & Wonacott (1969). Deviations from these standard values are probably not significant considering the resolution of the data.

* A list of structure factors and a table of all the bond lengths and angles for the four GpC have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 31869 (44 pp., 1 microfiche). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

Geometry of the phosphate group

The geometry of the phosphate group (Fig. 7) is important because it makes the greatest contribution of scattering matter (37% for an average nucleotide), and thus a knowledge of its geometry is important in deciphering diffraction data for tRNA or other larger nucleic acids. Fortunately, this unit exhibits the least variability among the four observed GpC. The four O atoms are arranged approximately tetrahedrally

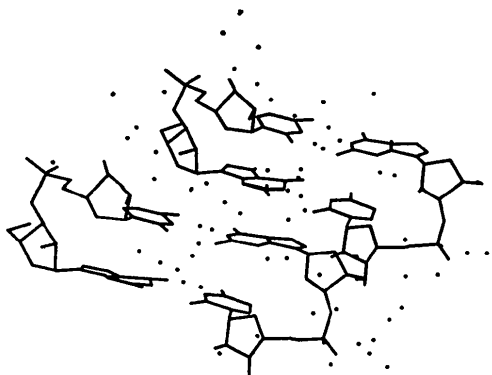


Fig. 5. View down two adjacent unit cells so that the stacking of the six-membered aromatic rings is illustrated. The dots in the figure are the water positions.

about the P atom. The two in-chain P–O distances are 1.59, as observed in nucleotides, while the other P–O distances are 1.47 Å, indicating that the phosphate group carries a charge of -1 distributed equally between these two shortened bonds. The bond angles about O(3') and O(5') are both 120° indicating the same sp^2 hybridization reported by Day *et al.* (1973) for NaGpC.

Hydrogen bonds (non-water)

Complementary base-pairing (Watson–Crick): Each GpC participates in six Watson–Crick hydrogen bonds (Table 11). These bonds are all quite reasonable and vary from 2.73 to 2.98 Å, as could have been expected from previous studies.

Other intermolecular hydrogen bonds: Each GpC exhibits two close contacts that are probably hydrogen bonds. These are O(1') (C) \cdots O(5') (G) and N(4) (C) \cdots O(3') (C). The GpC that participate in these bonds are given in Table 12. The distances range from 3.00

Table 11. Watson–Crick hydrogen-bond distances (Å) for GpC dimers

Parentheses show hydrogen-bonded atom of dimeric partner.

Molecule	O(6) [–N(4)]	N(1) [–N(3)]	N(2) [–O(2)]
(1)	2.83	2.94	2.78
(2)	2.90	2.92	2.85
(3)	2.98	2.92	2.86
(4)	2.79	2.95	2.73

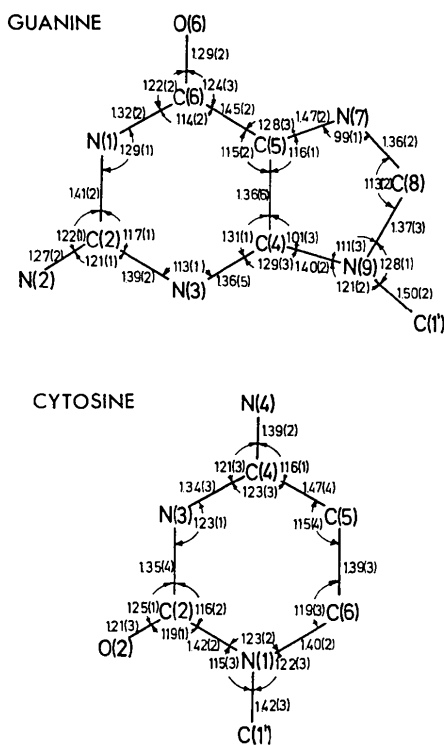


Fig. 6. Guanine and cytosine base plane bonds and angles (weighted mean values).

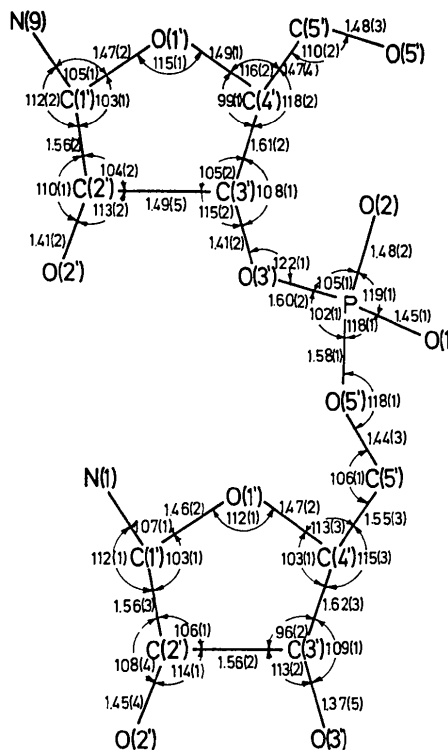


Fig. 7. Phosphate ribose backbone bonds and angles (weighted mean values) [O(5')–P–O(2)P, $107(1)^\circ$; O(2)P–P–O(3'), $111(1)^\circ$ not shown].

to 3.23 Å and 2.82 to 3.00 Å for O(1')(C) \cdots O(5')(G) and N(4)(C) to O(3')(C) respectively. These are reasonable values and are consistent with the results for NaGpC.

Table 12. Additional close interatomic contacts between atoms of different GpC

Atom	Molecule	Atom	Molecule	Distance (Å)
O(1')	C1	O(5')	G4	3.00
O(1')	C2	O(5')	G3	3.06
O(1')	C3	O(5')	G1	3.16
O(1')	C4	O(5')	G2	3.23
N(4)	C1	O(3')	C3	3.00
N(4)	C2	O(3')	C4	3.00
N(4)	C3	O(3')	C2	2.82
N(4)	C4	O(3')	C1	2.88

Intramolecular hydrogen bonds: The C(8) of guanine is in close contact with both O(1') and O(5') of the adjoining sugar. The distances range from 2.95 to 3.39 Å and from 2.54 to 2.84 Å for C(8) \cdots O(5') and C(8) \cdots O(1') respectively, indicating possible hydrogen bonds (Watson & Crick, 1953). For O(1') the angle C(8)–H \cdots O(1') may be too short to be a hydrogen bond.

Base-plane geometry

Least-squares planes were calculated for the eight bases of the asymmetric unit. The equations of the

planes and the atomic deviations from those planes are given in Table 13. The planes were calculated with only the six in-ring atoms for cytosine or nine in-ring atoms for guanine. The largest deviation of any atom from the plane was 0.13 Å. The angles between these planes are given in Table 7 and indicate a smaller angle (6.9° average) than the 11.6° result observed for NaGpC.

Table 13. *Least-squares planes of cytosine and guanine bases in GpC*

(a) Guanine [least-squares planes do not include N(2), O(6) or C(1')]

Equations

$$\begin{aligned} (1) & 0.69191x - 0.02318y + 0.72161z - 5.29703 = 0 \\ (2) & 0.68446x - 0.04135y + 0.72788z - 8.62029 = 0 \\ (3) & 0.67703x - 0.09687y + 0.72955z - 11.79849 = 0 \\ (4) & 0.66180x - 0.09490y + 0.74365z - 0.89278 = 0 \end{aligned}$$

Deviations from planes in Å

	(1)	(2)	(3)	(4)
N(9)	+0.002	+0.006	-0.007	-0.027
C(8)	-0.028	-0.011	+0.001	+0.059
N(7)	-0.029	-0.012	-0.031	+0.041
C(5)	+0.052	-0.024	+0.077	-0.104
C(6)	+0.008	+0.055	-0.006	-0.031
N(1)	-0.025	-0.034	-0.037	+0.054
C(2)	-0.015	-0.013	+0.032	+0.026
N(2)	-0.108	-0.016	-0.051	+0.005
C(4)	+0.032	+0.030	-0.005	-0.018
O(6)	-0.003	+0.119	-0.029	-0.042
N(3)	-0.003	+0.003	-0.023	+0.001
C(1')	+0.047	-0.083	+0.114	-0.177

(b) Cytosine [least-squares planes do not include N(4), O(2) or C(1')]

Equations

$$\begin{aligned} (1) & 0.57219x - 0.06605y + 0.81746z - 0.42550 = 0 \\ (2) & 0.59137x - 0.02195y + 0.80610z - 11.27240 = 0 \\ (3) & 0.60476x - 0.03271y + 0.79574z - 7.14308 = 0 \\ (4) & 0.57380x - 0.13771y + 0.80733z - 4.06293 = 0 \end{aligned}$$

Deviations from planes in Å

	(1)	(2)	(3)	(4)
N(4)	-0.079	+0.014	-0.025	+0.060
C(4)	+0.002	+0.045	-0.018	-0.004
C(5)	+0.016	-0.066	+0.025	-0.042
C(6)	-0.045	+0.060	-0.021	+0.056
N(1)	+0.053	-0.026	+0.007	-0.028
N(3)	+0.006	-0.013	+0.006	+0.035
C(2)	-0.032	-0.001	+0.001	-0.017
O(2)	-0.118	+0.047	-0.023	+0.132
C(1')	-0.034	-0.007	-0.069	-0.034

Solvent and cation effects

Water hydrogen bonds: With 36 water molecules per asymmetric unit, a very complex solvent hydrogen-bonding scheme was developed. The GpC are surrounded on all sides by water molecules. Many waters act as bridges coordinating GpC through hydrogen-bond interactions. In addition, the waters fill up pockets of the structure, as also observed with NaGpC. In many cases bridges are accomplished through two or more waters which are not tightly bound in the structure, as evidenced by their high temperature factors and site disorder. A list of these bonds is given in Table 14.

Calcium coordination system and its effects on the structure: The two Ca^{2+} ions are an important stabilizing factor in the crystal. Two of the six octahedral sites in the coordination sphere of each Ca^{2+} are occupied by O atoms of phosphate groups from two GpC, each of which belongs to a different dimer, while the remaining four sites are occupied by water. Fig. 8 shows this symmetry relationship and how two of the waters in the Ca^{2+} sphere are also hydrogen-bonded to the guanine groups of GpC(2) and GpC(4) through N(7) and O(6). The details of the coordination geometry are given in Table 15. The Ca^{2+} is present on only one side of the pseudo twofold axis, which therefore reduces the symmetry from $C2$ to $P2_1$ (Fig. 3). The waters which are present near the vacant Ca^{2+} site rearrange themselves in order to fill the hole. This further breaks the symmetry and explains why several waters with only partial occupancy were observed in the average structure. As indicated earlier, the Ca^{2+} ions also perturbed the conformations of the four GpC indirectly. Fig. 9 shows the water and cation interactions for each GpC, indicating the similarity in the interactions for (1,4) and (2,3). This may be compared to NaGpC in which there is exact $C2$ symmetry, so that each GpC interacts with one Na^+ ion, or four GpC and four Na^+ ions per unit cell, while our asymmetric unit has four GpC and only two Ca^{2+} ions. Our unit cell is then completed by the 2_1 screw axis along b , which results in a doubling of the cell.

The presence of divalent ions has important implications for polynucleotide systems (Lindahl, Adams & Fresco, 1966; Adams, Lindahl & Fresco, 1967; Danchin, 1972). In this case, the Ca^{2+} ions, with near perfect octahedral coordination, act to hold the structure together by binding directly phosphate O atoms of two different GpC which are not Watson-Crick paired. Some of the coordination waters then bind to guanine residues through N(7) and O(6), but these are not from the same GpC as the phosphate O atoms. This is in contrast to the postulated model of Zn^{2+} interacting with ATP (Sundaralingam, 1969) where the ion coordinates with the phosphate and base of the same species. The Ca^{2+} ions bind both directly to the phos-

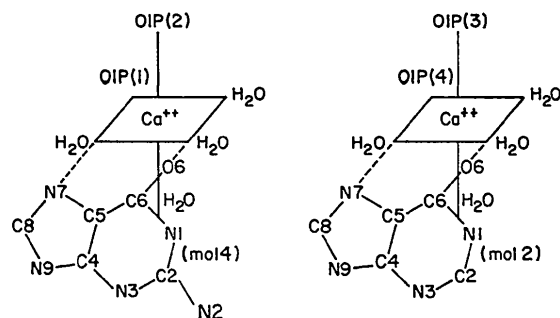


Fig. 8. Details of the geometry of the calcium coordination system showing the octahedral coordination bonding to the guanine rings of GpC(2) and (4).

phate and indirectly to the base moieties in another helical fragment. The only other Ca salt of a nucleotide, Ca²⁺ 5'TMP, (Trueblood, Horn & Luzzati, 1961) has sevenfold coordination with four phosphate O atoms from three different species plus three waters. Other divalent salts, Ba²⁺ 5'UMP (Shefter & Trueblood, 1965) and Ba²⁺ IMP (Nagashima & Iitaka, 1968), are less directly comparable, because of the size of the ion and coordination number. However, in Ba²⁺ 5'UMP, Ba²⁺ interacts with many water molecules (10 coordination) and the O(2') and O(3') of the ribose, but the interaction with the phosphate O atoms is through a water bridge rather than directly, as we observe here.

The Ca²⁺ and Na⁺ salts of GpC have significant differences, caused mostly by the different crystallographic environments. With only one Ca²⁺ per two GpC, compared with two Na⁺ in the structure solved by Day *et al.* (1973), the different coordination systems and unsymmetrical Ca²⁺ positions perturb the four GpC from exact symmetry. The GpC are therefore no longer identical and the result is four crystallographically independent observations of GpC in the asym-

Table 15. Calcium coordination geometry

Distances to 1Ca(1) (Å)		Distances to 4Ca(1) (Å)	
1W(3)	2.43	3W(3)	2.38
2W(4)	2.32	2W(4)	2.35
1W(1)	2.37	4W(1)	2.44
1W(2)	2.37	4W(2)	2.35
O1(P1)	2.40	O1(P3)	2.21
O1(P2)	2.20	O1(P4)	2.36
Angles about 1Ca(1) (°)		Angles about 4Ca(1) (°)	
1W(2)-1W(3)	87	4W(2)-3W(3)	88
-2W(4)	85	-3W(4)	86
-1W(1)	88	-4W(1)	94
-O1(P1)	96	-O1(P3)	175
-O1(P2)	178	-O1(P4)	99
1W(3)-2W(4)	80	3W(3)-3W(4)	80
-1W(1)	166	-4W(1)	167
-O1(P1)	86	-O1(P3)	95
-O1(P2)	94	-O1(P4)	88
2W(4)-1W(1)	87	3W(4)-4W(1)	87
-O1(P1)	166	-O1(P3)	90
-O1(P2)	94	-O1(P4)	168
1W(1)-O1(P1)	107	4W(1)-O1(P3)	83
-O1(P2)	90	-O1(P4)	104
O1(P1)-O1(P2)	86	O1(P3)-O1(P4)	86

Table 14. Water close interatomic contacts

	Partner	Distance (Å)		Partner	Distance (Å)		Partner	Distance (Å)		Partner	Distance (Å)
1W(1)	1Ca(1)	2.37	1W(4)	(O3')C1	2.80	3W(5)A	O2(P1)	2.85	4W(8)	(N2)G4	3.13
	1W(10)	2.70		(O6)G3	2.82		(O5')G1	2.89		4W(9)	2.86
	2W(10)	2.86		2W(1)	2.82		(O2')3C	2.40		4W(7)	2.78
2W(1)	O1(P1)	2.81		2W(4)	2.80	3W(5)B	1W(6)B	2.82		1W(6)A	2.82
	2W(2)	2.89	2W(4)	(O6)G4	2.69		O(2')3C	2.53		1W(6)B	2.87
	1W(4)	2.82		1Ca(1)	2.32		O2(P1)	2.54	1W(10)	O2(P1)	2.93
	2W(10)	2.84		1W(3)	3.06	4W(5)	O2(P2)	2.67		1W(1)	2.70
3W(1)	O1(P4)	2.82		1W(4)	2.80		(O2')4C	2.69		3W(7)	2.89
	3W(2)	2.81	3W(4)	(O6)G2	2.83		4W(9)	2.99	2W(10)	1W(1)	2.86
	4W(4)	2.69		4Ca(1)	2.35		4W(7)	2.78		2W(1)	2.84
	3W(10)	2.87		3W(3)	3.05	1W(7)	(O2')C1	2.91		4W(7)	3.00
4W(1)	O1(P3)	3.08		4W(4)	2.87		2W(7)	2.57	3W(10)	4W(1)	2.90
	4Ca(1)	2.44	4W(4)	(O6)G1	2.82		1W(8)	2.72		2W(7)	2.78
	3W(10)	2.90		(O3')C4	2.87		4W(10)	3.00		3W(1)	2.87
	4W(10)	2.67		3W(1)	2.69	2W(7)	1W(7)	2.57	4W(10)	O2(P4)	3.02
1W(2)	(O3')C1	2.78		3W(4)	2.87		3W(10)	2.78		4W(1)	2.67
	(O3')G1	2.95	1W(9)	(N3)G1	3.05	3W(7)	(O2')C3	3.04		1W(7)	3.00
	(O2')G1	2.91		(O2')G1	2.68		4W(7)	2.37	1W(6)A	O2(P2)	2.84
	1Ca(1)	2.37		1W(5)	2.92		3W(8)	2.87		3W(8)	2.82
2W(2)	(O3')C1	2.70		1W(8)	2.93		1W(10)	2.89		4W(8)	2.82
	(O3')G2	2.95		2W(8)B	3.08	4W(7)	4W(5)	2.78	1W(6)B	O2(P1)	2.60
	(O2')G2	2.95	2W(9)	(N3)G2	2.83		3W(7)	2.36		3W(5)B	2.82
	(N7)G3	2.82		(O2')G2	2.69		4W(8)	2.78		3W(8)	2.58
	2W(1)	2.89		(O5')G3	2.92	1W(8)	(N2)G1	2.95		4W(8)	2.87
3W(2)	(N7)G1	2.76		3W(6)B	2.69		1W(9)	2.93		3W(9)	3.06
	(O3')C3	2.66	3W(9)	(O5')G1	3.01		1W(7)	2.72	3W(6)A	O2(P3)	3.01
	(O3')G3	2.95		(N3)G3	2.94		2W(8)B	2.69		1W(8)	2.86
	(O2')G3	3.03		(O2')G3	2.85		3W(6)A	2.86		2W(8)A	2.95
	3W(1)	2.81		3W(6)B	2.75		3W(6)B	2.56		2W(8)B	2.18
4W(2)	(O3')C4	2.78		3W(8)	2.90	2W(8)A	3W(6)A	2.95	3W(6)B	1W(8)	2.56
	(O3')G4	3.00	4W(9)	(N3)G4	3.00		3W(6)B	2.98		2W(8)A	2.98
	(O2')G4	2.84		(O2')G4	2.89	2W(8)B	O2(P4)	2.73		2W(8)B	2.70
	4Ca(1)	2.35		4W(5)	2.99		1W(9)	3.08		O2(P3)	2.72
1W(3)	(N7)G3	3.03		4W(8)	2.86		1W(8)	2.69		2W(9)	2.69
	(N7)G4	2.78	1W(5)	(O2')1C	2.66	3W(8)	(N2)G3	2.83			
	1Ca(1)	2.43		O2(P4)	2.63		3W(9)	2.89			
	2W(4)	3.06		1W(9)	2.92		3W(7)	2.87			
3W(3)	(N7)G2	2.72	2W(5)	(O2')2C	2.70		1W(6)A	2.82			
	4Ca(1)	2.38		O2'(P3)	2.76		1W(6)B	2.58			
	3W(4)	3.05		(O5')G3	3.03						

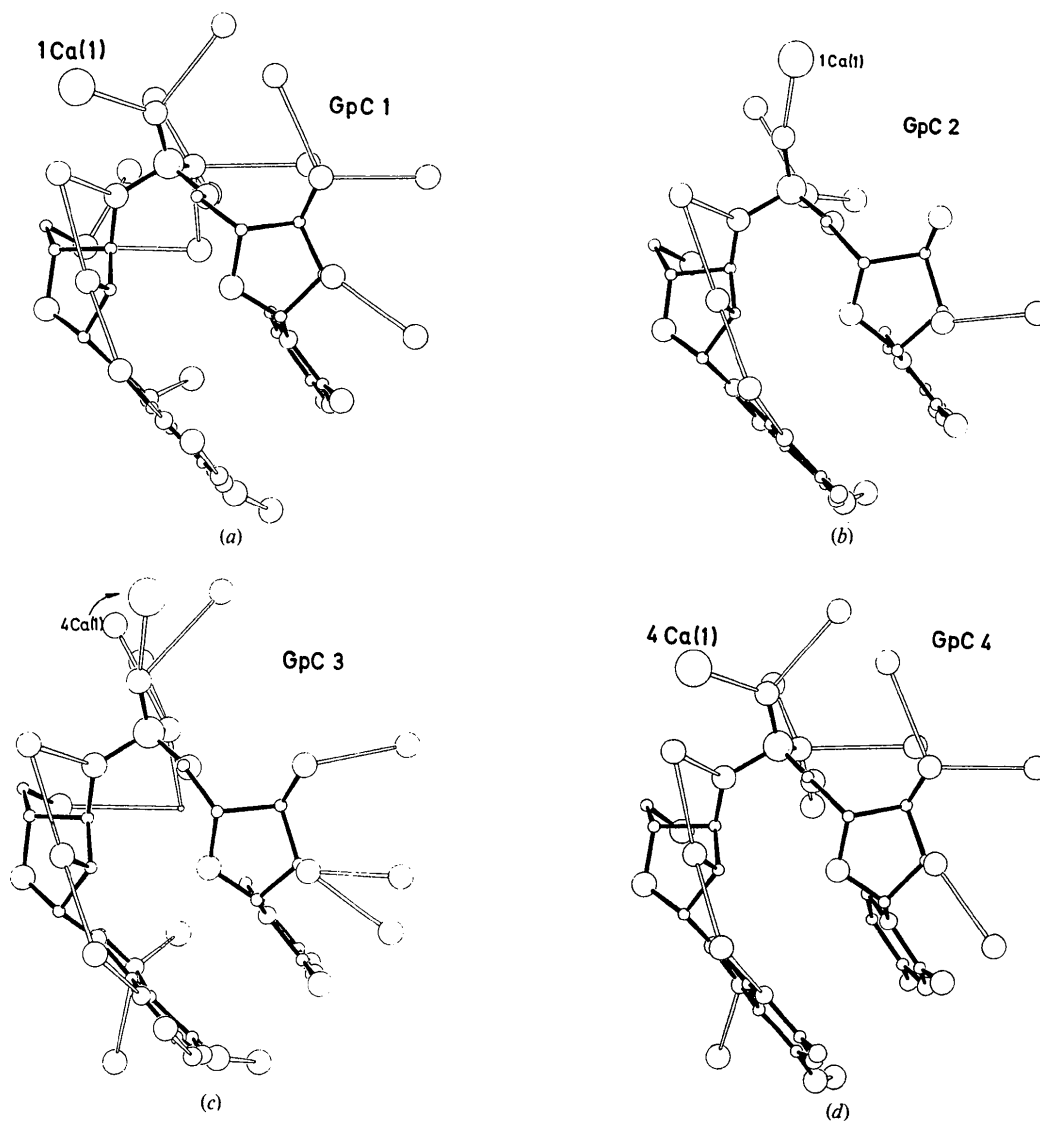


Fig. 9. Water and calcium interactions for (a) GpC(1) to (d) GpC(4) indicating similarity in (1,4) and (2,3) interactions.

metric unit. This shows the extent to which the backbone angles can be modified while still preserving the helical character of the structure.

The molecular model building was performed at the Princeton University Computer Graphics Laboratory, supported by NIH grant RR-00578; other support was received from NIH grant GM-16539 and NSF GB-28021. We thank Dr K. A. Kerr, Department of Chemistry and Physics, University of Calgary, Alberta, Canada for her helpful comments on this manuscript. Thanks also go to the Max Planck Institute for Experimental Medicine, Göttingen, BRD for computer time, Dr W. Saenger for helpful advice and Frau U. Bergmann for invaluable help with diagrams. Helpful comments by Dr A. Klug, MRC laboratory of Molecular Biology are also gratefully acknowledged. We thank Mr V.

Mallete, School of Physics, Georgia Institute of Technology for photographic work.

References

- ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* **A27**, 157–165.
 ADAMS, A., LINDAHL, T. & FRESCO, J. R. (1967). *Proc. Natl. Acad. Sci. U.S.* **57**, 1684–1691.
 ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Amer. Chem. Soc.* **94**, 8205–8212.
 ARNOTT, S., DOVER, S. D. & WONACOTT, A. J. (1969). *Acta Cryst.* **B25**, 2192–2206.
 DANCHIN, A. (1972). *Biopolymers*, **11**, 1313–1333.
 DAY, R. O., SEEMAN, N. C., ROSENBERG, J. M. & RICH, A. (1973). *Proc. Natl. Acad. Sci. U.S.* **70**, 849–853.
 DE CAMP, W. (1973). *Acta Cryst.* **A29**, 148–150.
 GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* **B26**, 274–285.

- HAMILTON, W. C. (1964). *Statistics in Physical Science*, pp. 157–162. New York: Ronald.
- HAUPTMAN, H., KARLE, I. L. & KARLE, J. (1960). *Acta Cryst.* **13**, 451–453.
- HINGERTY, B., SUBRAMANIAN, E., STELLMAN, S. D., BROYDE, S. B., SATO, T. & LANGRIDGE, R. (1975). *Biopolymers*, **16**, 227–236.
- LINDAHL, T., ADAMS, A. & FRESCO, J. R. (1966). *Proc. Natl. Acad. Sci. U.S.A.* **55**, 941–948.
- NAGASHIMA, N. & IITAKA, Y. (1968). *Acta Cryst.* **B24**, 1136–1138.
- POWELL, M. (1964). *Comput. J.* **7**, 155–162.
- ROGERS, D. (1965). *Computing Methods in Crystallography*. Edited by J. S. ROLLETT, pp. 117–132. New York: Pergamon Press.
- ROSENBERG, J. M., SEEMAN, N. C., DAY, R. O. & RICH, A. (1973). *Amer. Cryst. Assoc. Meeting*, Vol. 1, Ser. 2, 130.
- RUBIN, J., BRENNAN, T. & SUNDARALINGAM, M. (1972). *Biochemistry*, **11**, 3112–3128.
- SIEKER, L. C. & JENSEN, L. H. (1974). *Amer. Cryst. Assoc. Meeting*, Vol. 2, Ser. 2, 79.
- SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067–1077.
- STELLMAN, S. D., HINGERTY, B., BROYDE, S. B., SUBRAMANIAN, E., SATO, T. & LANGRIDGE, R. (1973). *Biopolymers*, **12**, 2731–2750.
- STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination: A Practical Guide*, p. 321. New York: Macmillan.
- SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.
- SUNDARALINGAM, M. (1973). *Jerusalem Symposia on Quantum Chemistry and Biochemistry*. Edited by E. BERGMANN & B. PULLMAN, Vol. V, pp. 417–455. New York: Academic Press.
- SUSSMAN, J. (1971). Ph.D. Dissertation, Massachusetts Institute of Technology.
- SUSSMAN, J., SEEMAN, N. C., KIM, S. H. & BERMAN, H. J. (1972). *J. Mol. Biol.* **66**, 403–421.
- TROTTER, J. (1965). Univ. British Columbia, Vancouver, Canada.
- TRUEBLOOD, K. N., HORN, P. & LUZZATI, V. (1961). *Acta Cryst.* **14**, 965–982.
- WATSON, J. D. & CRICK, F. H. C. (1953). *Nature, Lond.* **171**, 737–738.
- WUENSCH, B. J. & PREWITT, C. T. (1965). *Z. Kristallogr.* **122**, 24–59.

Acta Cryst. (1976). **B32**, 3013

Structural Studies of Polychlorinated Hydrocarbons. II.* Hexachloroxanthene and Hexachloroxanthone

BY M. SÖDERHOLM AND U. SONNERSTAM

University of Stockholm, S-104 05 Stockholm, Sweden

AND R. NORRESTAM† AND T.-B. PALM

Department of Structural Chemistry, Arrhenius Laboratory, University of Stockholm, S-104 05 Stockholm, Sweden

(Received 6 January 1976; accepted 28 April 1976)

The crystal structures of the two related compounds 1,3,4,6,7,9-hexachloroxanthene ($C_{13}H_4Cl_6O$) and 1,3,4,6,7,9-hexachloroxanthone ($C_{13}H_2Cl_6O_2$) have been determined by X-ray diffraction techniques. Both structures are monoclinic, space group $P2_1/n$ with $a = 16.114$ (4), $b = 11.950$ (2), $c = 7.254$ (1) Å, $\beta = 92.29$ (2)°, $Z = 4$, $D_x = 1.85$ g cm⁻³ and $a = 15.960$ (3), $b = 11.833$ (2), $c = 7.348$ (2) Å, $\beta = 93.11$ (2)°, $Z = 4$, $D_x = 1.93$ g cm⁻³, for the xanthene and xanthone derivatives respectively. Both structures were refined to R values of 0.06. The molecules of both compounds were found to be almost planar, with the largest deviation of the nonhydrogen atoms from planarity being shown by the pyranoid ring of the hexachloroxanthene molecule. The crystal structures are built up by successive layers (separation 3.4 Å) of nearly coplanar molecules, where the centroids of the molecules are approximately hexagonally distributed within one layer.

Introduction

The structure investigations of the two related hexachloro-substituted xanthene derivatives, 1,3,4,6,7,9-hexachloroxanthene and 1,3,4,6,7,9-hexachloroxanthone (cf. Fig. 1), reported here are part of a research project

concerning structural studies of polychlorinated hydrocarbons. Suitable specimens (Göthe & Wachtmeister, 1972; Eklund-Westlin, Göthe & Wachtmeister, 1976) were kindly supplied by Dr R. Göthe, University of Stockholm. The following experimental details apply to studies of both compounds. Single crystals suitable for X-ray studies were obtained by slow evaporation of ethanolic solutions. Preliminary examination of the crystals by photographic X-ray techniques revealed monoclinic symmetry with the systematic extinctions characteristic of the centrosymmetric space group

* Part I: Norrestam, Palm, Jensen & Sundström (1976).

† Present address: Chemistry Department B, The Technical University of Denmark, DTH 301, DK-2800 Lyngby, Denmark.