molecule reasonably well, and that on dimer formation there is little change in the molecule's folding and conformation.

Recently, large crystals have been obtained which diffract well to beyond 1·5 Å spacing. The refinement calculations are now being carried out at 1·7 Å spacing with newly collected data; these calculations will be further extended when the higher-resolution data become available.

References

ADAMS, M. J., BLUNDELL, T. L., DODSON, E. J., DODSON, G. G., VIJAYAN, M., BAKER, E. N., HARDING, M. M., HODGKIN, D. C., RIMMER, B. & SHEAT, S. (1969). *Nature (London)*, **224**, 491–495.

AGARWAL, R. C. (1978). *Acta Cryst.* A**34**, 791–800.

BENTLEY, G. A., DODSON, E. J., DODSON, G. G., HODGKIN, D. C. & MERCOLA, D. A. (1976). *Nature (London)*, **261**, 166–168.

BENTLEY, G. A., DODSON, E. J., DODSON, G. G. & LEVITOVA, A. (1978). *J. Mol. Biol.* **125**, 387–396.

BLUNDELL, T. L., DODSON, G. G., HODGKIN, D. C. & MERCOLA, D. A. (1972). *Adv. Protein Chem.* **26**, 279–402.

BRICOGNE, G. (1976). *Acta Cryst.* A**32**, 832–847.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–178. New York: Gordon and Breach.

CUTFIELD, J. F., CUTFIELD, S. M., DODSON, E. J., DODSON, G. G., EMDIN, S. O. & REYNOLDS, C. D. (1979). *J. Mol. Biol.* **132**, 85–100.

CUTFIELD, J. F., CUTFIELD, S. M., DODSON, E. J., DODSON, G. G., REYNOLDS, C. D. & VALLELY, D. (1981). *Structural Studies on Molecules of Biological Interest*, edited by G. G. DODSON, J. P. GLUSKER & D. SAYRE. Oxford Univ. Press.

CUTFIELD, J. F., CUTFIELD, S. M., DODSON, E. J., DODSON, G. G. & SABESAN, M. (1974). *J. Mol. Biol.* **87**, 23–30.

CUTFIELD, S. M. (1981). Private communication.

DODSON, E. J., DODSON, G. G., HODGKIN, D. C. & REYNOLDS C. D. (1979). *Can. J. Biochem.* **57**, 469–479.

DODSON, E. J., ISAACS, N. W. & ROLLETT, J. S. (1976). *Acta Cryst.* A**32**, 311–315.

GATTNER, H. G. (1975). *Hoppe–Seyler's Z. Physiol. Chem.* **386**, 1397–1404.

KIKUCHI, K., LARNER, J., FREER, R., DOY, A., MORRIS, H. & DELL, A. (1980). *J. Biol. Chem.* **255**, 9281–9288.

LI JIA-YAO, SONG SHIH-YAN, LI JUN, WANG JIA-HAI & YAU TSI-HO (1980). *Adv. Biochem. Biophys.* **4**, 43–45 (in Chinese).

MEYTS, P. D., OBBERGHEN, E. V., ROCH, J., WOLLMER, A. & BRANDENBURG, D. (1978). *Nature (London)*, **273**, 504–509.

NIXON, P. E. & NORTH, A. C. T. (1976). *Acta Cryst.* A**32**, 320–325.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 35–59.

PEKING INSULIN STRUCTURE RESEARCH GROUP (1974). *Sci. Sin.* **17**, 752–777.

PEKING INSULIN STRUCTURE RESEARCH GROUP (1976). *Sci. Sin.* **19**, 358–361.

POCKER, Y. & BISWAS, S. B. (1980). *Biochemistry*, **19**, 5043–5049.

REEVES, W. G. (1981). Private communication.

ROSSMANN, M. G. (1962). *Acta Cryst.* **15**, 24–31.

THE SHANGHAI INSULIN RESEARCH GROUP (1976a). *Sci. Sin.* **19**, 351–357.

THE SHANGHAI INSULIN RESEARCH GROUP (1976b). *Sci. Sin.* **19**, 475–485.

WANG JIA-HAI, LI JIA-YAO, YAU TSI-HO, KWEI LOO-LOO, LOU MAI-JAING, YU JUN-MENG & SONG SHIH-YAN (1980). *Sci. Bull.* **8**, 369–371 (in Chinese).

The Structure of the Ribodinucleoside Monophosphate Guanylyl-3',5'-cytidine as its Ammonium Octahydrate Salt

BY A. AGGARWAL, S. A. ISLAM, R. KURODA, M. R. SANDERSON* AND S. NEIDLE†

Cancer Research Campaign Biomolecular Structure Research Group, Department of Biophysics, King's College, 26–29 Drury Lane, London WC2B 5RL, England

AND H. M. BERMAN

Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Fox Chase, Philadelphia, Pennsylvania 19111, USA

Abstract

The crystal and molecular structure of an ammonium octahydrate salt of the ribodinucleoside monophosphate guanylyl-3',5'-cytidine ($C_{19}H_{24}N_8O_{12}P$) has been determined by X-ray methods, and refined to a final R of 0·111 for 1197 observed reflections. The salt crystallizes in space group $C2$ with cell dimensions $a = 20·987$ (6), $b = 16·470$ (4), $c = 9·566$ (1) Å and $\beta = 94·36$ (2)°; $Z = 4$ for a calculated density of 1·50 Mg m^{-3} and μ(Cu $K\alpha$) = 1·19 mm^{-1}. The unit-cell

* Present address: Gorlaeus Laboratories, The State University, Leiden, The Netherlands.

† To whom correspondence should be addressed.

dimensions are very similar to those of the sodium salt [Rosenberg, Seeman, Day & Rich (1976). *J. Mol. Biol.* **104**, 145–167]; however, the two structures are non-isomorphous. Conformational features of the ammonium salt are as expected for an RNA-like Watson–Crick base-paired duplex.

Introduction

Self-complementary phosphates may be considered to be the smallest fragments of double-helical nucleic acids that possess the conformational characteristics of polynucleotides themselves. The determinations of the crystal structures of guanylyl-3',5'-cytidine (GpC) as sodium (Rosenberg, Seeman, Day & Rich, 1976) and calcium salts (Hingerty, Subramanian, Stellman, Sato, Broyde & Langridge, 1976) showed the principal features of RNA-11 helical fragments – Watson–Crick base-pairing and backbone torsion angles that are close to those derived from X-ray fibre studies on double-stranded RNA (Arnott, Smith & Chandrasekaran, 1976).

This study is concerned with the crystal structure of a third salt of GpC, the ammonium one. In view of the important roles played by water molecules and counter ions in maintaining the integrity of nucleic acid and oligonucleotide structure, this work pays especial attention to these features.

Experimental

Crystals of ammonium GpC were grown from an aqueous solution containing free-base GpC, ammonium sulphate, and an intercalating drug, in the course of attempts to obtain a crystalline GpC–drug complex. The small prismatic crystals of the dinucleoside monophosphate were indicated to have monoclinic symmetry from preliminary X-ray photographs. They were sealed in quartz Lindemann capillary tubes with mother liquor so as to minimize loss of bound water. Initial measurement of cell dimensions indicated at least a close similarity to those reported for the sodium salt (Rosenberg *et al.*, 1976); the space group in both cases is $C2$. Accurate measurements of 25 θ values on an Enraf–Nonius CAD-4 diffractometer showed that, although very close to each other in size, the two cells are actually significantly distinct (Table 1). Intensity data were collected on the diffractometer with ω–2θ scans and graphite-monochromated Cu $K\alpha$ radiation

Table 1. *Cell dimensions for GpC salts*

	a (Å)	b (Å)	c (Å)	β (°)	V (Å^3)	Reference
Na$^+$	21·460	16·927	9·332	90·54	3389·7	Rosenberg *et al.* (1976)
NH$_4^+$	20·987 (6)	16·470 (4)	9·566 (1)	94·36 (2)	3297·0	This study
Ca^{2+}	21·224	34·207	9·372	90·53	6803·9	Hingerty *et al.* (1976)

$(1\cdot5 < \theta < 65\cdot0°)$, for a crystal of approximate dimensions $0\cdot08 \times 0\cdot05 \times 0\cdot01$ mm. A total of 3263 reflections were measured of which 3098 were unique. 1197 reflections had significant intensity with $I > 2\sigma(I)$. There was no crystal decomposition during the data collection.

Structure analysis

Since the cell dimensions for the ammonium salt are close to those for the sodium one, it was initially assumed that the two crystal structures were approximately isomorphous, with the asymmetric unit of GpC relating to another molecule by means of the crystallographic twofold axis to produce a double-stranded duplex. Accordingly, the Na$^+$GpC coordinates (Rosenberg *et al.*, 1976) were used as a starting-point for full-matrix least-squares refinement. This converged to an R of 0·45 with no apparent indication as to which regions of the structure were misplaced. It was thus assumed that the two crystal structures were non-identical, as far as this refinement procedure indicated. Nonetheless, marked differences in gross molecular position were ruled out at this stage. This conclusion was apparently reinforced by examination of a high-order $(\theta > 25°)$ Patterson map in a search for P–P interatomic vectors. A prominent vector in this map corresponded to a P position at $x = 0\cdot152$, $z = 0\cdot041$, compared to the P at $x = 0\cdot123$, $z = 0\cdot101$ in the Na$^+$GpC structure. However, attempts at translation of the structure to this position, followed by refinement, failed to produce acceptable F_o/F_c agreement. Numerous phasing trials starting with the P atoms were unsuccessful, as were attempts using an apparently correct phosphate group (obtained by superposition techniques).

At this point, reciprocal-space search methods (Tollin & Cochran, 1964) were used to find the orientation of the planar base pairs. An $I(\theta,\varphi)$ search gave a single strong maximum, which indicated that the normal to the bases lay in the ac plane at an angle of $\sim$−45° with respect to the a axis. This was confirmed using the Rosenberg *et al.* (1976) base-pair coordinates as a model for a rotation-function search in three dimensions; this gave two prominent peaks (A and B), both of which showed the normal to the base pairs at the orientation predicted by the $I(\theta,\varphi)$ function. Orientation A differed from the Rosenberg *et al.* (1976) orientation by a 90° rotation about the b axis, whilst orientation B differed by 180° about the a axis (Fig. 1). The net result was that both orientations gave the same P–P vector, but differed by a 180° rotation around an ac diagonal.

In order to find the translational component for each orientation, a computer program was written which calculates intermolecular distances between symmetry-related asymmetric units as a model is translated with respect to the symmetry elements. For orientation A,
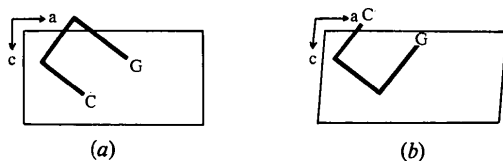
Fig. 1. (a) The orientation of a GpC molecule in the sodium salt (Rosenberg et al., 1976). (b) The final correct orientation found in this study for ammonium GpC.

three minima were found at which there were no intermolecular distances less than 2·7 Å, of which one gave good packing with the formation of GpC duplexes utilizing the twofold axes of the space group. However, a structure factor calculation produced a disappointingly high R factor. With orientation B, one of the two minima gave good packing. The transformation from the present structure to the sodium salt is:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0 & 0 \\ 0 & 0 & -1 \end{bmatrix} + \begin{bmatrix} -0·025 \\ 0·0 \\ -0·22 \end{bmatrix}.$$

This solution resulted in a low R of 0·402. Re-examination of the high-order Patterson map revealed a prominent peak corresponding to this position (which had previously been rejected). The model was gradually improved by a series of Fourier maps in which different parts of the structure were omitted in successive runs.

Refinement

At the point where R was 0·295, refinement was commenced. It was soon realized that unconstrained full-matrix refinement resulted in unacceptable geometries, presumably due to the paucity of observable reflections; accordingly the restrained/constrained refinement procedures of *SHELX* 76 (Sheldrick, 1976) were henceforth used. Idealized geometry used was based on that found for sodium GpC. Bases were almost fully constrained, whereas the backbone and sugar rings had relatively light restraints, with only valence-bonded distances kept close to their ideal values. In this way, the sugars and backbone atoms were left free to adopt their crystallographic conformations. After several cycles, R was reduced to 0·203. At this point, difference Fourier syntheses were examined in order to locate water molecules and the ammonium ion. Definitive assignment of the ammonium ion could not be made. It was thought most likely that this ion would be in close contact with the phosphate group. Accordingly, each 'water' molecule that was in such a position was tested in turn by least-squares refinement, to see which one gave improved agreement and a lower thermal parameter, when refined as an ammonium ion. Unfortunately, neither test differentiated between the possibilities, so the final assignment, made on distance criteria, is not a

unique one. A total of nine distinct water positions were found, of which two were situated on the twofold axis. Further refinement with the P atom assigned anisotropic thermal parameters, and with H atoms in calculated 'riding' positions, resulted in a final R of 0·111. Table 2 lists the final non-hydrogen position

Table 2. *Positional and thermal parameters and their e.s.d.'s*

	x	y	z	U_{iso} (Å^2)
C(6)C	0·317 (1)	−0·030 (2)	0·103 (3)	0·040 (9)
C(5)C	0·335 (1)	−0·105 (2)	0·070 (3)	0·045 (10)
C(4)C	0·388 (1)	−0·112 (2)	−0·020 (3)	0·050 (10)
N(4)C	0·406 (1)	−0·184 (2)	−0·058 (3)	0·060 (9)
N(3)C	0·415 (1)	−0·046 (2)	−0·065 (3)	0·043 (8)
C(2)C	0·395 (1)	0·029 (2)	−0·035 (3)	0·041 (9)
O(2)C	0·416 (1)	0·092 (1)	−0·089 (2)	0·047 (7)
N(1)C	0·346 (1)	0·037 (1)	0·054 (3)	0·048 (8)
C(1')C	0·322 (2)	0·121 (2)	0·078 (3)	0·056 (11)
C(2')C	0·281 (1)	0·153 (2)	−0·049 (3)	0·065 (12)
O(2')C	0·294 (1)	0·237 (2)	−0·054 (3)	0·093 (10)
C(3')C	0·215 (1)	0·136 (2)	0·006 (3)	0·041 (9)
O(3')C	0·166 (1)	0·185 (2)	−0·060 (2)	0·055 (7)
C(4')C	0·228 (1)	0·161 (2)	0·159 (3)	0·044 (9)
O(1')C	0·289 (1)	0·123 (2)	0·201 (2)	0·056 (7)
C(5')C	0·178 (2)	0·131 (2)	0·257 (3)	0·051 (10)
O(5')C	0·173 (1)	0·044 (1)	0·247 (2)	0·048 (6)
P	0·1238 (4)	0	0·3412 (9)	0·053 (6)*
O(1)P	0·121 (1)	−0·086 (1)	0·292 (3)	0·094 (10)
O(2)P	0·064 (1)	0·046 (2)	0·344 (3)	0·103 (10)
O(3')G	0·156 (1)	0·006 (2)	0·497 (2)	0·042 (6)
C(3')G	0·216 (1)	−0·031 (2)	0·535 (3)	0·044 (9)
C(4')G	0·221 (1)	−0·059 (2)	0·686 (3)	0·042 (9)
C(5')G	0·196 (2)	−0·141 (2)	0·717 (4)	0·08 (1)
O(5')G	0·210 (1)	−0·202 (2)	0·617 (3)	0·085 (9)
O(1')G	0·289 (1)	−0·060 (2)	0·726 (2)	0·041 (6)
C(1')G	0·324 (1)	−0·019 (2)	0·629 (3)	0·06 (1)
C(2')G	0·273 (1)	0·027 (2)	0·532 (3)	0·06 (1)
O(2')G	0·259 (1)	0·104 (1)	0·593 (2)	0·049 (6)
N(9)G	0·359 (1)	−0·079 (1)	0·552 (2)	0·034 (7)
C(8)G	0·359 (2)	−0·162 (2)	0·562 (4)	0·09 (2)
N(7)G	0·392 (1)	−0·199 (2)	0·467 (3)	0·062 (9)
C(5)G	0·422 (1)	−0·134 (1)	0·406 (3)	0·052 (10)
C(4)G	0·401 (1)	−0·060 (1)	0·452 (3)	0·06 (1)
N(3)G	0·420 (1)	0·015 (1)	0·417 (3)	0·049 (7)
C(2)G	0·463 (1)	0·014 (2)	0·317 (3)	0·049 (9)
N(2)G	0·482 (1)	0·086 (2)	0·274 (3)	0·066 (10)
N(1)G	0·485 (1)	−0·058 (2)	0·265 (2)	0·042 (8)
C(6)G	0·472 (1)	−0·135 (1)	0·307 (3)	0·06 (1)
O(6)G	0·492 (1)	−0·195 (2)	0·255 (3)	0·057 (7)
NH(4)	0·277 (2)	0·298 (3)	−0·301 (5)	0·13 (2)
OW(1)	0·133 (1)	0·157 (2)	0·677 (3)	0·080 (9)
OW(2)	0	0·162 (4)	0·500	0·14 (2)
OW(3)	0·371 (1)	0·176 (2)	0·518 (3)	0·10 (1)
OW(4)	0·042 (3)	−0·241 (5)	−0·057 (8)	0·30 (4)
OW(5)	−0·011 (3)	0·148 (5)	0·130 (7)	0·28 (3)
OW(6)	0·098 (3)	−0·089 (5)	−0·012 (8)	0·32 (4)
OW(7)	0·500	0·38 (1)	0·500	0·50 (9)
OW(8)	0·456 (5)	0·271 (7)	−0·35 (1)	0·43 (5)
OW(9)	0·063 (4)	0·018 (6)	0·846 (9)	0·40 (5)

* For P, $U_{eq} = \frac{1}{3} \sum_i \sum_j U_{ij} a_i^* a_j^* (\mathbf{a}_i . \mathbf{a}_j)$ and the anisotropic temperature factors are ($\times 10^3$):

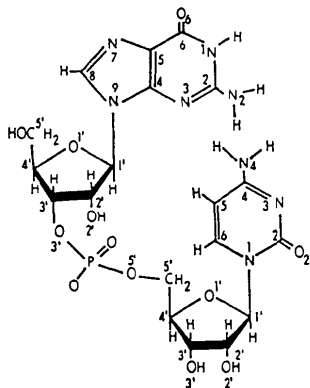| U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
27 (4)	80 (8)	55 (6)	3 (1)	27 (4)	−4 (6)

Fig. 2. Structure and atomic numbering for GpC.

parameters.* The atomic numbering scheme used is indicated in Fig. 2.

Most calculations were performed on PDP11/34 and 11/70 computers; refinements using *SHELX* were carried out on the University of London CDC7600 computer.

Discussion

Molecular structure

In view of the constrained/restrained refinement used, the values of individual bond distances and angles for ammonium GpC cannot be meaningfully analysed, and so will not be further discussed here. E.s.d.'s for all distances given below average 0·04 Å.

Fig. 3 shows two molecules of GpC related by the crystallographic twofold axis. The three Watson–Crick hydrogen bonds have similar distances to those found in the other salts (Table 3), confirming the broad trends noted in these earlier studies, with $N(3)C\cdots N(1)G$ and $N(4)C\cdots O(6)G$ having the longest, and $O(2)C\cdots N(2)G$ the shortest hydrogen bond. The base pairs themselves are planar within experimental error, having a twist [as defined by Berman & Shieh (1981)] of 0·6°. The sodium salt has a 3·2° twist to the base pairs, and those in the calcium salt (with bases of the two complete GpC duplexes per asymmetric unit) range from −1·2 to 6·1°. The individual bases are bent with respect to each other at 7·1°, in close agreement with the bending found in the other salts. Likewise, the dihedral angle between the planes of the guanine and cytosine bases on the same strand (7·1°) is very similar to the earlier values. The base turn angle† of 27·5° is

---

* Lists of structure factors and H-atom positions have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 38099 (7 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

† Defined as the angle subtended by the vectors connecting $C(1')$ atoms of each base pair, when projected on the average base plane viewed from a point perpendicular to this plane (Berman, Neidle & Stodola, 1978).

only slightly lower than the 30° value for the sodium salt. However, it is known that small differences such as this (as well as differences in other conformational variables) can result in considerable differences in helical parameters for RNA double helices derived from quasi-helical fragments such as GpC (Berman & Shieh, 1981). We are currently further investigating this topic (Islam, Neidle & Subbiah, unpublished studies).

The backbone and glycosidic torsion angles are given in Table 4, which shows that the close resemblance to the conformation of A-RNA (Arnott *et al.*, 1976) is maintained and, therefore, that the backbone is helical-like. The low value of the 5'-end glycosidic angle compared to that at the 3' end is in accord with the findings for the other GpC salts (Table 4). The two ribose sugars are broadly similar, both having $C(3')$-*endo* puckers (Table 5). However, in detail they are distinct; the 3' ribose has close to pure $C(3')$-*endo* character, with its pseudorotation parameter P close to the ideal value of 18° (Altona & Sundaralingam, 1972). The 5' ribose, on the other hand, is best described as $C(3')$-*endo*—$C(2')$-*exo* (which has a P of 0°), with atoms $C(3')$ and $C(2')$ both deviating from the least-squares plane of $C(4')$—$O(1')$—$C(1')$ (Table 5b). This pattern, with slightly distinct ring puckers, has been observed in both the sodium and the calcium salts of GpC and in sodium ApU (Seeman, Rosenberg, Suddath, Kim & Rich, 1976); as in the present structure, the 3' end displays purer $C(3')$-*endo* pucker, and this thus appears to be a general property of ribodinucleosides.
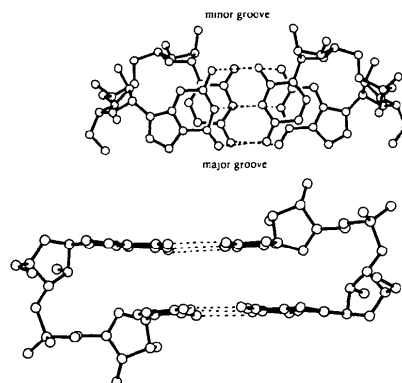

Fig. 3. Two views of a GpC duplex, with the two molecules in each duplex connected by hydrogen bonds.

Table 3. *Watson–Crick hydrogen-bond distances* (Å) *in GpC salts* (*average e.s.d.* 0·04 Å)

	$N(2)C$–$O(2)G$	$N(3)C$–$N(1)G$	$N(4)C$–$O(6)G$
Na⁺	2·86	2·95	2·91
NH₄⁺	2·88	2·95	2·96
Ca²⁺	2·81	2·93	2·87

Values for the Ca²⁺ salt are the averages from the four independent molecules in the structure.

Table 4. *Backbone torsion angles* (°) *for GpC salts, and for double-helical RNA* (*average e.s.d. for the* NH_4^+ *salt is* 3°)

	$\chi(5')$	C(4')–C(3')	C(3')–O(3')	O(3')–P	P–O(5')	O(5')–C(5')	C(5')–C(4')	C(4')–C(3')	$\chi(3')$	Reference
Na^+	13	89	211	292	285	184	50	77	32	Rosenberg *et al.* (1976)
NH_4^+	1	86	212	298	288	180	57	76	27	This work
Ca^{2+}	5	79	220	291	287	175	55	80	27	Hingerty *et al.* (1976)
A-RNA	14	95	202	294	294	186	49	95	14	Arnott *et al.* (1976)

Values for the calcium salt are the average for the four independent GpC molecules in the asymmetric unit.

Table 5. *Ribose-ring conformations*

(a) Ring torsion angles (°) (average e.s.d. 3°)

	τ_0	τ_1	τ_2	τ_3	τ_4	P	τ_m
Guanosine	11	−31	39	−34	14	3	39
Cytidine	−4	−24	40	−45	30	24	44

The pseudorotation parameter P is defined as:

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2\tau_2 (\sin 36° + \sin 72°)},$$

and the maximum degree of pucker τ_m as: $\tau_m = \tau_2/\cos P$.

(b) Deviations from planes (Å) (e.s.d.'s 0·04 Å)

	Guanosine	Cytidine
*C(1')	0·00	0·00
*O(1')	0·00	0·00
*C(4')	0·00	0·00
C(3')	−0·36	0·74
C(2')	−0·29	0·09

* These atoms constitute the plane.

Although the self-complementary ribodinucleoside phosphate structures of GpC and ApU salts show relative conformational uniformity, the backbone angles O(5')–C(5') and C(5')–C(4') both vary over a 20° range. These variations are in fact related, as shown in Fig. 4, in an inverse manner.
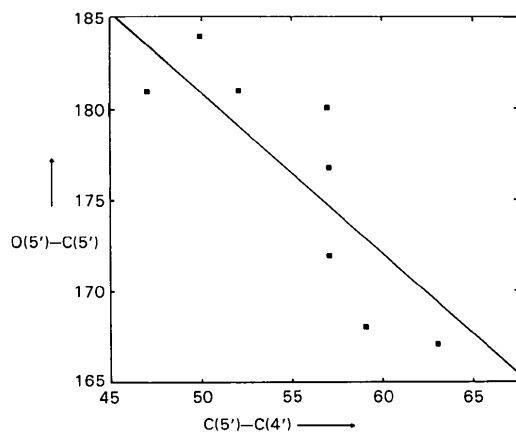

Fig. 4. Correlation plot of torsion angles (°), using data from sodium, ammonium and calcium GpC, and from sodium ApU.

Crystal structure

The GpC molecules in all three salts are heavily hydrated, with eight waters per GpC molecule in the ammonium salt, and nine for each of the other two. All the water molecules are involved in first-shell hydration. The slight increase in hydration for the present structure compared to these others is reflected in its ~3% smaller volume, and may well be a consequence of the likely difference in the hydration requirements possessed by the ammonium ion in this structure, compared to sodium and calcium ions in their GpC salts. We emphasize at this point, in view of the inconclusive ammonium ion assignment, that the following discussion is not to be taken as definite.

Examination of Table 6 shows that the 'ammonium' ion is involved in just three hydrogen bonds, none of which directly concern water molecules. A phosphate O atom O(1)P and O(5') of the guanosine ribose are bridged by the 'ammonium' ion (Fig. 5), which also links to the 2' hydroxyl group of a symmetry-related cytosine ribose group. This is in marked contrast with the role of the sodium ion in the sodium salt (Rosenberg *et al.*, 1976), which is coordinated only to water molecules and the phosphate group, and even more with the role of the calcium ion in its GpC salt (Hingerty *et al.*, 1976), which bridges phosphate groups from two GpC molecules and hydrogen-bonds to two guanine bases, as well as coordinating water molecules.

Two water molecules are involved in hydrogen-bonded bridges, as well as the 'ammonium' ion detailed above. OW(3) bridges N(3) of a guanine base and O(2') of its ribose, which itself bridges O(3') via OW(1) (Fig. 5). An identical arrangement is seen in both the sodium and the calcium salt crystal structures, and O(2')···O(3') water bridges are also found in the sodium ApU structure (Seeman *et al.*, 1976).

The various sites on the base pairs that are capable of participating in hydrogen bonds besides the Watson–Crick ones, do so in a manner similar to that in sodium GpC. Thus, N(3), N(7) and O(6) of the guanine hydrogen-bond to water atoms OW(3), OW(1), OW(2) and OW(5) (Table 6). The cytosine base is hydrated *via* its O(2) atom; N(4) is involved in a relatively strong interaction (2·91 Å) with O(3')C of

Table 6. *Intermolecular hydrogen-bond distances* (Å) (*average e.s.d.* 0·04 Å)

(a) Involving water molecules, NH_4^+ ions and GpC molecules

N(3)G···OW(3)	3·03	4 0 0 1	O(3')G···OW(1)	3·08	1 0 0 1	O(2)C···OW(4)	3·19	4 0 0 0
O(6)G···OW(5)	2·85	4 0 – 1 0	O(2)P···OW(2)	2·82	1 0 0 0	OW(1)···OW(2)	3·14	1 0 0 – 1
N(7)G···OW(1)	2·77	4 0 – 1 0	···OW(2)	2·82	2 0 0 1	···OW(2)	3·14	2 0 0 1
···OW(2)	3·22	3 0 – 1 0	···OW(5)	2·99	2 0 0 0	OW(3)···OW(8)	2·60	1 0 0 0
···OW(2)	3·22	4 0 – 1 1	···OW(9)	3·13	2 0 0 0	OW(4)···OW(6)	2·79	1 0 0 0
O(2')G···OW(1)	2·96	1 0 0 1	O(1)P···NH(4)	2·87	1 0 0 0	OW(5)···OW(5)	2·57	2 0 0 0
···OW(3)	2·78	4 0 0 1	···OW(6)	2·91	1 0 0 0	···OW(9)	2·42	1 0 0 0
O(5')G···NH(4)	3·06	1 0 0 0	···OW(8)	2·94	1 0 0 0	OW(7)···OW(8)	2·49	1 0 0 0
···OW(3)	2·87	1 0 0 0	O(2')C···NH(4)	2·57	4 0 0 0	···OW(8)	2·49	2 0 0 1

(b) Involving only GpC molecules

O(2)C···N(2)G	2·88	2 1 0 0	N(4)C···O(3')C	2·91	4 0 – 1 0	C(4)C···O(1')G	3·19	1 0 0 – 1
N(3)C···N(1)G	2·95	2 1 0 0	···O(6)G	2·96	2 1 0 0			

Equivalent positions are denoted (1 → 4); cell translations are 0, ±1 in the **a**, **b** and **c** directions

1 0 0 0	$x,$	$y,$	z	3 0 0 0	$\frac{1}{2}+x,$	$\frac{1}{2}+y,$	z
2 0 0 0	$-x,$	$y,$	$-z$				

(at right) 4 0 0 0 $\frac{1}{2}-x,$ $\frac{1}{2}+y,$ $-z$
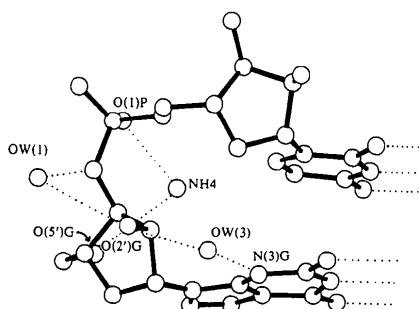
Fig. 5. A GpC molecule, with bridging 'ammonium' ion and water molecules.

Fig. 6. View of the crystal structure. Water molecules are shown as filled circles.

another GpC molecule neighbouring the major-groove environment of N(4)C. The major groove is defined as shown in Fig. 3; it is the side from which the glycosidic bonds are directed towards the sugars and away from the bases. The minor groove is directed to the reverse direction.

The major-groove hydration is relatively more extensive than the minor one, with O(6)G and N(7)G participating in hydrogen bonds to four water molecules and N(4)C linking to a symmetry-related O(3')C (see above). By contrast, N(3)G and O(2)C hydrogen-bond to just two water molecules, with N(2)G not being hydrated. A channel of water molecules primarily involving OW(1), OW(2) and OW(5) runs parallel to the c axis (Fig. 6). This channel also provides linkages to phosphate groups of symmetry-related duplexes. Therefore, these channels also serve to link separate columns of GpC duplexes; by contrast, the sodium GpC structure has the sodium ions dominating these gaps. This pattern of the major-groove hydration network being more extensive than the minor-groove one is typical of A-type double helices. It has been observed in a more marked form in the deoxy-tetranucleotide d(CpCpGpG), which has an A-DNA structure (Conner, Takamo, Tanaka, Itakura & Dic-
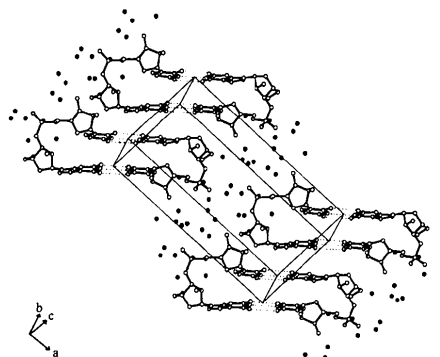
kerson, 1982), as well as in the drug-intercalated deoxydinucleoside d(CpG) (Neidle, Berman & Shieh, 1980).

Conclusions

This study has shown that substituting one counter ion for another may produce significant changes in the crystal structures and hence the patterns of water coordination for nucleic acid fragments. Furthermore, apparent equivalence of unit-cell parameters for these structures may not necessarily be a reliable indicator of isomorphism in packing arrangements for oligonucleotides.

References

ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

ARNOTT, S., SMITH, P. J. C. & CHANDRASEKARAN, R. (1976). *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. 2, Sect. *B*, edited by G. D. FASMAN, pp. 411–422. Cleveland, Ohio: Chemical Rubber Co.

BERMAN, H. M., NEIDLE, S. & STODOLA, R. K. (1978). *Proc. Natl Acad. Sci. USA*, **75**, 828–832.

BERMAN, H. M. & SHIEH, H.-S. (1981). *Topics in Nucleic Acid Structure*, edited by S. NEIDLE, pp. 17–32. London: MacMillan.

CONNER, B. N., TAKAMO, T., TANAKA, S., ITAKURA, K. & DICKERSON, R. E. (1982). *Nature (London)*, **295**, 294–299.

HINGERTY, B., SUBRAMANIAN, E., STELLMAN, S. D., SATO, T., BROYDE, S. B. & LANGRIDGE, R. (1976). *Acta Cryst.* B**32**, 2998–3013.

NEIDLE, S., BERMAN, H. M. & SHIEH, H.-S. (1980). *Nature (London)*, **288**, 129–133.

ROSENBERG, J. M., SEEMAN, N. C., DAY, R. O. & RICH, A. (1976). *J. Mol. Biol.* **104**, 145–167.

SEEMAN, N. C., ROSENBERG, J. M., SUDDATH, F. L., KIM, J. J. P. & RICH, A. (1976). *J. Mol. Biol.* **104**, 109–144.

SHELDRICK, G. M. (1976). *SHELX* 76. A system of computer programs for X-ray crystallography. Univ. of Cambridge, England.

TOLLIN, P. & COCHRAN, W. C. (1964). *Acta Cryst.* **17**, 1322–1324.

Theoretical Aspects of the OH···O Hydrogen Bond and Its Role in Structural and Kinetic Phenomena

BY M. D. NEWTON

Department of Chemistry, Brookhaven National Laboratory, Upton, New York 11973, *USA*

Abstract

Ab initio molecular-orbital theory, in conjunction with a flexible atomic-orbital basis set (4-31G), has given a good quantitative account of phenomena associated with the OH···O bond, using methanediol and one or two hydrogen-bonded water molecules as model systems. The calculations and the available crystal structure data both lead to the following conclusions: (1) OH···O bonds involving an anomeric OH donor group are shorter and stronger than those associated with non-anomeric O, and OH groups (whether anomeric or non-anomeric) which are also proton acceptors serve as more effective proton donors; in either case, the H···O bond-length shortening falls in the range 0.07 ± 0.02 Å; (2) although the anomeric OH group can serve as a proton acceptor, its affinity for this role is less than for a non-anomeric O, as inferred by consideration of bond energy and the length of the acceptor H···O bond; (3) the difference in lengths of the two CO bonds in the hemiacetal group bearing the anomeric O atom (O_a) is sensitive to the hydrogen-bonding environment, with variations of up to 0.04 Å being observed; cooperative coupling of proton donor and acceptor functions at O_a occurs at the expense of disproportionation in CO bond lengths; (4) studies of small clusters of water molecules indicate that extended sequential OH···O bonding, with each monomer donating and accepting one proton, yields cooperative shortening of O···O distances and lengthening of the short OH bonds (those involved as proton donors) which can be used to rationalize the limiting magnitudes observed in ice I. In contrast, the branched pentamer has the same O···O separation as the dimer because of the unfavorable double-donor and double-acceptor roles played by the central water molecule.

Introduction

Recent neutron diffraction studies of carbohydrate molecules (Jeffrey & Takagi, 1978; Jeffrey & Lewis, 1978; Ceccarelli, Jeffrey & Taylor, 1981; Jeffrey & Małuszyńska, 1981; Jeffrey, 1982; Jeffrey & Mitra, 1983; Saenger, 1979*a,b*; Lindner & Saenger, 1982) have yielded a body of precise crystal structure data pertaining to OH···O hydrogen bonds and have given new impetus to theoretical studies (Newton & Jeffrey, 1977; Tse & Newton, 1977; Newton, Jeffrey & Takagi, 1979). Much of the theoretical work on hydrogen bonding over the past decade has focused on the equilibrium characteristics of prototype systems, such as the dimers obtainable from the simple first-row (NH_3, H_2O, HF) and second-row (PH_3, H_2S, HCl) hydrides (Hankins, Moskowitz & Stillinger, 1970; Del Bene & Pople, 1970; Allen, 1975; Dill, Allen, Topp &