

## Polymorphs and Conformations of the Cytidylic Acids

S. FURBERG

*Institute of Chemistry, University of Oslo, Blindern-Oslo, Norway*

Unit cell dimensions and space groups have been determined for three polymorphs of cytidylic acid "a" and two polymorphs of cytidylic acid "b". The conformations of the molecules are discussed.

It was shown by Harris *et al.*<sup>1</sup> that cytidylic acid "a" (cytidine-2-phosphate) and cytidylic acid "b" (cytidine-3-phosphate) both exist in two stable crystalline forms called the *alc*-form and the  $H_2O$ -form. The different polymorphs were characterized by their infra-red spectra, but crystal data were not given. In connection with other X-ray crystallographic studies of nucleic acid components, we have determined unit cell dimensions of the polymorphs, thus providing an additional method for their identification. A third polymorphic form of the "a"-isomer, which is metastable, was also examined.

Crystal data were derived from oscillation and Weissenberg diagrams taken with Cu radiation ( $\lambda = 1.542 \text{ \AA}$ ). The cell dimensions given are believed to be accurate to within 1 %. The space groups were determined from systematic absences and the densities measured by the flotation method. None of the crystal forms contain water of crystallization<sup>1</sup>.

*Cytidylic acid "b"*. Following the procedure described by Harris *et al.*, the two polymorphs were obtained together by letting a saturated aqueous solution evaporate on a watch-glass.

The *alc*-form crystallizes in the monoclinic system, with  $a = 5.94 \text{ \AA}$ ,  $b = 16.96 \text{ \AA}$ ,  $c = 6.89 \text{ \AA}$  and  $\beta = 118^\circ$ . The space group is  $P2_1$  and the unit cell contains two molecules. The observed density is  $1.73 \text{ g/cm}^3$ , the calculated value  $1.75 \text{ g/cm}^3$ . This polymorphic form was found to occur in a variety of different crystal habits. Short polar needles elongated along  $b$  with  $c$  lying in the predominant face were observed, but usually the crystals are elongated along the diagonal  $[101]$ , being either rod-like, or developed as flat plates (010). The length of this diagonal is  $6.66 \text{ \AA}$ .

The  $H_2O$ -form is readily obtained as big flat plates several mm of extension. In general, they are somewhat elongated along  $c$ , with (010) as the dominating face. Forms  $\{011\}$ ,  $\{001\}$ ,  $\{100\}$ , and  $\{110\}$  were also observed. The crystals are biaxial positive with (001) as the optic axial plane. They crystallize in the

orthorhombic system, with unit cell dimensions  $a = 8.80 \text{ \AA}$ ,  $b = 21.7 \text{ \AA}$  and  $c = 6.85 \text{ \AA}$ . The space group is  $P2_12_12$ , and there are four molecules in the unit cell. The density was found to be  $1.66 \text{ g/cm}^3$  (calculated  $1.64 \text{ g/cm}^3$ ).

*Cytidylic acid "a"*. A sample of the *alc-form* was kindly supplied by Dr. R.J.C. Harris of the Central Laboratories, Imperial Cancer Research Fund. It consisted of very small needles with  $b$  as the needle axis. On the basis of a number of weak oscillation diagrams taken around the needle axis it was concluded that the crystals are monoclinic, with  $a = 16.97 \text{ \AA}$ ,  $b = 7.50 \text{ \AA}$ ,  $c = 11.80$  and  $\beta = 119^\circ$ . The space group is  $C2$ , as reflexions with  $h + k$  odd are absent and the molecules are optically active. The calculated density assuming four molecules per unit cell, is  $1.68 \text{ g/cm}^3$ .

By letting a saturated aqueous solution evaporate on a watch-glass, two other polymorphic forms were obtained, one as a crust at the edge and one deposited from the middle of the solution. The latter is probably the  $H_2O$ -form of Harris *et al.* The former is metastable and is transformed into the *alc-form* described above in the course of a few days. The transformation is made evident by the loss of extinction and by changes in the X-ray diagrams.

The *metastable* form yields needle-shaped crystals elongated along the  $c$  axis. The crystals grow either together in spherulitic formations of needles, or alternatively well-developed crystals with (100) as the dominating face are formed. They give straight extinction and may be recognized under the microscope by the angle of  $57^\circ$  between (010) and (021). The crystals are monoclinic, with  $a = 9.32 \text{ \AA}$ ,  $b = 22.8 \text{ \AA}$ ,  $c = 7.38 \text{ \AA}$  and  $\beta = 95^\circ$ . Space group  $P2_1$ . Four molecules in the unit cell give a calculated density of  $1.38 \text{ g/cm}^3$ . This density is much lower than that of the four other crystalline cytidylic acids, for which values between  $1.64$  and  $1.75 \text{ g/cm}^3$  are found. The packing of the molecules must be a very loose one and it is not surprising that the crystals are less stable than those of the other polymorphs.

The crystals of the  $H_2O$ -form were extremely small, and only feeble diffraction patterns were obtained. These show that the crystals probably are orthorhombic, with  $a = 7.35 \text{ \AA}$ ,  $b = 25.2 \text{ \AA}$  and  $c = 6.66 \text{ \AA}$ . The extinctions indicate the space group  $P2_12_12_1$  or  $P2_12_12$ . The calculated density is  $1.74 \text{ g/cm}^3$ . The crystals have the shape of flat prisms elongated along  $c$ . Faces (100), (001) and (010) were observed, the latter being predominant.

*Identification.* The data given above provide a method for identification of extremely small quantities of the different polymorphs and isomorphs of cytidylic acids. One single crystal of dimensions greater than say  $0.05 \text{ mm}$  will suffice with ordinary cameras and X-ray sources.

In general, the crystals are more or less elongated along one direction and the crystals may be identified by taking an oscillation diagram around this axis. The polymorphs are distinguishable by the length of the corresponding crystallographic axis, and by the symmetry, where some of them give diagrams with a plane of symmetry perpendicular to the axis of oscillation, others not (denoted "plane" and "no plane", respectively, in Table 1). Table 1 may serve as a key for the identification.

All the crystals give straight extinction when lying on the predominant face, except some of the habits of "b" *alc* and the metastable form.

Table 1. The characteristics of oscillation diagrams of crystals of cytidylic acid.

Length of axis, Å	Symmetry	Polymorph
6.66	No plane	"b" alc.
6.66	Plane	"a" H <sub>2</sub> O
6.85	Plane	"b" H <sub>2</sub> O
7.38	No plane	"a" metastable
7.50	Plane	"a" alc.
16.96	Plane	"b" alc.

*Molecular shape.* The molecular structure of nucleotides is a problem of considerable interest because of its relation to the structure of nucleic acids. We have carried out a complete structure analysis of the "b" H<sub>2</sub>O polymorph and determined the molecular structure<sup>2</sup>. In the crystals the NH<sub>2</sub>-group in one molecule is connected by hydrogen bonds to PO<sub>4</sub>-groups in neighbouring molecules and the molecular axis is roughly parallel to *b*, the longest cell edge. It seems reasonable to assume that the amino and phosphate groups are linked together by hydrogen bonds also in the other polymorphic forms. Taking also the unit cell dimensions and the space group into account, this permits some conclusions to be drawn about the general shape of the molecules and their orientation in the crystals.

Considering first the "b"-isomer the molecular dimensions in the known structure ("b" H<sub>2</sub>O) are roughly 4.5 × 7 × 11 Å, whereas in the "b" alc form the molecule probably extends only about 8.5 Å in the *b* direction. This indicates a considerably shorter molecule, and it seems likely that the molecule exhibits different shapes in the two polymorphic crystal forms.

Two main factors govern the general shape of the cytidylic acid molecule, namely (i) at the conformation of the ribose ring and (ii) the orientation of the phosphate group. From the crystal structure analysis of cytidine and cytidylic acid "b" H<sub>2</sub>O two different ribose conformations are known to occur, one with C<sub>2'</sub> out of the plane of the four other atoms, and one with C<sub>3'</sub> out of the plane<sup>2</sup>. As for the orientation of the phosphate group, one would expect a staggered arrangement of the bonds around the C<sub>3'</sub>-O<sub>3'</sub> linkage. There are three such orientations, but one of these corresponds to a *cis* relation between phosphate and ribose and would seem to be less favourable than the other two. Thus the molecule may be expected to occur in four general shapes:

- (1) C<sub>2'</sub> out of the plane; P—O<sub>3'</sub> nearly parallel to C<sub>2'</sub>—C<sub>3'</sub>
- (2) » » » » » » » » C<sub>3'</sub>—C<sub>4'</sub>
- (3) C<sub>3'</sub> » » » » » » » » C<sub>2'</sub>—C<sub>3'</sub>
- (4) » » » » » » » » C<sub>3'</sub>—C<sub>4'</sub>

Conformations (1) and (2) are shown in Fig. 1. The length of the molecule, as measured by the distance P—N<sub>3</sub>, is about 9.8 Å, 8.8 Å, 9.1 Å and 8.5 Å in the four conformations, respectively. These values are based on bond lengths and angles derived by the X-ray analysis of cytidylic acid "b" H<sub>2</sub>O. Minor

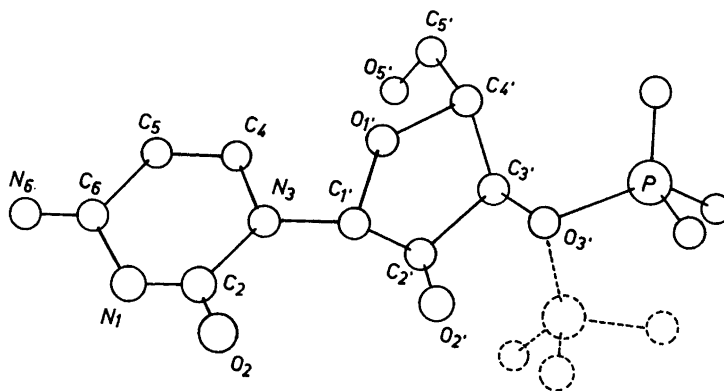


Fig. 1. Conformations of cytidylic acid "b" with atom  $C_3$ , out of the ring plane.

variations in molecular shape may also be obtained by rotation around the bonds  $N_3-C_1'$  and  $P-O_3$ , but these changes do not alter the length of the molecule and need not be considered in this connection.

In the known structure of the  $H_2O$  polymorph the molecule has conformation (1), the most extended one. As stated above, the molecule would seem to be less extended in the alc polymorph and one of the other three conformations occurs. It may also be mentioned that current models of DNA are based on a nucleotide conformation of type (4).

Turning now to cytidylic acid "a", there are also in this case four stable conformations. It is seen from models, however, that the molecule is less extended than that of the "b" isomer, the amino and phosphate groups being in closer proximity. The  $P-N_6$  distances are 7.5 Å and 5.8 Å, respectively, with  $C_2$  out of the plane, and 7.3 Å and 8.0 Å with  $C_3$  out of the plane. In the three first-mentioned conformations the molecule is very roughly equally extended in all directions, in agreement with the cell dimensions of the  $H_2O$  form, which may be interpreted as indicating four molecules of dimensions  $7 \times 7 \times 6$  Å along the  $b$  axis. Similar molecular dimensions may be derived from the data obtained for the alc-form. On the basis of the data given above it is not possible to decide whether or not the molecule has the same conformation in the two cases. As for the unstable modification, the low density of this form makes a comparison with the other polymorphs difficult.

Finally, it may be remarked that the discussion of the conformations of the cytidylic acids given above is valid also for other 2' and 3' nucleotides. For each isomer there should be four conformations as far as variations in ribose puckering and orientation of the phosphate group are concerned. In the case of the 5'-nucleotides, there are more conformations because of the additional single bond  $C_4'-C_5'$ .

The author wishes to thank Dr. W. E. Cohn of the Oak Ridge National Laboratories and Dr. R. J. C. Harris of the Central Laboratories, Imperial Cancer Research Fund, for supplying samples of the isomers. The investigation was supported by *Norsk Hydros fond*.

## REFERENCES

1. Harris, R. J. C., Orr, S. F. D., Roe, E. M. F. and Thomas, J. F. *J. Chem. Soc.* **1953** 489.
2. Alver, E. and Furberg, S. *Acta Chem. Scand.* **13** (1959) 910.

Received September 8, 1959.