Crystal and Molecular Structure of 3,N^4-Ethenocytidine Hydrochloride. A Study of the Dimensions and Molecular Interactions of the Fluorescent ϵ-Cytidine System

Andrew H. -J. Wang, Jorge R. Barrio, and Iain C. Paul*

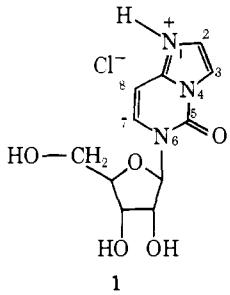
Contribution from the Department of Chemistry, School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801. *Received March 3, 1976*

Abstract: The crystal structure of 3,N^4-ethenocytidine hydrochloride (**1**, ϵ-Cyd·HCl) has been determined to provide the molecular dimensions of this fluorescent analogue of a cytosine nucleoside and to probe its capability for molecular interaction. The crystals of **1** are monoclinic, with a = 6.330 (1), b = 16.968 (2), c = 6.695 (1) Å, and β = 115° 22 (1)′, and there are two $C_{11}H_{14}N_3O_5^+Cl^-$ entities in the space group $P2_1$. The structure has been refined to an R factor of 0.045 on 1675 nonzero reflections (Mo Kα). The ϵ-cytosine moiety is slightly nonplanar with a maximum deviation of 0.037 Å from the best plane through the ten nonhydrogen atoms; there are some minor differences in bond lengths and angles from those found in protonated unbridged cytosine derivatives. The arrangement about the glycosyl bond is anti ($\chi_{C\text{-}N}$ = 42.6°) and the ribose ring exists in the C(2′)-endo or the C(2′)-endo-C-(3′)-exo conformation. The C(4′)-C(5′) exocyclic bond is in a gauche-gauche arrangement. There appears to be a C(6)-H(6)- - -O(5′) intramolecular hydrogen bond. The N-H bond in the base and the three hydroxyl groups in the sugar are involved in an intermolecular hydrogen bonding arrangement in which the chloride anion acts as an acceptor for three hydrogen bonds. Translationally equivalent ϵ-cytosine bases crystallize such that a chloride anion lies between them in the plane of the rings in a fashion similar to that observed for other unbridged cytosine salts. This arrangement results in some short H- - -Cl^- and H- - -O contacts. This packing also results in another Cl^- lying 3.21 Å above the plane of the ϵ-cytosine cation. This latter type of "ion pair" interaction is quite common in anhydrous halide salts of nucleic acid bases and nucleosides. There is no base-base overlap.

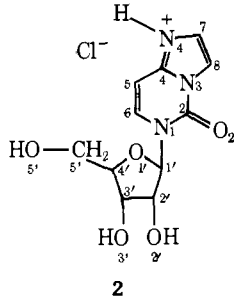
The systematic examination of the interactions of coenzyme analogues, together with a comparison of their kinetic behavior relative to the natural substrate, provides information concerning the nature of the binding sites of the functioning enzymes. Analogues have been used in the past, e.g., in studies of the mechanisms of enzyme action[1] and of the evolution of enzyme structures.[2-4] Recent progress in the determination of the three-dimensional structure of enzymes[5] has increased understanding of the mechanisms of enzyme action.[6]

Fluorescent substrate analogues of cytosine nucleotides, obtained by reaction of cytosine derivatives with chloroacetaldehyde,[7-10] have been shown to replace the natural substrates in several enzymatic reactions. The introduction of the second ring on the cytosine portion gives the new molecule a similar spatial outline and similar potential binding areas to those of the corresponding adenine nucleotides (Figure 1). Indeed, the chloroacetaldehyde-modified cytosine nucleotides mimic the structural features of the natural coenzymes and replace adenosine nucleotides in enzymatic phosphorylation[9] and photophosphorylation.[11]

Since various substituted ϵ-Cyd compounds are readily accessible[8,9,12] and in view of the observations concerning the behavior of ϵ-cytidine coenzyme analogues, we undertook an x-ray study of 3,N^4-ethenocytidine hydrochloride (**1**, ϵ-Cyd·HCl).[8,13] This study should provide the molecular dimensions of the ϵ-cytidine molecule and allow examination of the molecular interactions of **1** in the crystalline state.



1



2

Knowledge of the structure of the coenzyme binding sites of the enzyme and the molecular dimensions of 3,N^4-ethenocytidine should aid in the correlation of the activity of the modified cytidine nucleotides and the spatial dimensions of the enzymes. In order to facilitate comparisons with related cytidine compounds, the conventional atom numbering for cytidine, as shown in **2**, is used in this paper. This paper is the second in a series on the structure of chloroacetaldehyde-modified nucleosides. Previously, a derivative of ϵ-adenosine was reported.[14]

Experimental Section

ϵ-Cytidine hydrochloride[8] was recrystallized from ethanol-water by standing at room temperature. The crystals are colorless, transparent, elongated, rectangular-shaped plates.

Crystal data: $C_{11}H_{14}N_3O_5^+\cdot Cl^-$; mol wt 303.7; monoclinic; a = 6.330 (1), b = 16.968 (2), c = 6.695 (1) Å; β = 115° 22 (1)′; V = 649.8 Å^3; D_{measd} = 1.54 g cm^{-3}; Z = 2; D_{calcd} = 1.55 g cm^{-3}; $F(000)$ = 316; μ (Mo Kα) = 3.2 cm^{-1}; systematic absences for $0k0$ when k = 2 n + 1; the space group is either $P2_1$ or $P2_1/m$. Since the compound is optically active, the former must be the correct space group. The density was measured by flotation in a mixture of hexane and carbon tetrachloride and the cell dimensions were obtained by a least-squares fit to the hand-centered settings for 12 reflections on a Picker FACS-1 diffractometer (Mo Kα_1, λ = 0.70926 Å).

A crystal with dimensions ca. 0.40 × 0.30 × 0.25 mm was used for data collection. The general procedures for data collection were essentially the same as described previously,[15] except that Nb-filtered Mo K$\bar{\alpha}$ (λ = 0.7107 Å) radiation was used. The octants of data, hkl and $\bar{h}kl$, were measured to 2θ = 70° (sin θ/λ = 0.807). No evidence for crystal deterioration or loss of intensity was noted. Out of 2961 possible independent reflections, 1675 were considered to be nonzero, using a 3σ criterion based on counting statistics. The data were corrected for Lorentz and polarization effects but not for absorption; the maximum and minimum transmission factors (based on intensities) were estimated to be 0.86 and 0.90.

The structure was solved by the heavy atom method based on chlorine. The hydrogen atoms were located from a difference map. Full-matrix, least-squares refinement of positional and anisotropic thermal parameters for the nonhydrogen atoms and of positional and isotropic thermal parameters for the hydrogen atoms converged with

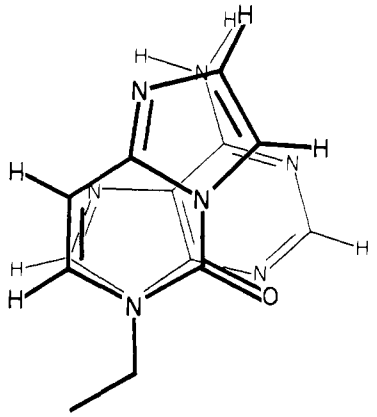


Figure 1. Overlay of the line representations of the N(9)-substituted adenine and N(1)-substituted ϵ-cytidine showing similar polarities at various points. Drawing similar to one depicted in ref 78.

values for R and R_2, $R_2 = [\Sigma w(|F_o| - |F_c|)^2/\Sigma w|F_o|^2]^{1/2}$, of 0.045 and 0.037, respectively. The final value of $\Sigma w\Delta^2/(NO - NV)$, where NO is the number of observations and NV is the number of variables, is 1.55. All reflections were assigned weights using a program written by Dieterich[16] using the scheme proposed by Corfield et al.[17] A final difference map showed a negative peak of 1 electron/Å^3 near the center of the position occupied by the chloride ion.[18] There were also some peaks and troughs with electron densities of 0.2–0.3 electron/Å^3. Most of these peaks lie close to the midpoint between two atoms which are covalently bonded and may be due at least in part to the effect of the valence electrons. A final structure-factor calculation showed no significant abnormality for any individual reflection. The final values for the atomic coordinates and thermal parameters are listed in Tables I and II.[19]

The scattering curves for Cl^-, C, N, and O used in the analysis were taken from the compilation by Cromer and Mann;[20] that for hydrogen was the one calculated by Stewart et al.[21] The curve for Cl^- was corrected for the effects of anomalous dispersion.[22]

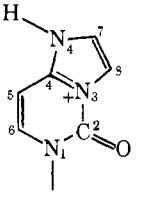
Results and Discussion

Molecular Dimensions of the ϵ-Cytosine Ring. A stereoscopic view of the ϵ-CydH$^+$ cation is shown in Figure 2. The bond lengths and angles involving the nonhydrogen atoms are given in Figure 3. The distances and angles involving hydrogen atoms have been deposited. An analysis[23] of thermal vibrations of the ϵ-cytosine ring system showed that it could be treated as a moderately good rigid body (root mean square Δ of $U(i,j)$ is 0.0019 Å^2) and resulted in adjustments to the bond lengths of a maximum of 0.006 Å for N(4)–C(7) and an average change of 0.003 Å. In subsequent discussions, the unadjusted values are used. The cytosine ring system is bridged between N(3) and N(4) by the etheno group; N(4) has an attached hydrogen atom. This is the first example of an x-ray analysis on a cytosine ring with an ϵ bridge. It is well established[24] that substitution at nitrogen has a definite effect on the lengths and angles in pyrimidine rings. In 1971, Viswamitra et al.[25] compiled the bond lengths and angles found in x-ray structures of unsubstituted cytosine bases, neutral N(1)-substituted cytosine bases, and in protonated N(1)-substituted cytosine derivatives. We have brought this compilation up to date (Dec 1975),[26] and we compare the dimensions found for the above three types with those found for the ϵ-cytosine ring system in Figure 4. In the earlier comparison the greatest differences between the protonated [Figure 4 (c)] and neutral species [Figure 4(a) and (b)] were in the N(4)–C(4)–N(3)–C(2)–O(2) region of the heterocyclic base. With the new and considerably expanded structural data, the differences are generally similar to those described before. It should be pointed out that some of the dimensions, particularly those involving peripheral atoms, could be significantly influenced by hydrogen bonding. Thus, in addition to experimental errors, one could anticipate some differences within a particular class. The complete list of data including maximum and minimum values for each dimension is included in the microfilm edition of the Journal. The agreement between the values found for the ϵ-CydH$^+$ cation [Figure 4 (d)] and in N(1)-substituted and protonated cytosine structures (Fig. 4c) is quite good, with the greatest differences involving the C(2)–N(3) bond (longer in ϵ-CydH$^+$), the C(4)–C(5) bond (shorter), the C(4)–N(4) bond (longer), and the N(1)–C(6) bond (longer). In both protonated species [Figure 4(c) and (d)], the C(2)–O(2) distance is less than in the uncharged forms [Figure 4(a) and (b)]. The inferred decrease in this distance upon protonation in the ϵ-Cyd series is accompanied by an increase in the C–O stretching frequencies in the infrared spectra.[53] There are also significant changes in the internal ring bond angles at N(1), C(2), C(4), and C(5). The two bond angles external to the six-membered ring at C(4) are greatly different from each other [133.1 (4) and 106.8 (3)°] in the ϵ-CydH$^+$ cation. The two main contributors to the resonance hybrid of the ϵ-CydH$^+$ cation would appear to be **3** and **4**.

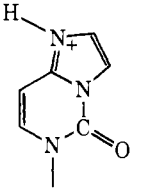
Table I. Final Atomic Coordinates in the ϵ-Cyd·HCl Structure. Standard Deviations in Parentheses

Atoms	x	y	z
Cl	0.21124 (16)	0.50000[a]	0.11675 (16)
N(1)	1.1163 (5)	0.2967 (2)	0.1790 (4)
C(2)	1.2071 (6)	0.3355 (2)	0.3787 (5)
O(2)	1.4072 (4)	0.3319 (2)	0.5167 (4)
N(3)	1.0383 (5)	0.3799 (2)	0.4129 (5)
C(4)	0.8137 (6)	0.3888 (2)	0.2629 (6)
N(4)	0.7130 (6)	0.4377 (2)	0.3521 (6)
C(5)	0.7317 (6)	0.3488 (2)	0.0621 (6)
C(6)	0.8859 (6)	0.3030 (2)	0.0259 (6)
C(7)	0.8738 (8)	0.4600 (3)	0.5596 (8)
C(8)	1.0749 (8)	0.4247 (2)	0.5996 (7)
C(1′)	1.2751 (6)	0.2456 (2)	0.1260 (6)
C(2′)	1.2405 (6)	0.1582 (2)	0.1498 (5)
C(3′)	1.3328 (6)	0.1244 (2)	−0.0081 (6)
C(4′)	1.2442 (7)	0.1844 (2)	−0.1973 (6)
C(5′)	1.0129 (7)	0.1671 (3)	−0.3908 (6)
O(1′)	1.2294 (4)	0.2591 (2)	−0.0954 (4)
O(2′)	1.3455 (5)	0.1300 (2)	0.3669 (4)
O(3′)	1.5805 (5)	0.1241 (2)	0.1060 (5)
O(5′)	0.8226 (5)	0.1587 (2)	−0.3327 (5)
NH(4)	0.556 (9)	0.447 (3)	0.268 (7)
H(5)	0.585 (6)	0.350 (2)	−0.026 (6)
H(6)	0.863 (8)	0.263 (3)	−0.106 (8)
H(7)	0.836 (7)	0.492 (3)	0.648 (6)
H(8)	1.207 (9)	0.424 (3)	0.736 (8)
H(1′)	1.423 (6)	0.258 (2)	0.231 (6)
H(2′)	1.069 (5)	0.145 (2)	0.082 (5)
H(3′)	1.275 (6)	0.074 (2)	−0.048 (5)
H(4′)	1.353 (6)	0.192 (2)	−0.252 (6)
H(5′A)	0.974 (7)	0.220 (3)	−0.478 (7)
H(5′B)	1.034 (7)	0.125 (3)	−0.451 (7)
OH(2′)	1.463 (8)	0.137 (3)	0.402 (8)
OH(3′)	1.617 (7)	0.093 (3)	0.028 (7)
OH(5′)	0.834 (9)	0.126 (3)	−0.261 (9)

[a] This coordinate was held constant in the refinement to determine the origin in the y direction.



3



4

Table II. Final Thermal Parameters, Expressed in the Form $\exp[-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)]$ and Isotropic Thermal Parameters $\exp[-(B_{\text{iso}} \sin^2\theta/\lambda^2)]$

	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{33}
Cl	0.0208 (3)	0.00212 (3)	0.0217 (3)	0.00053 (9)	0.0103 (2)	0.00045 (9)
N(1)	0.0139 (8)	0.0015 (1)	0.0142 (8)	−0.0000 (2)	0.0052 (7)	−0.0010 (2)
C(2)	0.0171 (11)	0.0017 (1)	0.0131 (9)	−0.0003 (3)	0.0064 (9)	−0.0004 (3)
O(2)	0.0196 (8)	0.0029 (1)	0.0155 (7)	0.0013 (3)	0.0039 (6)	−0.0013 (2)
N(3)	0.0167 (9)	0.0017 (1)	0.0138 (8)	0.0007 (3)	0.0060 (7)	0.0002 (2)
C(4)	0.0152 (10)	0.0017 (1)	0.0208 (10)	−0.0002 (3)	0.0090 (9)	0.0007 (3)
N(4)	0.0205 (10)	0.0022 (1)	0.0250 (11)	0.0017 (3)	0.0130 (10)	0.0003 (3)
C(5)	0.0109 (11)	0.0027 (2)	0.0191 (11)	0.0002 (3)	0.0026 (9)	−0.0001 (3)
C(6)	0.0158 (10)	0.0022 (1)	0.0180 (10)	−0.0011 (3)	0.0062 (9)	−0.0012 (3)
C(7)	0.0318 (16)	0.0025 (2)	0.0224 (13)	0.0019 (4)	0.0152 (12)	−0.0007 (4)
C(8)	0.0274 (14)	0.0026 (2)	0.0161 (11)	0.0022 (4)	0.0082 (11)	−0.0008 (3)
C(1′)	0.0134 (10)	0.0019 (1)	0.0132 (9)	−0.0004 (3)	0.0067 (8)	−0.0007 (3)
C(2′)	0.0149 (11)	0.0018 (1)	0.0147 (10)	−0.0003 (3)	0.0073 (8)	−0.0005 (3)
C(3′)	0.0165 (10)	0.0020 (1)	0.0148 (10)	−0.0003 (3)	0.0067 (9)	−0.0012 (3)
C(4′)	0.0171 (12)	0.0026 (2)	0.0153 (10)	−0.0010 (3)	0.0093 (9)	−0.0009 (3)
C(5′)	0.0235 (14)	0.0027 (2)	0.0156 (11)	−0.0016 (4)	0.0093 (10)	−0.0008 (4)
O(1′)	0.0228 (9)	0.00189 (9)	0.0157 (7)	−0.0009 (2)	0.0108 (6)	−0.0002 (2)
O(2′)	0.0157 (9)	0.0029 (1)	0.0147 (7)	−0.0003 (3)	0.0063 (7)	0.0008 (2)
O(3′)	0.0166 (8)	0.0043 (1)	0.0216 (9)	0.0016 (3)	0.0072 (7)	−0.0029 (3)
O(5′)	0.0182 (9)	0.0027 (1)	0.0173 (8)	−0.0002 (3)	0.0072 (7)	0.0009 (3)

	B_{iso}(Å^2)		B_{iso}(Å^2)
H(5)	2.8 (9)	OH(2′)	3.9 (1.3)
H(6)	5.2 (1.1)	H(3′)	2.1 (7)
H(7)	3.7 (9)	OH(3′)	4.0 (1.0)
H(8)	5.4 (1.2)	H(4′)	1.6 (7)
NH(4)	4.4 (1.1)	H(5′A)	4.2 (1.0)
H(1′)	1.9 (7)	H(5′B)	2.8 (9)
H(2′)	1.1 (6)	OH(5′)	4.9 (1.5)

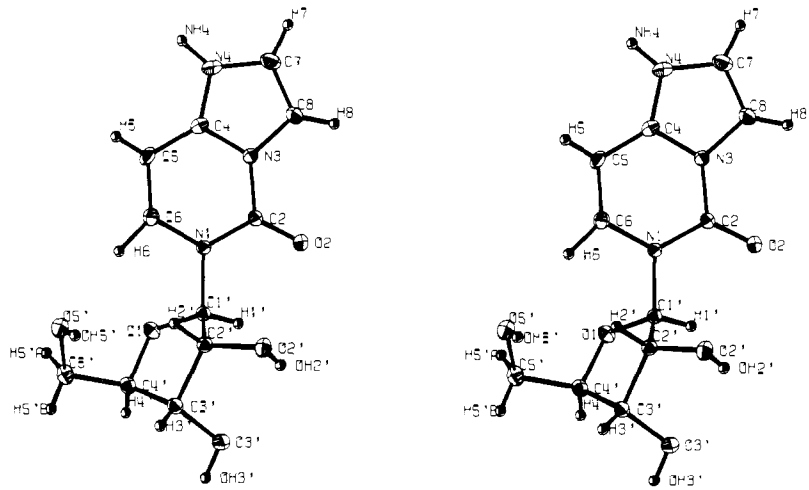


Figure 2. Stereoscopic view of the ϵ-Cyd cation. The percentage probability of the ellipsoids is 20%.

The ϵ-cytosine moiety is slightly, but significantly, nonplanar (Table III), with a maximum deviation of 0.037 Å from the best plane through the ten atoms which were included in the plane calculation. The C(1′) atom is practically coplanar with the ϵ-cytosine ring, as was the case in the ϵ-adenosine derivative.[14] The pyrimidine ring is slightly nonplanar with atoms deviating from the plane by distances of −0.020 to +0.018 Å. The five atoms of the etheno ring are coplanar as was also observed in the study of the ϵ-adenosine cation.[14] The best planes through the two rings of the ϵ-cytosine moiety are inclined at an angle of 1° 27′. The NMR spectra (100 MHz) of both the ϵ-CydH$^+$ cation and its free base in either D_2O or Me_2SO-d_6 clearly show a long-range coupling between H(5) and H(8) with a coupling constant of 0.6 Hz.[53] Coupling constants of this magnitude occur for essentially planar configurations and fall off rapidly with large departures from planarity. From these data, near planarity of the free base of ϵ-Cyd can be inferred.

Molecular Dimensions of the Ribose Ring. The best four-atom plane in the ribose ring contains the atoms C(1′), C(3′), C(4′), and O(1′) with C(2′) lying 0.594 Å on the same side as C(5′) (Table III). The conformation of the ribose is thus C(2′)-endo, which is one of the common puckering modes for pyrimidine nucleosides and nucleotides.[54-56] When referred to the three-atom plane, C(1′), O(1′), and C(4′), the puckering of the ribose ring is C(2′)-endo-C(3′)-exo ($^2T_3$)[55] with C(2′) and C(3′) lying on the same and opposite sides of the plane as the C(5′) atom at distances of 0.466 and −0.169 Å, respectively. Using the pseudorotation rotation introduced by Altona and Sundaralingam,[57] the values of τ_m and P are −38.7 and 170.5°, respectively. The values for the torsion angles around the ring are O(1′)-C(1′)-C(2′)-C(3′), 36.1°; C(1′)-C(2′)-

Table III. Least-Squares Planes for the Base and Sugar and the Deviations of the Atoms from Planes I–VI in Å[a,b]

Atoms	I	II	III	IV	V	VI
N(1)	*−0.008*	*−0.002*				
C(2)	*0.009*	*0.018*	0.060			
O(2)	*0.018*	0.031				
N(3)	*−0.025*	*−0.020*	*0.001*			
C(4)	*0.011*	*0.012*	*−0.001*			
N(4)	*0.025*	0.025	*0.000*			
C(5)	*0.007*	*0.007*	−0.020			
C(6)	*−0.012*	*−0.010*				
C(7)	*−0.002*	0.001	*0.001*			
C(8)	*−0.037*	−0.030	*−0.002*			
C(1′)	−0.021	−0.012		*−0.190*	*0.027*	*0.000*
C(2′)				*0.223*	0.594	0.466
C(3′)				*−0.222*	*−0.024*	−0.169
C(4′)				*0.126*	*0.044*	*0.000*
O(1′)				*0.028*	*−0.025*	*0.000*
C(5′)				1.474	1.252	1.228
χ^2	257	81.4	0.33	9344	246.8	0.00
P[c]	<0.005	<0.005	0.80	<0.005	<0.005	

[a] The distances italicized are for the atoms included in the best plane calculations. [b] In these calculations, the atoms were weighted as $1/\sigma^2$, where σ is the standard deviation from the least-squares results. [c] The probability (on the basis of the χ^2 test) that the deviations of the atoms from the plane form a normal distribution.

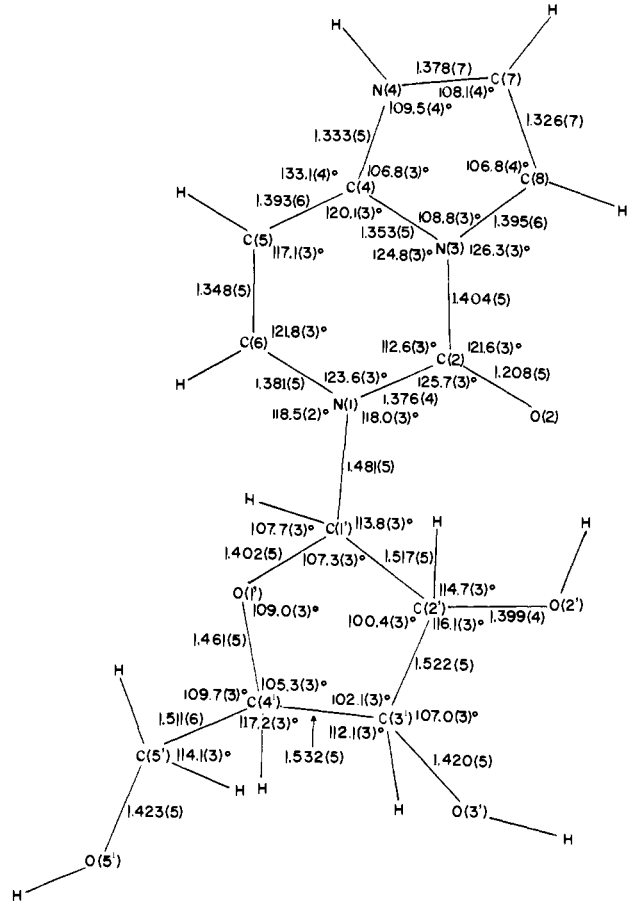


Figure 3. Bond distances (Å) and angles (degree) involving nonhydrogen atoms in the ϵ-cytidine hydrochloride. The range of C–H distances is 0.86 (5)–1.07 (5) Å and of OH distances is 0.69 (6)–0.84 (5) Å. The full list of distances and angles involving hydrogen atoms has been deposited.

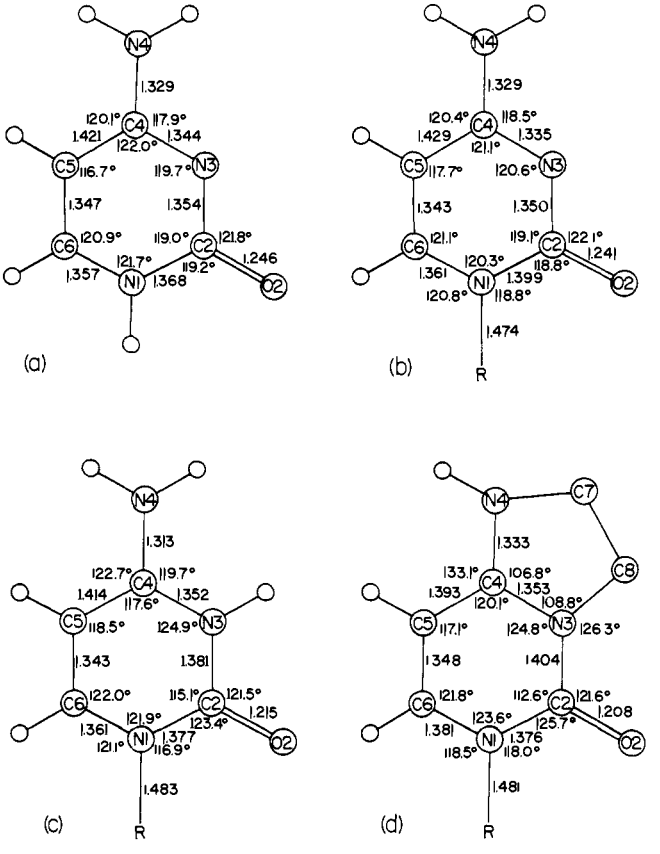


Figure 4. (a) Molecular dimensions averaged over six neutral and unsubstituted at N(1) cytosine derivatives.[26] (b) Dimensions averaged over 11 neutral but N(1)-substituted cytosine derivatives. (c) Dimensions averaged over 12 N(3)-protonated cytosine derivatives. (d) Dimensions in the ϵ-cytosine ring in **2**.

C(3′)-C(4′), −38.2; C(2′)-C(3′)-C(4′)-O(1′), 28.7°; C(3′)-C(4′)-O(1′)-C(1′), −6.6°; C(4′)-O(1′)-C(1′)-C(2′), −18.8°; using the definition that the angle A-B-C-D is positive if, when looking along the B-C bond, atom A has to be rotated in a clockwise direction to eclipse atom D. The O(2′) atom is in an equatorial position, while the atoms O(3′) and C(5′) are axial.

The C(4′)-O(1′) bond [1.461 (5) Å] in the ϵ-CydH$^+$ cation is considerably longer than the corresponding bond in the ϵ-Ado derivative [1.437 (6) Å].[14]

The torsion angles about the C(4′)-C(5′) exocyclic bond show that the side chain assumes a gauche-gauche confor-

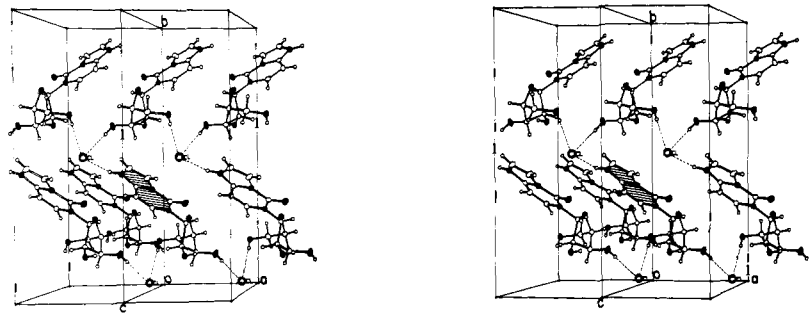


Figure 5. Stereoscopic view of the packing in the crystal of **2** looking approximately along the *c* axis. The ϵ-cytosine group in the reference molecule is shaded.

Table IV. Molecular Dimensions[a] Relevant to the Hydrogen Bonding Scheme and Short Contacts (Distances in Å, Angles in Degrees). D Refers to Donor and A to Acceptor

Hydrogen bond D—H- - -A	Distances D- - -A	H- - -A	Angle D—H- - -A
N(4)—NH(4)- - -Cl^{I}	3.067 (4)	2.17 (6)	164 (5)
O(3′)—OH(3′)- - -Cl^{II}	3.175 (3)	2.35 (4)	168 (4)
O(5′)—OH(5′)- - -Cl^{III}	3.107 (3)	2.40 (6)	163 (4)
O(2′)—OH(2′)- - -O(5′)^{IV}	2.863 (4)	2.24 (6)	152 (5)
C(6)—H(6)- - -O(5′)^{I}	3.331 (5)	2.27 (5)	168 (3)

[a] I, x,y,z; II, $2-x,-\frac{1}{2}+y,-z$; III, $1-x,-\frac{1}{2}+y,-z$; IV, $1+x,y,1+z$.

mation with $\phi_{O_1'-O_5'} = -61.6°$ and $\phi_{C_3'-O_5'} = 58.4°$,[58] respectively. This conformation has been predicted to be the most stable one for purine and pyrimidine nucleosides and nucleotides, irrespective of the puckering of the sugar.[59,60] The fact that the ϕ_{O-O} and ϕ_{C-O} angles are nearly of equal magnitude but of opposite sign has the effect of bringing O(5′) into a position where it can accept the intramolecular C—H- - -O hydrogen bond (see below).

The arrangement of the present structure appears to be further stabilized by an intramolecular hydrogen bond, C(6)—H(6)- - -O(5′), with an H(6)- - -O(5′) distance of 2.27 (5) Å, a C(6)-H(6) distance of 1.07 (5) Å, and a C(6)—H(6)- - -O(5′) angle of 168 (4)°. It is of interest to note that the H(6) atom is bent toward the O(5′) atom; the C(5)-C(6)-H(6) and N(1)-C(6)-H(6) angles are 131 (3) and 107 (3)°, respectively. Similar types of intramolecular hydrogen bonds have been observed in all four structurally independent nucleotide units in the crystals of uridylyl(3′→5′)adenosine hemihydrate,[61,62] and their role in contributing to the stability of the structures of nucleic acids has been suggested. However, in the structure of adenylyl(3′→5′)uridine,[63] this type of bonding occurred in only one of the adenosine residues.

Molecular Conformation of the ϵ-Cyd Cation. The glycosyl torsion angle, χ_{CN} [τ[O(1′)-C(1′)-N(1)-C(6)]], which describes the relative orientation of the base with respect to the sugar is 42.6° in the structure of ϵ-Cyd·HCl (see Figure 2).[64] Thus, the conformation about the glycosyl bond is anti. This glycosyl torsion angle is in the range frequently found in many pyrimidine nucleosides with C(2′)-endo puckering of the sugar.[55,56]

Crystal Structure. A stereoview of the crystal structure of ϵ-Cyd·HCl is shown in Figure 5. Layers of a nonpolar or hydrophobic nature alternate with those of a more polar nature (consisting of the sugar rings and the anions) in the *b* direction. These nonpolar and polar layers extend infinitely in the *a* and *c* directions. Somewhat similar structures have also been observed in the crystals of 2′-deoxycytidine hydrochloride,[43] 1-(β-D-arabinofuranosyl)cytosine hydrochloride,[47] and adenosine hydrochloride.[65]

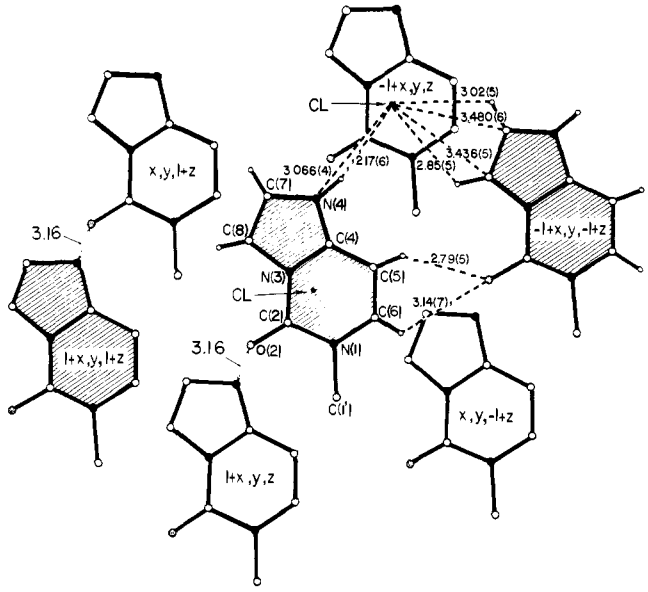


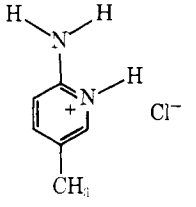
Figure 6. The projection of six ϵ-cytosine rings and the chloride ion onto the best plane through the reference molecule (the central shaded molecule). Two of the ϵ-cytosine rings that also lie in the plane of the reference molecule are also shaded. Chloride ion is shown as a small black dot which is 3.23 Å above the ring. The near coplanar arrangement of ϵ-cytosine rings and Cl^- ions that is found in many cytosine hydrochloride salts[47] is illustrated by the reference molecule and the ϵ-cytosine ring to its right.

Hydrogen Bonding. Compared to protonated cytosine nucleosides, the ϵ-cytosine has lost considerable hydrogen bonding capability. The only hydrogen bond involving the atoms in the ϵ-cytosine cation (with the exception of the intramolecular C—H- - -O hydrogen bond mentioned earlier) is one of length of 3.067 (4) Å, involving the N(4) proton to the chloride ion, Cl^-, at x,y,z (Table IV). As is also the case for many protonated cytidine derivatives,[43-45,47] the carbonyl atom, O(2), does not participate in hydrogen bonding. However, the O(2) atom in the structure of deoxycytidine 5′-phosphate monohydrate[25] was hydrogen bonded to a water molecule. The chloride ion is an acceptor for three hydrogen bonds and thus serves to link three ϵ-cytidine molecules. The donor atoms are N(4) in the molecule at x,y,z, O(3′) in the molecule at $2-x,\frac{1}{2}+y,-z$, and O(5′) in the molecule at $1-x,\frac{1}{2}+y,-z$.

All the hydroxyl groups in the sugar are involved in hydrogen bonding. The O(5′) atom is an acceptor for two hydrogen bonds, with one being an intramolecular C(6)—H(6)- - -O(5′) hydrogen bond and the other being the hydrogen bond, O(2′)—OH(2′)- - -O(5′)^{IV}, which is the only one connecting the two ϵ-cytidine molecules directly.

The chloride ions lie close to the planes of the ϵ-cytosine rings (Figure 6), apparently interacting with adjacent ϵ rings. This type of "in-plane" interaction, as pointed out by Sherfinski and Marsh,[47] appears to be characteristic of the hydrochloride salts

of cytosine derivatives and is found here despite the presence of the etheno bridge in the ϵ-Cyd·HCl. The two hydrogen atoms at the 3,N^4-etheno bridge replace those of N(3) and N(4) in other cytosine derivatives, although in the ϵ-cytidine structure, the distances between the hydrogen atoms and the chloride ion are longer [2.85 (5) Å and 3.02 (6) Å] than those found in other cytosine derivatives (2.22–2.49 Å and 2.74–3.05 Å), where they were presumed to be hydrogen bonds. The O(2) atom in the molecule at $-1 + x,y,-1 + z$ approaches the C(5)–H and C(6)–H atoms with O(2)- - -H distances of 2.79 (5) and 3.14 (7) Å, respectively (Figure 6). The corresponding values found in other cytosine derivatives range from 2.51 to 2.84 Å for H(5)- - -O(2) and 2.60 to 2.94 Å for H(6)- - -O(2) distances.[47] A similar arrangement has also been observed in 1-methyluracil hydrobromide.[66] In an interesting comparison, Sherfinski and Marsh[67] studied the crystal structure of 2-amino-5-methylpyridine hydrochloride (**5**) [which can be considered somewhat similar to cytosine hydrochloride with O(2) replaced by a hydrogen atom] and they found an entirely different type of crystal packing. They therefore concluded that the presence of carbonyl oxygen atom is essential for this type of interaction between two cytosine rings and a chloride anion and that the C—H- - -O interactions found in the cytosine structures are a definite stabilizing influence.



5

Stacking. In contrast to the extensive stacking in the crystal of the ϵ-adenosine derivative,[14] the crystal of ϵ-Cyd·HCl shows no direct overlap between the bases. Figure 6 also depicts the relation between the reference ϵ-cytosine cation and the six closest neighboring bases and the nearest chloride ion. The shortest interplanar distance is 3.164 (5) Å between N(4) and O(2) in the molecule at $-1 + x,y,z$.

The Cl^- ion at $1 + x,y,z$ lies immediately above the plane through the pyrimidine ring of the ϵ-cytosine cation at a distance of 3.21 Å, a value which is significantly shorter than the sum of the ionic radius of the chloride ion (1.81 Å)[68] and the half thickness of an aromatic ring (1.70 Å).[69] The N(3)- - -Cl^- and the C(2)- - -Cl^- distances are 3.338 (3) and 3.303 (4) Å, respectively. The Cl^- ion, sandwiched between ϵ-cytosine rings as described previously, occupies the position at which other ϵ-cytosine cations could have stacked, thus precluding the possibility of overlap between bases.

Table V lists some examples of anhydrous salts of nucleic acid bases and nucleosides where there is little or no overlap between the rings of the bases in the solid state. They all show the common features of direct interaction between the negatively charged anions (e.g., Cl^-, Br^-, I^-) and the positively charged heterocyclic rings. These anions lie above or nearly above the rings of the bases in the crystal. The interaction between the anions and the base cation is presumably due to ionic forces. This type of "ion-pair" interaction between the heterocyclic ring and anion has also been observed frequently in the crystal structures of flavin derivatives, such as riboflavin hydrobromide monohydrate,[70] 1,3,10-trimethylisoalloxazinium iodide,[71] 1,10-ethylene-7,8-dimethylisoalloxazinium iodide monohydrate,[72] and 10-methylisoalloxazine hydrobromide dihydrate.[73] A similar case was also found in the structure of *N*-methylacridinium iodide,[74] which has been suggested as an organic charge-transfer salt. This latter possibility was also commented upon in a preliminary communication on the structure of 1,7-dimethylguanosine iodide.[75]

Table V. Halide Ion to Plane of Heterocyclic Base Distances in Some Anhydrous Nucleic Acid Base and Nucleoside Structures

Structure	Distance, Å
Nucleic Acid Bases and Related Compounds	
9-Methylguanine hydrobromide[a]	3.74
1-Methyluracil hydrobromide[b]	3.41
1-Methylcytosine hydrobromide[c]	3.40
1-Methylcytosine hydrochloride[d]	3.35 (molecule A)
	3.54 (molecule B)
9-β-Chloroethyl-7,8-dihydro-9*H*-imidazo[2,1-*i*]-purine hydroiodide[e]	3.52
2-Amino-7-methyl-5-propyl-*s*-triazolo[2,3-*c*]-pyrimidine hydrochloride[f]	3.14
Nucleosides and Related Compounds	
ϵ-Cytidine hydrochloride[g]	3.23
2′-Deoxycytidine hydrochloride[h]	3.37
1-(β-D-Arabinofuranosyl)cytosine hydrochloride[i]	3.40
Adenosine hydrochloride[j]	3.28
Guanosine hydrobromide hemihydrate[k,l]	3.39
	3.55
	−3.80
5′-Methylammonium-5′-deoxyadenosine iodide hydrate[l,m]	3.82
Aristeromycin hydrobromide[n]	3.59
	−3.63

[a] H. M. Sobell and K. Tomita, *Acta Crystallogr.*, **17**, 126 (1964). [b] H. M. Sobell and K. Tomita, *ibid.*, **17**, 122 (1964). [c] Reference 52. [d] Reference 46; there are two molecules in the crystallographic asymmetric unit. [e] W. M. MacIntyre and R. F. Zahrobsky, *Z. Kristallogr., Kristallgeom. Kristallphys. Kristallchem.*, **119**, 226 (1963). [f] P. G. Owston and J. M. Rowe, *Acta Crystallogr.*, **15**, 231 (1962). [g] Present work. [h] Reference 43. [i] Reference 47. [j] K. Shikata, T. Ueki, and T. Mitsui, *Acta Crystallogr., Sect. B*, **29**, 31 (1973). [k] P. Tougard, *Jerusalem Symp. Quantum Chem. Biochem.* **4**, 217 (1972); P. Tougard and J.-F. Chantot, *Acta Crystallogr., Sect. B*, **30**, 214 (1974). There are two crystallographically independent guanosine molecules in the crystallographic asymmetric unit. For one ring there is a Br- - -ring distance of 3.39 Å, while for the other there are bromine atoms disposed on opposite sides at 3.55 and 3.80 Å. [l] It should be emphasized that this molecule has some water in the crystal. [m] W. Saenger, *J. Am. Chem. Soc.*, **93**, 3035 (1971). [n] T. Kishi, M. Muroi, T. Kusaka, M. Nishikawa, K. Kamiya, and K. Mizuno, *Chem. Pharm. Bull.*, **20**, 940 (1972).

The type of anion–nucleoside interaction found here and for other *anhydrous hydrohalide* salts of nucleic acid bases and nucleosides is in marked contrast to the stacking of heterocyclic bases found in the ϵ-adenosine derivative[14] and in many other hydrated salts of nucleic acid bases and nucleosides.[76] When such compounds are crystallized with water molecules in the unit cell, the anions tend to interact with the sugar residue and with the water molecules by means of polar interactions such as ion–dipole forces and hydrogen bonding and so create a largely polar region in the crystal. The heterocyclic bases are effectively excluded from such a region and, as a result of the classic "hydrophobic" effect, form effective stacks of near-planar residues. *Thus the presence of water molecules in the crystal often has a decisive influence on whether the heterocyclic bases are stacked or not.*

It has been noted that halogen substituents on the purine and pyrimidine rings affect the solid-state stacking patterns of nucleic acid constituents.[76,77] In general, the presence of a halogen substituent results in a stacking pattern which permits close contact between the halogen atom and adjacent purine or pyrimidine rings.

The physical and biological properties of 3,N^4-ethenocytidine and related nucleotides hold considerable interest since differences in their behavior in relation to the corresponding naturally occurring cytosine or adenine compounds might be related to defined geometrical changes. On this basis the possible interaction of structural models of ϵ-NCD$^+$, an analogue of NAD$^+$, with Kendrew models of dogfish lactate dehydrogenase and lobster muscle glyceraldehyde 3-phosphate dehydrogenase has been examined by Greenfield et al.[78] In both enzymes the replacement of the adenine of the bound NAD$^+$ with ϵ-cytosine requires no apparent change in the coenzyme binding sites.

This similarity is also reflected in the dinucleoside phosphates of the general formula 3,N^4-ethenocytidyl(3′→5′) nucleoside (ϵ-CpN) which are not hydrolyzed by pancreatic RNase A.[79] This is consistent with the previous observations that substitution of a bulky function at N-3 of the pyrimidine moiety renders normal substrates resistant to RNase A[80] and with the results of Kochetkov and co-workers with ϵ-CpU.[81] The high degree of resistance of ϵ-CpN phosphodiester bonds to RNase A action can actually be used with advantage in sequencing chloroacetaldehyde-treated oligonucleotides and tRNA.

In conclusion, the molecular dimensions and conformation of the fluorescent ϵ-cytidine hydrochloride have been determined and analyzed in comparison with other cytosine derivatives. The nature of the interactions of the ϵ-cytosine base with chloride ions, both "in-plane" and "out-of-plane" is described; particular emphasis is placed on the probable role played by water molecules in the crystal structures of hydrohalide salts of heterocyclic bases and nucleosides.

Acknowledgment. We thank Professor Nelson J. Leonard for suggesting the ϵ-Cyd derivative as an attractive molecule for an x-ray study, for useful discussions during the course of the work, and for his generous and discerning comments upon the manuscript. Dr. Kwo-Tsair Wei and Miss Mary K. Greensley assisted with several calculations; Miss Jennifer Jones and Mrs. Patti Eckert made several of the drawings. The work was supported by grants from NIH (GM 19336) and also by Dr. Leonard's Grant GM 05829.

Supplementary Material Available: listing of h,k,l, $|F_{obsd}|$, $|F_{calcd}|$, α, hydrogen atom bond lengths and angles, and the complete compilation of the dimensions in the various types of cytosine derivatives (20 pages). Ordering information is given on any current masthead page.

References and Notes

(1) H. Sund in "Biological Oxidations", T. P. Singer, Ed., Interscience, New York, N.Y., 1968, p 641.
(2) N. O. Kaplan, M. M. Ciotti, M. Hamolsky, and R. E. Bieber, *Science*, **131**, 392 (1960).
(3) N. O. Kaplan, *Proc. Int. Congr. Biochem., 5th*, **3**, 97 (1963).
(4) S. T. Rao and M. G. Rossmann, *J. Mol. Biol.*, **76**, 241 (1973).
(5) A. Liljas and M. G. Rossmann, *Annu. Rev. Biochem.*, **43**, 475 (1974).
(6) A. S. Mildvan, *Annu. Rev. Biochem.*, **43**, 357 (1974).
(7) N. K. Kochetkov, V. N. Shibaev, and A. A. Kost, *Tetrahedron Lett.*, 1993 (1971).
(8) J. R. Barrio, J. A. Secrist III, and N. J. Leonard, *Biochem. Biophys. Res. Commun.*, **46**, 597 (1972).
(9) J. R. Barrio, L. G. Dammann, L. H. Kirkegaard, R. L. Switzer, and N. J. Leonard, *J. Am. Chem. Soc.*, **95**, 961 (1973).
(10) A. P. Razjivin, A. A. Kost, and V. N. Shibaev, *Stud. Biophys.*, **51**, 29 (1975).
(11) Y. Shahak, D. M. Chipman, and N. Shavit, *Proc. Int. Congr. Photosynth., 3rd*, 859 (1974).
(12) C. Ivancsics and E. Zbiral, *Monatsh. Chem.*, **106**, 417 (1975).
(13) Abbreviations following the IUPAC–IUB Commission on Biochemical Nomenclature recommendations [*J. Mol. Biol.*, **55**, 299 (1971)] are used throughout. The abbreviation "ϵ" stands for etheno, so that ϵ-Cyd is 3,N^4-ethenocytidine or 5,6-dihydro-5-oxo-6-β-D-ribofuranosylimidazo[1,2-*c*]pyrimidine (see ref 8); ϵ-NCD$^+$, nicotinamide 3,N^4-ethenocytosine dinucleotide.
(14) A. H.-J. Wang, L. G. Dammann, J. R. Barrio, and I. C. Paul, *J. Am. Chem. Soc.*, **96**, 1205 (1974).
(15) J. K. Frank and I. C. Paul, *J. Am. Chem. Soc.*, **95**, 2324 (1973); R. S. Miller, I. C. Paul, and D. Y. Curtin, *ibid.*, **96**, 6334 (1974).
(16) D. A. Dieterich, Ph.D. Thesis, University of Illinois, 1973.
(17) P. W. R. Corfield, R. J. Doedens, and J. A. Ibers, *Inorg. Chem.*, 6, 197 (1967).
(18) We have been unable to find a satisfactory explanation for this anomalous negative peak. One possibility that we considered is that approximately 5% of the Cl$^-$ ions in the crystal are randomly disordered throughout various regions of the unit cell. However, no evidence for such disorder was obtained from our analysis. Another possibility that we considered was that we had carried out an x-ray study on a cocrystallite of ϵ-Cyd·HCl and ϵ-Cyd in the ratio of ~20:1. However, microanalytical data do not support this hypothesis, although the possibility remains that some reaction took place in the crystal.
(19) See paragraph at end of paper regarding supplementary material.
(20) D. T. Cromer and J. B. Mann, *Acta Crystallogr., Sect. A*, **24**, 321 (1968).
(21) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, **42**, 3175 (1965).
(22) "International Tables for X-Ray Crystallography", Vol. III, C. H. MacGillavry and G. D. Rieck, Ed., Kynoch Press, Birmingham, 1962, pp 213–216.
(23) V. Schomaker and K. N. Trueblood, *Acta Crystallogr., Sect. B*, **24**, 63 (1968).
(24) See, for example, M. Sundaralingam and L. H. Jensen, *J. Mol. Biol.*, **13**, 930 (1965).
(25) M. A. Viswamitra, B. S. Reddy, G. H.-Y. Lin, and M. Sundaralingam, *J. Am. Chem. Soc.*, **93**, 4565 (1971).
(26) In the unprotonated and unsubstituted cytosine class of compounds, we included the results from cytosine,[27] cytosine monohydrate,[27] the neutral cytosine ring in the 2:1 complex of cytosine and resorcylic acid monohydrate,[28] the two molecules of thiocytosine (omitting the C–S bond from the averages),[29] and cytosine-5-acetic acid.[30] The earlier studies on cytosine[31] and cytosine monohydrate[32] were not included nor was the rather imprecise study of the 1:1 complex of cytosine and 5-fluorouracil.[33] Among unprotonated, but N(1)-substituted cytosine derivatives, we included cytidine,[34] 1-methylcytosine in a 1:1 complex with 9-ethylguanine,[35] 1-methyl-5-fluorocytosine in a 1:1 complex with 9-ethylguanine,[35] 1-methylcytosine in a 1:1 complex with 5-fluorouracil,[36] the two independent molecules of 2′-deoxycytidine,[37] 1-β-D-arabinofuranosylcytosine,[38] 3′-*O*-methyl-1-β-D-arabinofuranosylcytosine,[39] 2-thiocytidine dihydrate (omitting the C–S bond from the compilation),[40] and the two independent molecules in sodium β-cytidine 2′,3′-cyclic phosphate.[41] We did not include the results from the 1:1 complex of 5-bromodeoxycytidine and deoxyguanosine.[42] The group of N(1)-substituted and N(3)-protonated cytosine derivatives that we considered consisted of 2′-deoxycytidine hydrochloride,[43] the two crystalline forms of cytidylic acid b,[44,45] deoxycytidine 5′-phosphate monohydrate,[25] 1-methylcytosine hydrochloride,[46] 1-β-D-arabinofuranosylcytosine hydrochloride,[47] the protonated cytosine salt of palladium tetrachloride,[48] the cytidine in the 1:1 cytidine–N-carbobenzoxyglutamic acid dihydrate complex,[49] the protonated cytosine ring in the 1:1 complex of cytosine and *N*-benzoylglycine monohydrate,[50] the protonated cytosine ring in the 2:1 complex of cytosine and resorcylic acid monohydrate,[28] and 3-methylcytidine methosulfate monohydrate.[51] The structure of 1-methylcytosine hydrobromide was not included.[52] A full list of the compilations including maximum and minimum values for each bond and angle will appear in the microfilm edition.[19]
(27) R. J. McClure and B. M. Craven, *Acta Crystallogr., Sect. B*, **29**, 1234 (1973).
(28) C. Tamura, S. Sato, and T. Hata, *Bull. Chem. Soc. Jpn.*, **46**, 2388 (1973). In this 2:1 complex, one of the cytosine rings is protonated at N(3) while the other is neutral.
(29) S. Furberg and L. H. Jensen, *Acta Crystallogr., Sect. B*, **26**, 1260 (1970).
(30) R. E. Marsh, R. Bierstedt, and E. L. Eichhorn, *Acta Crystallogr.*, **15**, 310 (1962).
(31) D. L. Barker and R. E. Marsh, *Acta Crystallogr.*, **17**, 1581 (1964).
(32) G. A. Jeffrey and Y. Kinoshita, *Acta Crystallogr.*, **16**, 20 (1963).
(33) D. Voet and A. Rich, *J. Am. Chem. Soc.*, **91**, 3069 (1969).
(34) S. Furberg, C. S. Petersen, and Chr. Rømming, *Acta Crystallogr.*, **18**, 313 (1965).
(35) E. J. O'Brien, *Acta Crystallogr.*, **23**, 92 (1967).
(36) S.-H. Kim and A. Rich, *J. Mol. Biol.*, **42**, 87 (1969).
(37) D. W. Young and H. R. Wilson, *Acta Crystallogr., Sect. B*, **31**, 961 (1975).
(38) P. Tougard and O. Lefebvre-Soubeyran, *Acta Crystallogr., Sect. B*, **30**, 86 (1974).
(39) G. I. Birnbaum, E. Darżynkiewicz, and D. Shugar, *J. Am. Chem. Soc.*, **97**, 5904 (1975).
(40) G. H.-Y. Lin, M. Sundaralingam, and S. K. Arora, *J. Am. Chem. Soc.*, **93**, 1235 (1971).
(41) C. L. Coulter, *J. Am. Chem. Soc.*, **95**, 570 (1973).
(42) A. E. V. Haschemeyer and H. M. Sobell, *Acta Crystallogr.*, **19**, 125 (1965).
(43) E. Subramanian and D. J. Hunt, *Acta Crystallogr., Sect. B*, **26**, 303 (1970).
(44) M. Sundaralingam and L. H. Jensen, *J. Mol. Biol.*, **13**, 914 (1965).
(45) C. E. Bugg and R. E. Marsh, *J. Mol. Biol.*, **25**, 67 (1967).
(46) B. L. Trus and R. E. Marsh, *Acta Crystallogr., Sect. B*, **28**, 1834 (1972).
(47) J. S. Sherfinski and R. E. Marsh, *Acta Crystallogr., Sect. B*, **29**, 192 (1973).
(48) B. L. Kindberg and E. L. Amma, *Acta Crystallogr., Sect. B*, **31**, 1492 (1975).
(49) T. Hata, M. Yoshikawa, S. Sato, and C. Tamura, *Acta Crystallogr., Sect. B*, **31**, 312 (1975).
(50) C. Tamura, T. Hata, S. Sato, and N. Sakurai, *Bull. Chem. Soc. Jpn.*, **45**, 3254 (1972).

(51) E. Shefter, S. Singh, T. Brennan, and P. Sackman, *Cryst. Struct. Commun.*, **3**, 209 (1974).
(52) R. F. Bryan and K.-I. Tomita, *Acta Crystallogr.*, **15**, 1174 (1962).
(53) J. R. Barrio, P. D. Sattsangi, B. A. Gruber, L. G. Dammann, and N. J. Leonard, *J. Am. Chem. Soc.*, following paper in this issue.
(54) M. Sundaralingam, *Biopolymers*, **7**, 821 (1969). For more recent compilations, see ref 55 and 56.
(55) M. Sundaralingam, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 417 (1973); see also M. Sundaralingam, *J. Am. Chem. Soc.*, **93**, 6644 (1971).
(56) S. Arnott and D. W. L. Hukins, *Biochem. J.*, **130**, 453 (1972).
(57) C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, **94**, 8205 (1972).
(58) E. Shefter and K. N. Trueblood, *Acta Crystallogr.*, **18**, 1067 (1965). In this paper distinction was not made between the possible signs for the O–C–C–O and C–C–C–O torsion angles. In later work this distinction is sometimes made and sometimes not. While we do not believe great confusion will arise, we here adopt the standard chemical convention that the A–B–C–D torsion angle will be positive if, when viewing along B to C, atom A has to be rotated clockwise to eclipse atom D.
(59) N. Yathindra and M. Sundaralingam, *Biopolymers*, **12**, 297 (1973).
(60) A. Saran, B. Pullman, and D. Perahia, *Biochim. Biophys. Acta*, **287**, 211 (1972).
(61) J. L. Sussman, N. C. Seeman, S.-H. Kim, and H. M. Berman, *J. Mol. Biol.* **66**, 403 (1972).
(62) J. Rubin, T. Brennan, and M. Sundaralingam, *Biochemistry*, **11**, 3112 (1972).
(63) J. M. Rosenberg, N. C. Seeman, J. J. P. Kim, F. L. Suddath, H. B. Nicholas, and A. Rich, *Nature (London)*, **243**, 150 (1973).
(64) A minor cautionary note: we have adopted the convention that the torsion angle A–B–C–D is positive if, when looking along the B–C bond, atom A has to be rotated clockwise to eclipse atom D; thus, the O(1′)–C(1′)–N(1)–C(6) torsion angle is +42.6°. This is the opposite sign for the angle that would be given by the convention described by Sundaralingam.[54] Among cytosine nucleosides, both conventions, the standard chemical one[41,47] and that of Sundaralingam,[44,45] are in use. The O(1′)–C(1′)–N(1)–C(2) angle has also been used[37] to define the relative orientation of sugar and base.
(65) K. Shikata, T. Ueki, and T. Mitsui, *Acta Crystallogr., Sect. B*, **29**, 31 (1973).
(66) H. M. Sobell and K.-I. Tomita, *Acta Crystallogr.*, **17**, 122 (1964).
(67) J. S. Sherfinski and R. E. Marsh, *Acta Crystallogr., Sect. B*, **31**, 1073 (1975).
(68) L. Pauling in "The Nature of the Chemical Bond", 3d ed, Cornell University Press, Ithaca, N.Y. 1960, pp 511–519.
(69) Reference 68, pp 257–264.
(70) N. Tanaka, T. Ashida, Y. Sasada, and M. Kakudo, *Bull. Chem. Soc. Jpn.*, **42**, 1546 (1969).
(71) M. V. Glehn, P. Kierkegaard, R. Norrestam, O Ronnquist, and P.-E. Werner, *Acta Chem. Scand.*, **24**, 3701 (1970).
(72) M. Leijonmarck and P.-E. Werner, *Acta Chem. Scand.*, **24**, 2916 (1970).
(73) B. L. Trus and C. J. Fritchie, Jr., *Acta Crystallogr., Sect. B*, **25**, 1911 (1969).
(74) K. Nakamura, N. Yasuoka, N. Kasai, and H. Mikawa, *Chem. Commun.*, 1135 (1970).
(75) E. Shefter, S. Singh, and P. Sackman, *J. Chem. Soc., Chem. Commun.*, 261 (1974).
(76) C. E. Bugg, J. M. Thomas, M. Sundaralingam, and S. T. Rao, *Biopolymers*, **10**, 175 (1971).
(77) C. E. Bugg and H. E. Sternglanz, *Jerusalem Symp. Quantum Chem. Biochem.* **7**, 473 (1974).
(78) J. C. Greenfield, N. J. Leonard, and R. I. Gumport, *Biochemistry*, **14**, 698 (1975).
(79) G. L. Tolman, J. R. Barrio, and N. J. Leonard, *Biochemistry*, **13**, 4869 (1974).
(80) F. M. Richards and H. W. Wyckoff, *Enzymes*, **4**, 647 (1971).
(81) N. K. Kochetkov, V. N. Shibaev, and A. A. Kost, *Dokl. Akad. Nauk SSSR*, **205**, 100 (1972).

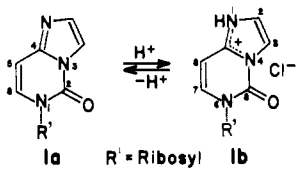
Species Responsible for the Fluorescence of 3,N^4-Ethenocytidine

Jorge R. Barrio, Prem D. Sattsangi, Bruce A. Gruber, Laurence G. Dammann, and Nelson J. Leonard*

Contribution from Roger Adams Laboratory, School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801. *Received July 8, 1976*

Abstract: The fluorescence properties of 3,N^4-ethenocytidine (ε-cytidine), substituted derivatives, and closely related two-ring heterocycles have been examined. The chloroacetaldehyde-modified cytidine is fluorescent only in its protonated form (**1b**). The fluorescence emission maximum is 340 nm and the pK_a* is 4.0, very close to the value for the ground state. N1-alkylation of ε-cytidine at the same position as protonation (**2** and **3**) makes reversion to the nonfluorescent type of structure **1a** impossible, upon changing the pH, and accordingly the fluorescence emission characteristics are preserved over a wide range of pH. The presence of the n → π* transition of the carbonyl group in ε-cytidine is considered to be responsible for the lack of fluorescence in neutral solution. Even ε-cytidine hydrochloride has a low fluorescence quantum yield ($\Phi < 0.01$) and a short fluorescence lifetime ($\tau = 30 \pm 5$ ps). Ring substitution produces a red shift of the π → π* transition, due to inductive or mesomeric effects, and a clear improvement in the fluorescence emission characteristics; e.g., 2-acetylamino-5,6-dihydro-5-oxo-6-β-D-ribofuranosylimidazo[1,2-*c*]pyrimidine (**9**), in its protonated form, shows $\Phi = 0.85$ and $\tau = 4$ ns. Imidazo[1,2-*a*]pyridine, a close model for ε-cytidine lacking the carbonyl group and accordingly the n → π* transition, has a high fluorescence quantum yield and long lifetime in either neutral solution or organic solvents. Both N1-protonated (**10**) and N1-alkylated (**11**) imidazo[1,2-*a*]pyridines have emission maxima similar to that observed for ε-cytidine hydrochloride but higher quantum yields.

It has been shown that the 3,N^4-ethenocytosine (ε-cytosine) moiety can replace *adenine* in nucleotides for enzymatic phosphorylation[1,2] and photophosphorylation[3a] and in dinucleotide derivatives as substrates for certain nucleases.[3b] In addition, the structural analogue of NAD^+, nicotinamide 3,N^4-ethenocytosine dinucleotide (ε-NCD^+),[4] had activity comparable with the natural coenzymes in several enzyme systems including representative dehydrogenases.[5] If the chloroacetaldehyde-modified cytidine (ε-Cyd)[6-8] and related nucleotides[1,2,5] are to be investigated as fluorescent as well as biologically active probes, this requires an understanding of the excited state of ε-Cyd (**1a**) or its conjugate acid (**1b**) because changes in quantum efficiencies, lifetimes, and spectra



1a **1b**

reflect environmental conditions surrounding the fluorophore. To contribute to such an understanding, the present work examines the effect of pH in aqueous solution and of solvent polarity on the fluorescence properties of 3,N^4-ethenocytidine, related derivatives, and model compounds. This study follows