

The Crystal and Molecular Structure of *N*-(2-Chloroethyl)-D-gluconamide

BY L. O. G. SATZKE AND M. F. MACKAY

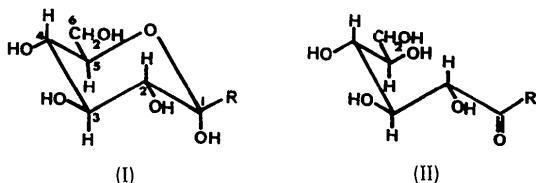
Department of Physical Chemistry, La Trobe University, Bundoora, Victoria 3083, Australia

(Received 2 December 1974; accepted 7 December 1974)

Triclinic crystals of *N*-(2-chloroethyl)-D-gluconamide, $C_8H_{16}ClNO_6$, belong to the space group $P1$, with $a=4.807$, $b=5.093$, $c=11.888$ Å, $\alpha=91.25$, $\beta=84.48$, $\gamma=94.48^\circ$ and $Z=1$. The structure was solved by Patterson methods subsequent to the study of E maps calculated with phases determined by direct methods. Full-matrix least-squares refinement yielded an R of 0.045 for 827 counter reflexions. The application of phase correction procedures also led to the structure solution. The gluconamide molecule has a linear conformation with its long axis approximately parallel to c , while hydrogen bonds along a and b link the molecules in the crystal into infinite sheets parallel to the ab plane.

Introduction

The interaction between glycosidases and glyconamides is not fully understood. For instance, it is not known why the enzymes have an affinity for glyconamides and not linear glycoside analogues, or why D-gluconamides, for example, interact with β -glucosidases but not with α -glucosidases (Jermyn, 1972). The compound *N*-(2-chloroethyl)-D-gluconamide, a potent inactivator of β -glucosidases, was thought to have a cyclic structure, *e.g.* (I) ($R = NHCH_2CH_2Cl$), analogous to α -D-glucopyranose. The possibility of the linear conformation (II), however, was not discarded.



To establish the conformation of alkyl glyconamides, an X-ray analysis of crystals of *N*-(2-chloroethyl)-D-gluconamide (Fig. 1) was undertaken. For glycosides, it is generally accepted that molecular conformation in the crystal and in solution is similar (Bentley, 1972). If this is also true for glyconamides, establishment of their conformation should assist in a further understanding of their interaction with glycosidases. A brief report of this structure has been presented earlier (Jermyn, Mackay & Satzke, 1973).

Experimental

Crystals of *N*-(2-chloroethyl)-D-gluconamide, $C_8H_{16}ClNO_6$, grown from an *n*-propyl alcohol solution in the form of triclinic prisms, were supplied by Dr M. A. Jermyn. Weissenberg photographs revealed that the crystals were triclinic. The cell parameters were determined by a least-squares fit of measured θ , φ and χ angles for 11 strong reflexions on a four-circle diffractometer. The density was determined by flotation

in a mixture of bromoform and xylene and indicated that $Z=1$. The space group of the optically active crystal was therefore $P1$.

Crystal data

N-(2-Chloroethyl)-D-gluconamide, $C_8H_{16}ClNO_6$. Triclinic, space group $P1$. F. W. 289.67, $F(000)=136$, $a=4.807$ (5), $b=5.093$ (5), $c=11.888$ (7) Å, $\alpha=91.25$ (6), $\beta=84.48$ (6), $\gamma=94.48$ (6) $^\circ$; $U=288.75$ Å 3 , $D_m=1.48$ (1), $D_x=1.475$ g cm $^{-3}$, $Z=1$; $\lambda=1.5418$ Å, $\mu(Cu K\alpha)=31.0$ cm $^{-1}$.

A set of non-equivalent intensities was measured on a Picker four-circle automatic diffractometer with Ni-filtered Cu $K\alpha$ radiation and an ω - 2θ scan. Of the 961 terms measured out to 0.85 Å, 827 had values significantly greater than background. The 200 reflexion was taken as a standard and measured after every 50 reflexions to check the stability of the crystal. During data collection the reference intensity showed a gradual decrease to 95% of its initial value; the intensities were therefore scaled to the reference standard. The usual corrections were made for Lorentz and polarization factors, but not for absorption. Scattering factors were those given in *International Tables for X-ray Crystallography* (1962).

Structure determination and refinement

An electron density synthesis phased on the Cl atom was calculated, the Cl atom being positioned at the origin. A pseudo centre of symmetry thus introduced into the $P1$ distribution and the presence of a large number of peaks made location of the lighter atoms difficult. Electron density and Patterson syntheses, sharpened by the modification function of Wunderlich (1965), did not help to clarify the situation.

In another attempt to elucidate the structure, direct methods were applied with *MULTAN* (Main, Woolfson & Germain, 1971). Phases for 284 terms with $|E| > 1.0$ were derived by application of the tangent

formula (Karle & Hauptman, 1956). From the resultant E map, peaks were selected which when assembled could feasibly correspond to atom sites forming a six-membered ring of the chair form. Phases derived from these atomic coordinates were recycled, and in the subsequent E map the peaks assigned to two ring atom sites had decreased considerably in height. Other peaks relatively enhanced in height now indicated five atoms, C(2), C(3), C(4), C(5) and C(6) lying in a

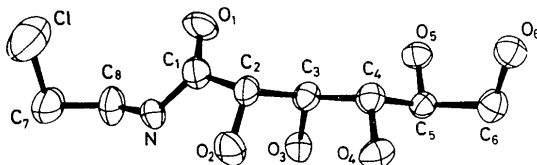


Fig. 1. A perspective view of the molecule with thermal ellipsoids scaled to 60% probability. The diagram was prepared from the output of the *ORTEP* program (Johnson, 1965).

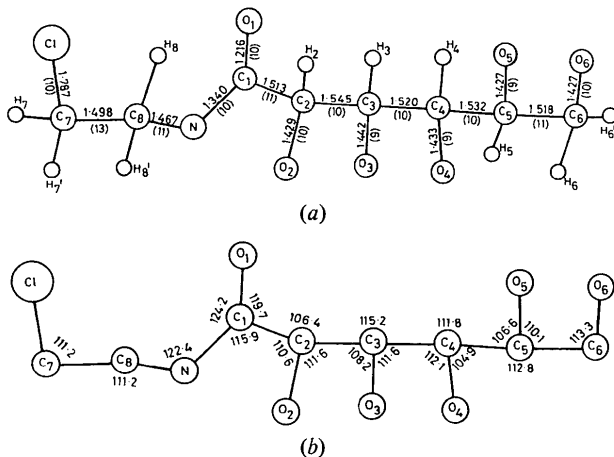


Fig. 2. Bond lengths and angles involving the non-hydrogen atoms. (a) Bond lengths with their estimated standard deviations given in parentheses. (b) Bond angles. The estimated standard deviations are 0.6°.

plane, linked in a linear fashion and extending along c [Fig. 3(a)]. Re-examination of the vector maps at this stage led to the location of three additional atomic sites and it was apparent that the molecule had a linear conformation. The remaining seven non-hydrogen atoms were located on a difference map and the atoms distinguished as C, N or O by chemical considerations. After full-matrix least-squares refinement in which anisotropic temperature factors were given to all the located atoms, an R of 0.082 for the 827 observed terms was attained; $R = \sum ||F_o| - |F_c|| / \sum |F_o|$.

Location of the H atoms was only partially successful. The ten H atoms bonded to the C atoms were located on difference maps, but those bonded to N and O, all of which presumably participate in hydrogen bonding, were not located. The final refinement reduced R to 0.045. The ten included H atoms were given the same isotropic thermal parameter as the C atom to which they were bonded and only their positional parameters were refined. The refinement was carried out with unit weight given to all terms with *ORFLS* (Busing, Martin & Levy, 1962), while the Fourier summations were calculated with *MUFR* (White, 1966).

Final atomic coordinates and thermal parameters are given in Table 1.* Bond lengths and angles are given in Fig. 2 and Table 2; no correction has been made for thermal effects. Some short intermolecular approaches including hydrogen-bond distances are given in Table 3 and Fig. 3.

Phase correction

The difficulty encountered in deriving the structure of the gluconamide molecule arising from the low Cl atom contribution to the total scattering and the centre of symmetry introduced into the Fourier maps when

* A list of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 30812 (8 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

Table 1. Final atomic parameters

The hydrogen atom coordinates are multiplied by 10^3 and all other parameters by 10^4 . Estimated standard deviations are given in parentheses and the expression used for the temperature factor was $\exp [-(h^2\beta_{11} + k^2\beta_{22} + l^2\beta_{33} + 2hk\beta_{12} + 2hl\beta_{13} + 2kl\beta_{23})]$.

	x/a	y/b	z/c	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
Cl	0000	0000	0000	818 (17)	791 (16)	77 (2)	-150 (12)	-54 (4)	-37 (4)
C(1)	-147 (17)	1288 (16)	-3040 (6)	441 (37)	307 (34)	45 (5)	-21 (29)	-8 (11)	28 (11)
C(2)	2074 (16)	1265 (15)	-4032 (6)	376 (36)	264 (30)	41 (5)	-33 (27)	0 (11)	11 (9)
C(3)	656 (16)	2036 (15)	-5077 (6)	283 (32)	273 (30)	44 (5)	-40 (26)	-2 (10)	0 (10)
C(4)	2651 (15)	2809 (15)	-6106 (6)	288 (32)	247 (29)	43 (5)	4 (25)	-7 (10)	-2 (9)
C(5)	1076 (16)	3554 (15)	-7099 (6)	317 (36)	247 (29)	41 (5)	51 (26)	-8 (10)	-2 (10)
C(6)	3029 (17)	4379 (16)	-8132 (6)	351 (37)	354 (35)	39 (5)	-31 (29)	-23 (11)	9 (11)
C(7)	-2306 (22)	-2406 (20)	-627 (7)	742 (61)	476 (43)	52 (6)	-75 (42)	-7 (16)	28 (13)
C(8)	-3288 (20)	-1394 (19)	-1683 (8)	456 (39)	484 (45)	58 (6)	66 (34)	40 (12)	42 (12)
O(1)	-1153 (16)	3340 (12)	-2744 (5)	885 (44)	203 (22)	74 (5)	107 (24)	91 (11)	-2 (8)
O(2)	3137 (12)	-1275 (11)	-4175 (5)	342 (25)	301 (22)	61 (4)	73 (19)	-1 (8)	30 (7)
O(3)	-1354 (11)	-92 (11)	-5335 (4)	235 (21)	284 (21)	49 (3)	-37 (16)	-18 (7)	11 (7)
O(4)	4253 (11)	679 (11)	-6528 (4)	242 (23)	300 (21)	52 (4)	83 (18)	10 (7)	-15 (7)
O(5)	-557 (12)	5674 (11)	-6710 (4)	320 (25)	263 (20)	53 (4)	49 (18)	-14 (8)	-3 (7)
O(6)	5079 (11)	6452 (11)	-7903 (4)	281 (24)	325 (22)	49 (4)	-30 (18)	-13 (8)	15 (7)
N	-976 (14)	-1056 (12)	-2580 (5)	401 (30)	252 (25)	42 (4)	-21 (23)	29 (9)	18 (8)

Table 1 (cont.)

	<i>x/a</i>	<i>y/b</i>	<i>z/c</i>
H(2)	-631 (17)	278 (17)	-384 (7)
H(3)	-76 (19)	378 (17)	-476 (8)
H(4)	-623 (18)	437 (19)	-588 (7)
H(5)	-47 (19)	185 (17)	-729 (7)
H(6)	-597 (18)	258 (19)	-844 (8)
H(6')	-817 (18)	504 (17)	-872 (8)
H(7)	-398 (22)	-295 (21)	3 (9)
H(7')	-154 (21)	-446 (20)	-72 (8)
H(8)	-406 (20)	83 (20)	-154 (8)
H(8')	-518 (22)	-318 (19)	-183 (8)

Table 2. Bond lengths (Å) and angles (°) involving the ten located hydrogen atoms

Estimated standard deviations for the bond lengths are 0.09 Å and for the bond angles 5°.

C(2)-H(2)	1.09	C(6)-H(6')	1.03
C(3)-H(3)	1.19	C(7)-H(7)	1.09
C(4)-H(4)	0.97	C(7)-H(7')	1.14
C(5)-H(5)	1.13	C(8)-H(8)	1.22
C(6)-H(6)	1.11	C(8)-H(8')	1.26
C(1)-C(2)-H(2)	106	C(4)-C(5)-H(5)	109
O(2)-C(2)-H(2)	112	C(6)-C(5)-H(5)	112
C(3)-C(2)-H(2)	110	O(5)-C(5)-H(5)	106
C(2)-C(3)-H(3)	105	C(5)-C(6)-H(6)	106
C(4)-C(3)-H(3)	112	H(6')-C(6)-H(6)	111
O(3)-C(3)-H(3)	104	H(6')-C(6)-O(6)	108
C(3)-C(4)-H(4)	106	C(5)-C(6)-H(6')	108
C(5)-C(4)-H(4)	107	O(6)-C(6)-H(6)	111
O(4)-C(4)-H(4)	114		

Table 3. Intermolecular distances less than 3.70 Å

Symmetry code			
(i)	<i>x</i> -1	<i>y</i>	<i>z</i>
(ii)	<i>x</i>	<i>y</i> +1	<i>z</i>
(iii)	<i>x</i>	<i>y</i>	<i>z</i> +1
(iv)	<i>x</i> -1	<i>y</i> -1	<i>z</i> +1
(v)	<i>x</i> -1	<i>y</i> +1	<i>z</i>
Cl...C(6 ¹¹¹)	3.45	O(5)...C(6 ¹)	3.67
Cl...O(6 ^{1v})	3.67	O(5)...O(4 ¹)	3.42
C(3)...O(2 ¹)	3.69	O(5)...O(4 ^v)	3.69
C(3)...O(4 ¹)	3.68	O(5)...O(4 ¹¹¹)	3.32
C(5)...O(6 ¹)	3.56	N...O(2 ¹)	3.55
C(5)...O(4 ¹)	3.50	O(1)...N ¹¹¹	2.85
C(7)...O(6 ^{1v})	3.41	O(3)...O(2 ¹)	2.89
O(1)...O(2 ¹¹¹)	3.65	O(3)...O(4 ¹)	2.71
O(1)...C(8 ¹¹¹)	3.14	O(5)...O(3 ¹¹¹)	2.71
O(3)...C(2 ¹)	3.49	O(6)...O(4 ¹¹¹)	2.71
O(3)...C(4 ¹)	3.54	O(5)...O(6 ¹)	2.70
O(5)...C(4 ¹)	3.49		

phased only on the Cl contribution, suggested that phase correction procedures (Hoppe & Gassmann, 1968) could have led more directly to the structure. The applicability of this procedure in this present instance has therefore been investigated.

Initial electron densities were determined with the phasing atoms Cl+C(7) and Cl+C(7)+C(8) respectively, the approximate coordinates of C(7) and C(8) being derived from the Cl-phased Fourier map (Table 4). Both linear (Hoppe, Gassmann & Zechmeister, 1970) and non-linear modifications, including maximum and minimum density truncation (Barrett &

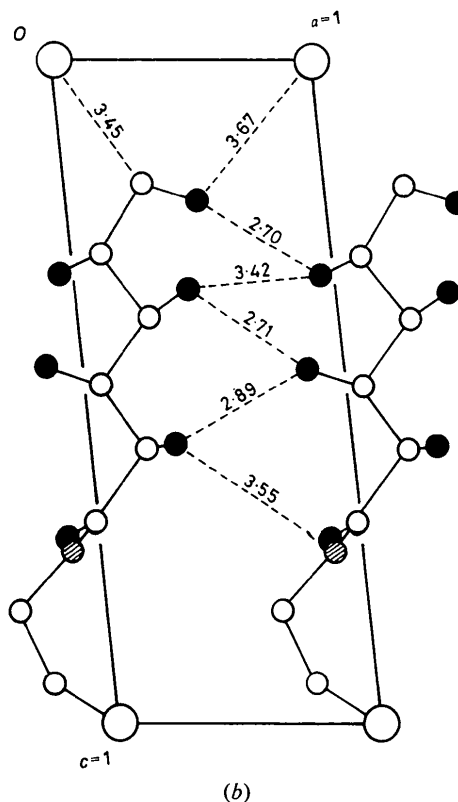
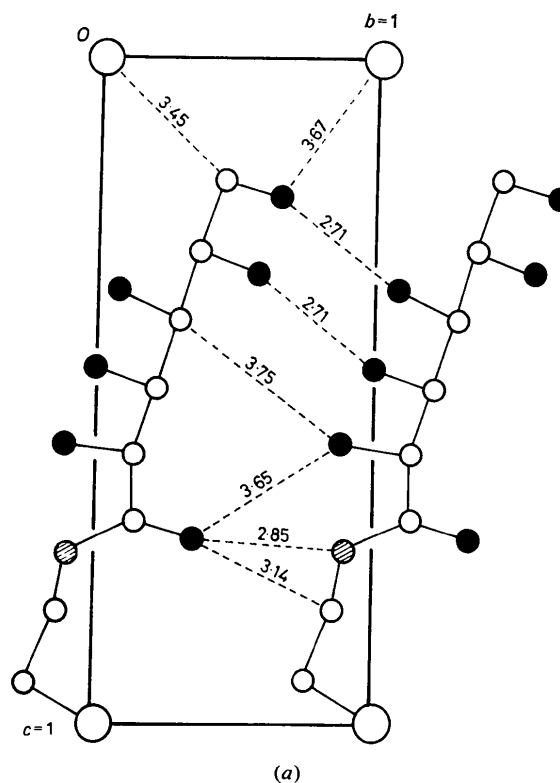


Fig. 3. The projection of the structure down (a) the *a* axis, (b) the *b* axis.

Zwick, 1971), were applied to the electron densities after subtraction of the contribution of one, two (or three) atoms of the partial structure from the initial density. The linear density modification was of the form, $p' = a + bp$, where $a = -\frac{1}{3}$ and $b = \frac{2}{9}$, and the non-linear modification was that given by Hoppe & Gassmann (1968) with $T = 0.3$. The program employed was written specifically for P1 (Satzke, 1974).

The series of phase refinements which were carried out are summarized in Table 4 while Fig. 4 illustrates

Table 4. Summary of phase correction refinements

(a) Atomic coordinates for the partial structures derived from the Cl-phased Fourier map. The refined values have been included in parentheses.

	x/a	y/b	z/c
Cl	0.0000	0.000	0.000
C(7)	-0.233 (-0.2306)	-0.250 (-0.2406)	-0.046 (-0.0627)
C(8)	-0.358 (-0.3288)	-0.104 (-0.1394)	-0.158 (-0.1683)

(b) The density modifications applied to the initial densities.

	Partial structure	Density subtracted before refinement	Modification function
1	Cl + C(7)	Cl	Linear
2	Cl + C(7)	Cl	Non-linear
3	Cl + C(7)	Cl + C(7)	Linear
4	Cl + C(7)	Cl + C(7)	Non-linear
5	Cl + C(7) + C(8)	Cl	Linear
6	Cl + C(7) + C(8)	Cl	Non-linear
7	Cl + C(7) + C(8)	Cl + C(7)	Linear
8	Cl + C(7) + C(8)	Cl + C(7)	Non-linear
9	Cl + C(7) + C(8)	Cl + C(7) + C(8)	Linear

the cyclic variation of the mean phase deviation of phases for the 460 largest $|F|$ terms from the true phases (*i.e.* phases calculated from the atomic parameters given in Table 1) for each of these refinements. The results indicate that phase correction could lead to the correct structure when the initial density corresponded to $> 20\%$ of the total scattering – see (6), (7), and (8), Fig. 4 – and that the nature of the density modification function was important as illustrated by comparison of (5) with (6) and (7) with (8) respectively. The convergence of the calculated phases to the true values was more rapid as the amount of scattering subtracted from the initial density was increased – see (6) and (8) – moreover, with a similar comparison of (5) and (7) the former has failed to converge after seven refinement cycles. The termination of phase refinement also noted in (9) can be attributed to the inaccurate location of C(8) in accord with similar results of Hoppe & Gassmann (1968). The most rapid phase refinement was obtained in (7) in which the mean phase deviation decreased from 50° to 31° after nineteen refinement cycles. In the subsequent Fourier map, the complete gluconamide structure was revealed and the highest background peak present had a height of only one-third the height of the lowest peak corresponding to an authentic atom.

Description of the structure

The molecule has a linear conformation as shown in Fig. 1. Five atoms in the carbon backbone, C(2), C(3),

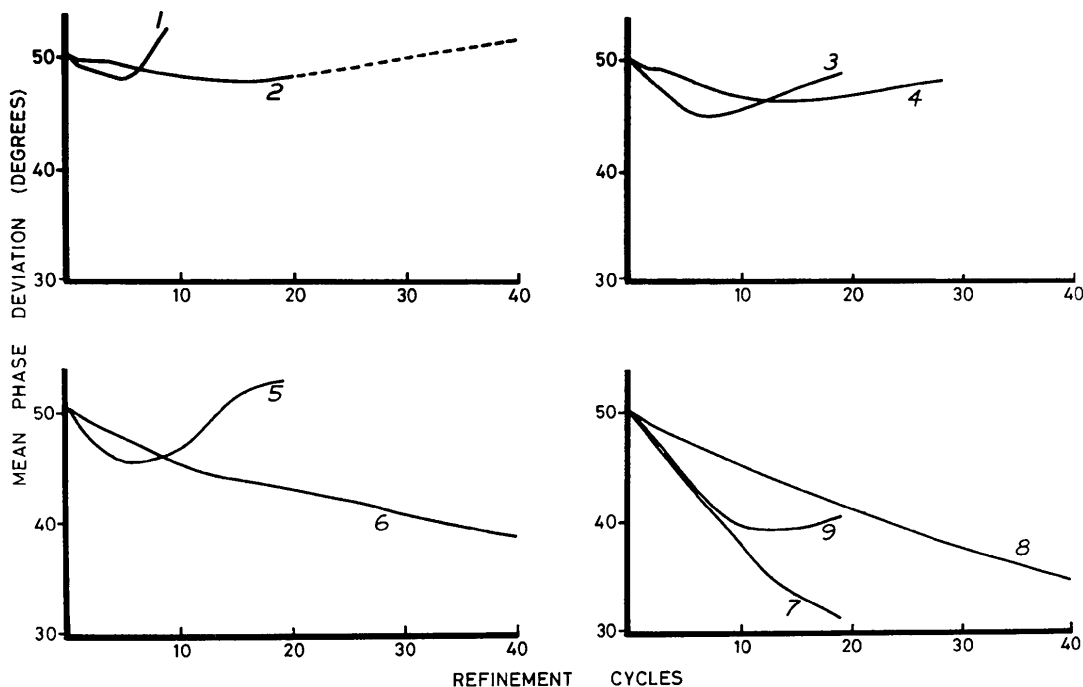


Fig. 4. Variation of the mean phase deviation of phases for the 460 largest $|F|$ terms from the true phases per refinement cycle, for the series of refinements in Table 4.

C(4), C(5) and C(6) are planar within 0.01 Å, the plane equation being

$$-0.0118X + 0.9467Y + 0.3219Z + 5.6264 = 0,$$

where *X*, *Y* and *Z* refer to orthogonal coordinates in Å. The planarity of the amide group is described by a root mean square distance of 0.02 Å of the five atoms, C(8), N, C(1), C(2) and O(1), from the least-squares plane

$$0.7438X + 0.1328Y + 0.6551Z + 6.5929 = 0,$$

the angle between the two groups being 69.8°. The N atom has the greatest deviation 0.037 Å from the amide plane as was also observed by Johnson (1966) in the structure of an α -D-glucosamine. The dimensions of the group do not differ significantly from those of amide groups in structures of dipeptides (Biswas, Hughes, Sharma & Wilson, 1968; Koch & Germain, 1970) and other simple peptides and related substances (Corey & Pauling, 1953), with the C(1)–N bond of 1.340 Å exhibiting partial double-bond character typical of amides.

The C–C bonds have a mean of 1.521 Å (mean deviation 0.012 Å), and the five C–OH bonds a mean of 1.432 Å (maximum deviation 0.010 Å) which agrees with other values reported for comparable compounds (Berman & Kim, 1968). The C(7)–Cl bond of 1.787 Å is typical of values reported for such bonds (Sutton, 1965).

The molecules are oriented in the crystal with their long molecular axis approximately parallel to *c*. Hydrogen bonds link the molecules related by the *a* and *b* translations to form infinite sheets parallel to the *ab* plane (Fig. 3), whereas along *c* the molecules are held together only by van der Waals interactions. The N and all the O atoms are involved in intermolecular hydrogen bonding and it is probable that some of these atoms also participate in intramolecular hydrogen bonding. As the H atoms which are bonded to the O and N atoms were not located in the analysis, a detailed description of the hydrogen bonding cannot be given. Four hydrogen-bonding O...O distances have values 2.70 or 2.71 Å (Table 3) which are comparable with those observed in carbohydrates and related compounds (Gatehouse & Poppleton, 1971; Park, Kim & Jeffrey, 1971). A weaker intermolecular hydrogen bond, O(3)...O(2), of 2.89 Å is also present, O(3) forming two other hydrogen bonds which involve O(4) and O(5) respectively. The N atom acts as a donor to the carbonyl O(1) of the adjacent molecule related by the translation along *b* [Fig. 3(a)], with N...O(1) = 2.85 Å. Other short intermolecular distances are listed in Table 3.

Since the extended structure is maintained by intermolecular hydrogen bonding in the crystal, the only portion that would be conserved in aqueous solution is

the planar amide group. The X-ray analysis therefore gives little indication of the conformation in which the molecule interacts with the active site of an enzyme.

We are grateful to Dr M. A. Jermyn, Division of Protein Chemistry, CSIRO, Melbourne for introducing this interesting problem to us, for supplying the crystals and for valuable discussion. We also thank Dr A. McL. Mathieson and Dr W. A. Denne, both of the Division of Chemical Physics, CSIRO, Melbourne, the former for his interest and for allowing us use of the Picker diffractometer and the latter for assistance with the data collection.

References

- BARRETT, A. N. & ZWICK, M. (1971). *Acta Cryst.* **A27**, 6–11.
- BENTLEY, R. (1972). *Ann. Rev. Biochem.* **41**, 953–996.
- BERMAN, H. M. & KIM, S. H. (1968). *Acta Cryst.* **B24**, 897–904.
- BISWAS, A. B., HUGHES, E. W., SHARMA, B. D. & WILSON, J. N. (1968). *Acta Cryst.* **B24**, 40–50.
- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. Oak Ridge National Laboratory Report ORNL-TM-305.
- COREY, R. B. & PAULING, L. (1953). *Proc. Roy. Soc.* **B141**, 10–20.
- GATEHOUSE, B. M. & POPPLETON, B. J. (1971). *Acta Cryst.* **B27**, 871–876.
- HOPPE, W. & GASSMANN, J. (1968). *Acta Cryst.* **B24**, 97–107.
- HOPPE, W., GASSMANN, J. & ZECHMEISTER, K. (1970). In *Crystallographic Computing*, pp. 26–36, edited by F. R. AHMED. Copenhagen: Munksgaard.
- International Tables for X-ray Crystallography* (1962). Vol. III, pp. 202–203. Birmingham: Kynoch Press.
- JERMYN, M. A. (1972). *Proc. Aust. Biochem. Soc.* **5**, 12.
- JERMYN, M. A., MACKAY, M. F. & SATZKE, L. (1973). *Proc. Aust. Biochem. Soc.* **6**, 17.
- JOHNSON, C. K. (1965). *ORTEP*. Oak Ridge National Laboratory Report ORNL-3794.
- JOHNSON, L. N. (1966). *Acta Cryst.* **21**, 885–891.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
- KOCH, M. H. J. & GERMAIN, G. (1970). *Acta Cryst.* **B26**, 410–417.
- MAIN, P., WOOLFSON, M. & GERMAIN, G. (1971). *MULTAN. A Computer Program for the Automatic Solution of Crystal Structures*. Univ. of York.
- PARK, Y. J., KIM, H. S. & JEFFREY, G. A. (1971). *Acta Cryst.* **B27**, 220–227.
- SATZKE, L. O. G. (1974). Ph.D. thesis, Department of Physical Chemistry, La Trobe Univ.
- SUTTON, L. E. (1965). *Interatomic Distances and Configuration in Molecules and Ions*, Supplement. Special Publication No. 18, M63s. London: The Chemical Society.
- WHITE, J. C. B. (1966). *MUFR. A Three-Dimensional Fourier Synthesis Program*. Chemistry Department, Univ. of Melbourne.
- WUNDERLICH, J. A. (1965). *Acta Cryst.* **19**, 200–202.