for the C-6 methyl groups in the product acetates. The peak area for the *p*-methyl group in the aromatic ring was used as an internal standard. A solution of 100 mg (0.316 mmol) of the tosylate **16** in 1 ml of acetic-d_4 acid (0.316 *M*) was placed in an nmr tube, and the spectrum was run (probe temperature ~35°). The tube was placed in the vapors of a refluxing ether bath (34.5°). The tube was withdrawn periodically, and the spectrum was obtained. The following results were obtained (time in seconds and per cent unreacted tosylate **16** remaining given): 0, 90.1; 2940, 46.1; 4440, 24.7; 5710, 15.0; 7080, 8.8. The rate constant, 4.0×10^{-4} sec^{-1}, and standard derivation, $\pm 0.2 \times 10^{-4}$ sec^{-1}, were calculated by the least-squares method described previously.[3b]

(b) *anti*-Tosylate 23. In a manner analogous to that described in part a, the following results were obtained for the *anti*-tosylate **23** (time in seconds and per cent unreacted tosylate **23** remaining given): 0, 89.2; 1290, 64.3; 2700, 34.2; 4240, 17.6; 5495, 9.9; 6900, 5.9. The first-order rate constant and standard derivation, calculated as in part a, were $4.3 \pm 0.4 \times 10^{-4}$ sec^{-1}.

Addition of Acetic Acid to Olefin 19. (a) *p*-Toluenesulfonic Acid Catalyzed at 45°. A solution of *p*-toluenesulfonic acid monohydrate (189.0 mg, 0.994 mmol) and acetic anhydride (102.1 mg, 1.000 mmol) in 4.0 ml of glacial acetic acid was heated at 45° for 21 hr. The solution was cooled to ~25°, and the olefin **19** (147.8 mg, 1.025 mmol) was added along with another 3.0 ml of acetic acid. The resulting solution (0.14 *M* each in olefin **19** and toluenesulfonic acid) was mixed thoroughly and heated at 45 ± 1° for 4.0 hr. After being cooled to ~25°, the dark purple solution was worked up as described above for the acetolyses to give 161.6 mg (77%) of a colorless oil. Nmr analysis indicated a composition of 69 ± 1% *syn*-acetate **17** and 31 ± 1% *anti*-acetate **24** (planimeter area measurements of the acetoxy methyl resonances on spectra run at 50 Hz sweep width). The remainder of the nmr spectrum and the infrared spectrum of the reaction mixture were consistent with this composition.

(b) *p*-Toluenesulfonic Acid Catalyzed at ~25°. A solution of 308 mg (1.62 mmol, 0.8 *M*) of *p*-toluenesulfonic acid monohydrate and 204 mg (2.0 mmol) of acetic anhydride in 2 ml of acetic acid was stirred for 4 hr. The olefin **19** (288 mg, 2.0 mmol, 1.0 *M*) was added, and the solution was stirred at ~25° for 3 hr. The reaction was greater than 95% complete (gc analysis). The dark blue solution was poured into water and extracted with ether. The ether extracts were washed with water, saturated sodium bicarbonate solution, and again with water. After the solution was dried, the solvent was removed under vacuum to give 380 mg (93%) of a 68:32 mixture of *syn*-**17** and *anti*-**24** acetates, respectively.

(c) Attempted Uncatalyzed Addition at 45°. A solution of the olefin **19** (151.7 mg, 1.052 mmol) in 7.0 ml of glacial acetic acid (0.15 *M*) was heated at 45 ± 1° for 48 hr. Gc analysis (10 ft × 0.25 in. column packed with 20% E-60 Silicone Nitrile on 60–80 mesh Gas-Chrom Z at 150°, flow rate 44 ml/min) showed no (<0.4% conversion) acetates **17** or **24** (both t_R's 30.3 min); only olefin **19** (t_R 8.2 min) was detected in addition to the solvent.

(d) Uncatalyzed Addition at 100°. The solution of unreacted olefin **19** which was obtained in the preceding section (part c) was heated at 100 ± 1° for 168 hr. Gc analysis as in part c after 48 and 168 hr indicated conversions of 39 and 83% (based only on peak areas of acetates **17** and **24** and olefin **19**, and assuming equal thermal conductivities for **17**, **19**, and **24** on a molar basis, see below), respectively, to the acetates **17** and **24**; no other products were detected. These conversions represent a first-order rate constant of 2.9×10^{-6} sec^{-1}. The amber reaction solution was worked up as in part a above to give 161.8 mg (64% conversion to acetates **17** and **24**, 16% of olefin **19** recovered, 76% yield of **17** and **24**) of pale yellow oil. Nmr analysis ($CDCl_3$) indicated a molar ratio of total acetates **17** and **24** to olefin **19** (two olefinic proton doublets centered at 4.62 and 4.56 ppm) of 80:20 and a *syn*-acetate **17** to *anti*-acetate **24** ratio of 63 ± 1:37 ± 1. The infrared spectrum was consistent with this composition.

Acknowledgment. The authors wish to thank Professors J. C. Martin, H. C. Brown, and D. G. Farnum for helpful discussions.

Semiempirical Calculations on the Electronic Structure and Preferred Conformations of Thiamine (Vitamin B_1) and Thiamine Pyrophosphate (Cocarboxylase)

Frank Jordan

Contribution from the Department of Chemistry, Rutgers University, Newark, New Jersey 07102. *Received January 16, 1974*

Abstract: Extended Hückel and CNDO/2 molecular orbital calculations were performed on the title compounds in their various states of ionization. The electronic structures are discussed. The charges then were used to define the regions of preferred conformations in these molecules *via* the sum of a Lennard-Jones and a Coulombic potential. The conformational energy maps for thiamine hydrochloride with unperturbed and ylide (C-2 deprotonated) thiazolium rings are fairly similar. It appears that full 360° rotation around either of the bonds connecting the two aromatic rings (C–C and C–N bond) is subject to substantial barrier of at least 15–18 kcal/mol. The energetically most favored conformational regions apparently can interconvert readily at room temperature. A suggestion is made for possible conformational changes occurring in the coenzyme upon binding of substrates and enzyme.

Thiamine or vitamin B_1 (thiamine hydrochloride, THC) is the important precursor to the coenzyme thiamine pyrophosphate (TPP). TPP has been strongly implicated in a number of enzymic processes among the most prominent of these being the α-keto acid decarboxylases and transketolases.[1]

Considerable evidence has been gathered especially concerning the mode of action of the enzyme pyruvate decarboxylase, PDC (from brewer's or baker's yeast). PDC requires TPP and Mg(II) for its action.[1] The model systems for this action strongly suggest the existence of a covalently bound pyruvate–TPP intermediate which intermediate is the actual decarboxylating specie.

The proposed mechanism involves several steps in the process.[2] First, the ylide of the coenzyme is formed by loss of a proton from C-2 of the thiazolium (THZ)

(1) L. O. Krampitz, *Annu. Rev. Biochem.*, **38**, 213 (1969).

(2) (a) R. Breslow, *J. Amer. Chem. Soc.*, 82, 3719 (1960); (b) J. Crosby and G. E. Lienhard, *ibid.*, **92**, 5707 (1970).

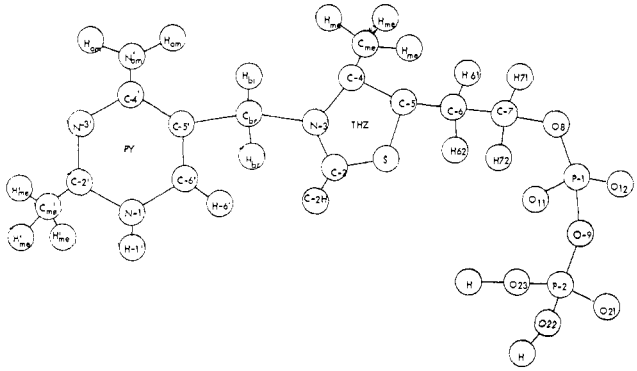


Figure 1. Numbering system in the thiamine pyrophosphate molecule: PY, 4-aminopyrimidine ring, THZ, thiazolium ring; prime positions refer to PY ring; br stands for bridge between two rings.

ring (see Figure 1 for numbering). Next the resulting carbanion attacks nucleophilically at some carbonyl carbon of the substrate forming the covalent substrate–TPP intermediate. This intermediate undergoes condensation or decarboxylation followed finally by a step involving the loss of product from the covalently bound product–TPP intermediate.

Because of the author's interests in both the experimental and the theoretical aspects of enzymic and thermal keto acid decarboxylations, the study was undertaken to calculate the electronic properties of the vitamin (THC), the coenzyme (TPP), and their corresponding ylides with a view of eventual elucidation of the binding and catalytic mechanism of such reactions from a theoretical point of view.

This report summarizes results of calculations on the electronic structures of, and of the allowed conformational regions of the coenzyme and of its precursor vitamin.

Methods of Calculation

(1) The molecular orbital calculations employed were of necessity two types. The electronic properties of THC were calculated both by the extended Hückel[3–5] and by the CNDO/2[6] method.

This allows for direct comparison of the results by two rather different approaches, with the CNDO/2 Self-Consistent Field (SCF) approach providing perhaps better insight to ionization potentials and dipole moments.[6] Because of the size of TPP no SCF calculations were performed on it, only the uniterated extended Hückel ones.

Solid state high quality X-ray crystallographic data are available in the literature for both THC[7] and TPP[8] and these geometries were used in the calculations. THC has a proton on the N-1′ of the pyrimidine ring in the reported structure.[7] Calculations were performed on this one (THC; THZ^+, PY^+), on the deprotonated pyrimidine (THC; THZ^+, PY^0) and finally on the ylide structure (THC; THZ^0, PY^0) always without the accompanying counterion. The solid-state

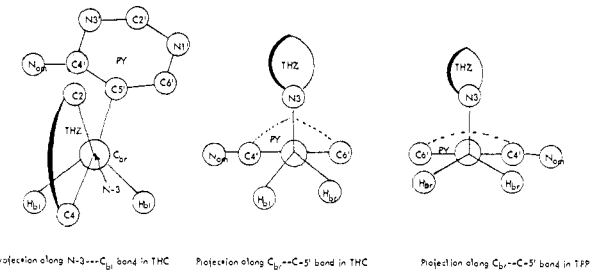


Figure 2. Newman projections for the ϕ_{C-N} and ϕ_{C-C} rotations.

structure of TPP also has a proton on the N-1′ atom of the pyrimidine ring[8] and a mononegative pyrophosphate with the negative charge residing on the inner phosphorus (P1) oxygens (the α position). Calculations were performed on TPP^+ (THZ^+, PY^+, PP^-), on TPP^0 (THZ^+, PY^0, PP^-), and on TPP^- (THZ^0, PY^0, PP^-) the last one possessing on ylide thiazolium ring (*i.e.*, C-2 deprotonated).

(2) Conformational energy calculations were performed assuming a simple nonbonding interaction potential made up of a Lennard-Jones (L-J) 6-12 term and of a Coulombic monopole–monopole interaction term ($E_Q = kq_iq_j/r_{ij}\epsilon$; where q_i and q_j are the net atomic charges of the nonbonded atoms, r_{ij} the distance separating the charges, k a numerical constant, ϵ the dielectric constant here assumed to be 1). Such a potential has been shown to give results in qualitative agreement with extended Hückel molecular orbital results in some nucleosides and nucleotides.[9,10] This type of calculation is orders of magnitude cheaper than molecular orbital calculations and provides a respectable description of the allowed conformational regions. The conformational problem is, in the simplest approximation, a two-dimensional one. It involves rotation of the two aromatic rings (THZ and PY) with respect to the bridge methylene group. An inspection of simple molecular models indicated that the pyrophosphate has considerable freedom of movement and does not have to couple its degrees of freedom to that of the two aromatic rings. Because of this observation, the conformational map of THC was calculated only, with the assumption that the relative disposition of the two aromatic rings should remain independent of the substituent on the O-8 atom (*i.e.*, whether THC or TPP is being considered). The following description of the geometry is helpful (see Figure 2 for appropriate Newman projections). Looking down the N-3 (THZ)–C_{br} bond the environment around the bridge carbon was rotated counterclockwise with 10° increments (ϕ_{C-N}). For each 10° rotations with respect to the ϕ_{C-N} angle, further 10° incremental counterclockwise rotations were performed looking down the C_{br}–C-5′ (PY) bond, this time rotating the entire pyrimidine ring system (ϕ_{C-C}). For each of these points specified by a ϕ_{C-N} and ϕ_{C-C} angle all nonbonded interactions were computed using the potential described above. The crystallographic geometry of THC[7] was assumed to have ϕ_{C-N} and ϕ_{C-C} of 0° (as in Figure 2).

It needs to be pointed out that no torsional contribution is included in the calculations. Since the two aromatic rings are bonded to the bridge methylene

(3) R. Hoffmann, *J. Chem. Phys.*, **39**, 1397 (1963) for original theory.
(4) D. B. Boyd and W. N. Lipscomb, *J. Theor. Biol.*, **25**, 403 (1969), for P, O, N, H, and C.
(5) D. B. Boyd, *J. Amer. Chem. Soc.*, **94**, 6513 (1972), for S parameters.
(6) J. A. Pople and G. A. Segal, *J. Chem. Phys.*, **44**, 3289 (1966); Quantum Chemistry Program Exchange, No. 141.
(7) J. Kraut and H. J. Reed, *Acta Crystallogr.*, **15**, 747 (1962).
(8) J. Pletcher and M. Sax, *J. Amer. Chem. Soc.*, **94**, 3998 (1972).
(9) F. Jordan, *J. Theor. Biol.*, **41**, 23, 375 (1973).
(10) F. Jordan, *Biopolymers*, **13**, 289 (1974).

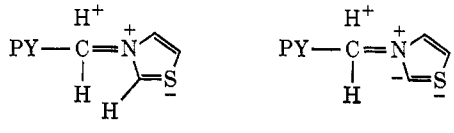
carbon with sp^2 hybrids, as a first approximation the torsional barrier should be very small as also evidenced by the magnitude of the torsional force constants of some model compounds.[11] Interestingly, the theoretical values for torsional barriers in the ethyl cation appear to be very near zero, a situation also involving an sp^2 hybridized C rotated with respect to the CH_3 group.[12]

Results and Discussion

Summary of Molecular Orbital Calculations. To the author's knowledge this is the first report on the electronic structure of the entire vitamin and coenzyme molecule. There is a report in the literature, however, on the π-electronic structure of the aromatic rings.[13]

There are several types of data given by the molecular orbital calculations. Among these are properties related to the wave function (charges, dipole moments, overlap populations) as well as properties related to the orbital energies (extended Hückel) and total molecular (or ionic) energies (CNDO/2) such as ionization potentials and proton affinities (the energy lowering associated with the attachment of a proton to a Brønsted base).

Charge Density–Population Analysis in THC. For this system both EH and CNDO/2 results are available as a function of the state of ionization. Figure 3 presents the charge density matrix for the three species calculated according to CNDO/2. First of all, in the THZ system the quaternary N (N-3) carries only 0.65 positive charge, the rest being delocalized throughout the two rings in the THZ^+ PY^0 specie (middle figures). Upon forming the ylide (top numbers) most of the positive charge is removed from N-3 (0.6 electron added) only 0.26 electrons added to C-2 (site of deprotonation), 0.4 electrons removed from S, and 0.2 electrons added to C-4. The effects of ylide formation are transmitted along the σ system of the dimethylene side chain as well as through the bridge carbon onto the pyrimidine ring. It should be pointed out that the charge densities quoted are already rounded off and the figures quoted are significant. The effect on the PY ring appears interesting, C-5′, C-6′, and N′-amino becoming more negative and N-3′, N-1′, and C-2′ less negative (more positive) upon ylide formation in the THZ ring. It is likely that the source of this phenomenon is inductive since it is not evident how else the bridge methylene could mediate these changes. As to whether the bridge methylene can participate in hyperconjugative interactions with the thiazolium ring of the type shown is difficult to ascertain. Certainly in



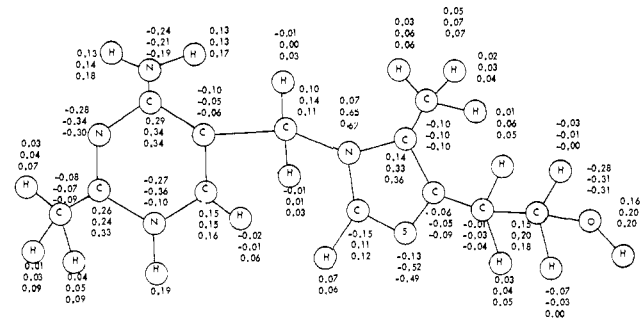


Figure 3. CNDO/2 net atomic charges in THZ species: top number, THZ ylide-Py^0, next; THZ^+–Py^0; bottom line, THZ^+–Py^+.

the ylide these should be much less significant than in the positively charged ring. The fact that in the ylide the sulfur atom is less negative (more positive) perhaps does point to such an effect. If such an effect were operative toward the PY ring more negative charges on C-2′, C-4′, and C-6′ should result in the PY system (true at C-4′ only). Conceivably both factors operate synchronously. While the present results support the principal resonance contribution suggested in ref 8 for THZ^+, they do not agree with the other suggestions.

Protonation on the pyrimidine ring distributes the effect of the positive charge throughout the ring also affecting *all* pyrimidine hydrogens significantly. A quick glance at changes in charges upon PY protonation indicates that the two methyls (PY and THZ) as well as the two aromatic protons (on C-6′ and C-2) behave differently and this should be an aid in assigning ^1H and ^{13}C resonances[14] in the nmr spectrum. Aside from the delocalization of the positive charge in the PY ring (also found in a prior study on nucleic acid base protonation[15]), a strong inductive effect mediated onto the THZ ring is again evident. The extended Hückel calculations present the results of the Mulliken population analysis[16] including net atomic populations (corresponding to charges) bond overlap populations (corresponding to bond orders). These results (Figures 4–6) are very different from the CNDO/2 results quoted above. First of all, the changes on both ylide formation and pyrimidine N-1 protonation are confined to only the immediate neighbors. On ylide formation the C-2 atoms become very negative (site of deprotonation), the N-3 more positive, and S less positive. The only significant change in overlap populations concerns the C-2–S bond which apparently becomes stronger upon ylide formation (the ring geometry is not being changed).

PY protonation again only appears to affect the vicinity at the site of protonation, N-1 and its hydrogen totally absorbing the effects. Evidently the σ protonation does not affect the aromatic nature of the pyrimidine ring significantly as indicated by the very similar overlap populations around the ring all indicative of partial double bonds both in neutral and in N-1 protonated structures. The amino group also is apparently part of the conjugated system based on the C-4′ $\cdots$ N_{am} overlap population (and bond length).

(11) A comparison of the H—C—C=O torsional force constant of 0.003 in acetaldehyde to the H—C—C—H force constant of 0.008 in propane (both in the same units) would give a maximum of about 1 kcal/mol torsional barrier for acetaldehyde (see W. Spindel, M. J. Stern, and E. U. Monse, *J. Chem. Phys.*, **52**, 5022 (1970) for the force constants). The only vaguely related carbon compound, toluene, has a measured sixfold barrier of only 14 cal/mol; see J. P. Lowe, *Progr. Phys. Org. Chem.*, **6**, 1 (1968). Thus it appears reasonable to assume that the torsional barrier with respect to the C_{br} environment should be smaller than 1 kcal/mol.

(12) L. Radom, J. A. Pople, and P. v. R. Schleyer, *J. Amer. Chem. Soc.*, **94**, 5935 (1972).

(13) B. Pullman and C. Spanjaard, *Biochim. Biophys. Acta*, **46**, 576 (1961).

(14) J. Suchy, J. J. Mieyal, G. Bantle, and H. Z. Sable, *J. Biol. Chem.*, **247**, 5905 (1972).

(15) F. Jordan and H. D. Sostman, *J. Amer. Chem. Soc.*, **94**, 7898 (1972); **95**, 6544 (1973).

(16) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833 (1955).

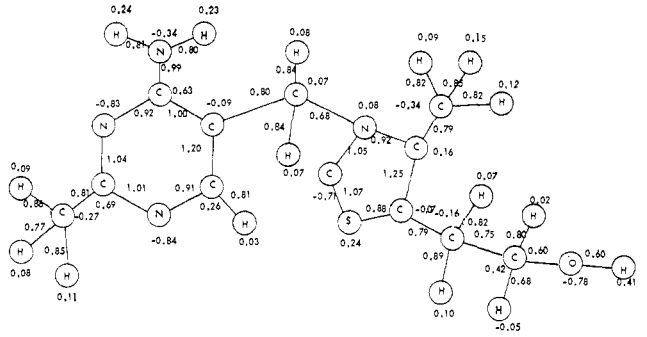


Figure 4. Extended Hückel results on net atomic and overlap populations in THC^0: thiazolium ylide, pyrimidine^0.

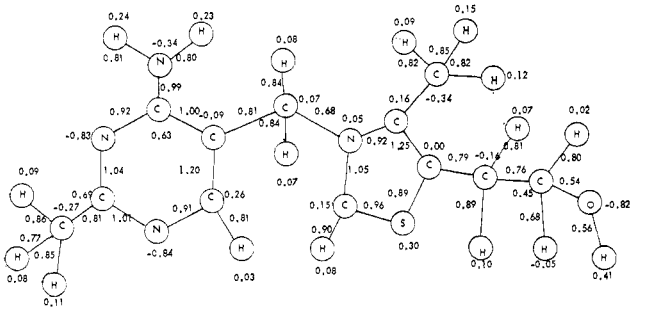


Figure 5. Extended Hückel results on net atomic and overlap populations in THC^+: thiazolium$^+$, pyrimidine^0.

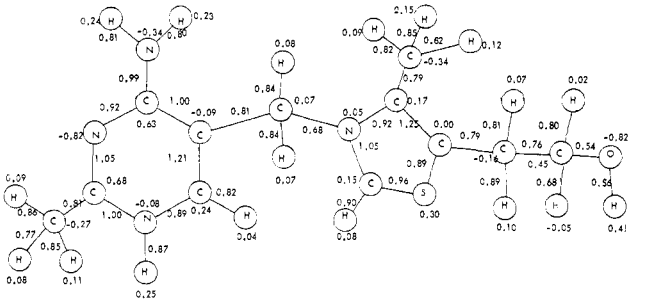


Figure 6. Extended Hückel results on net atomic and overlap populations in THC^{2+}: thiazolium$^+$, pyrimidine$^+$.

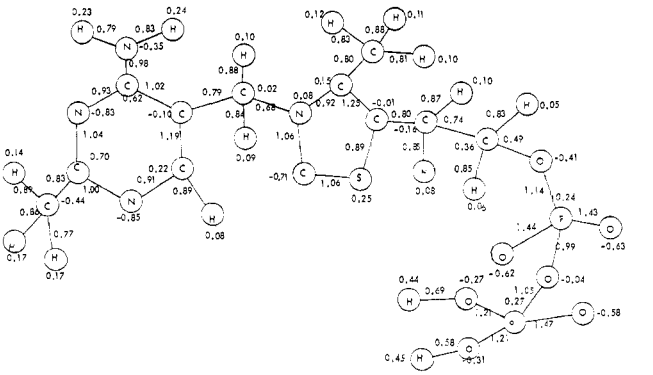


Figure 7. Extended Hückel results on net atomic and overlap populations in TPP^{-1}: thiazolium ylide, pyrimidine^0, phosphate$^-$.

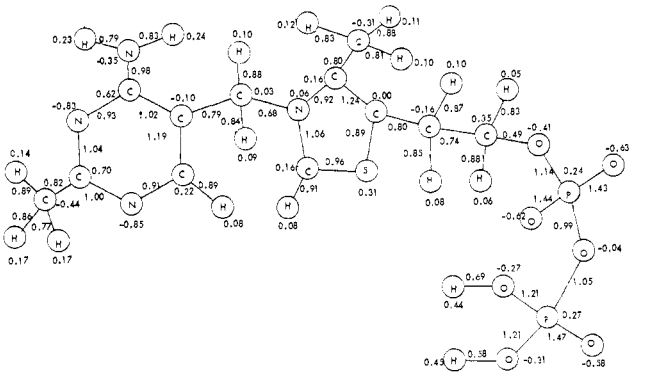


Figure 8. Extended Hückel results on net atomic and overlap populations in TPP^0: thiazolium$^+$, pyrimidine^0, phosphate$^-$.

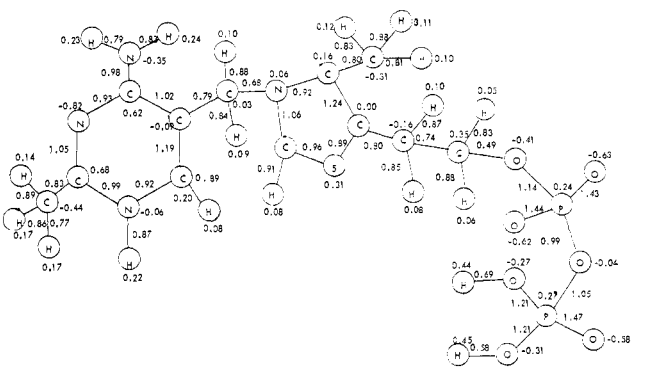


Figure 9. Extended Hückel results on net atomic and overlap populations in TPP^{-1}: thiazolium$^+$, pryimidine$^+$, phosphate$^-$.

Population Analysis in TPP. The results on the three TPP structures are presented in Figures 7–9. Again, the phosphate environment is retained unchanged throughout. The changes upon ylide formation (Figures 7 and 8) and upon pyrimidine protonation (contrast Figures 8 and 9) are very similar to those quoted for THC above.

According to the crystal structure of this diphosphate (pyrophosphate) moiety,[8] the inner (α) phosphorus environment carries the burden of deprotonation with the two oxygens (O-11 and O-12) having essentially equal negative charges, slightly more negative than the one on O-21. While the two phosphorus atoms carry nearly equal positive charges (0.24 for P-1 and 0.27 on P-2), the one further from the dimethylene chain (P-2) is slightly more positive. A comparison of O-8 and O-9 is of interest. O-8 carries a much larger negative charge, −0.41 (between C-7 and P-1) than does O-9 (−0.04, located between P-1 and P-2). Apparently the O-9 atom never participates in hydrogen bonding,[8] perhaps due to its small negative charge. The bond overlap populations are also interesting to consider. There is a clear demarcation of single and double type P–O bonds. A predominantly double bond overlap population is exhibited by the P-1 to O-11, P-1 to O-12, and P-2 to O-21 (1.44, 1.43, and 1.47, respectively) bonds. O–P single bond overlap populations are of the order of 1.14 along O-8–P-1, 0.99 on P-1–O-9, 1.05 on P-2–O-9, and 1.21 along each P-2–O-22 and P-2–O-23. The two O–P bonds between P-1 and P-2 appear to be the weakest (smallest overlap population) with the P-1 to O-9 bond being the weakest of all. This latter bond would probably be the one most susceptible to hydrolytic cleavage (such hydrolysis presumably occurring *via* nucleophilic attack at P-1) to the thiamine monophosphate. An encouraging feature of the population analysis is the consistency of predictions

Table I. Molecular Orbital Results on THC and TPP

Species	Method	E_{Total},[a] kcal/mol	Dipole moment, D	Ionization[b] potential, eV	Proton affinity,[c] kcal/mol	Q(PY)[d]	Q(THZ)[e]
THC;THZ^+,PY^0	CNDO/2	−108678.41		13.59	−93.64	0.24	0.72
	EH	−42696.44		12.01	−223.14	−0.01	1.02
THC;THZ^0,PY^0 (ylide)	CNDO/2	−108451.3	8.12	10.42	−227.28	0.13	−0.17
	EH	−42433.76	6.23	11.30	−262.68	0.02	−0.01
THC;THZ^+ PY^+	CNDO/2	−108772.05		18.83		1.12	0.84
	EH	−42919.58		12.01		1.02	0.99
TPP;THZ^+,PY^0,P^-	EH	−67407.50	19.51	11.99	−213.81	0.02	−0.01
TPP;THZ^0,PY^0,P^- (ylide)	EH	−67138.44		11.22	−269.06	−0.01	−0.97
TPP;THZ^+,PY^+,P^-	EH	−67621.31		11.99		1.00	0.00

[a] According to EH sum of orbital energies times the appropriate occupation number; according to CNDO/2 including total electronic energy and nuclear repulsions. [b] Negative of the energy of the highest occupied molecular orbital. [c] Total energy of protonated species minus that of conjugate base. [d] Charges summed including the bridge atoms. [e] Charges summed including dimethylene chain with its substituents, OH or pyrophosphate.

in THC and TPP based on two different crystal structure models.

The site of pyrimidine protonation is correctly predicted by all methods on both THC and TPP to be the N-1 atom (as evidenced by the X-ray results[7,8]), this atom consistently having the largest negative charge (or net atomic population). Solution experimental evidence also points to this site for the protonation near pH 5.[14] There seems to be no overwhelming preference for N-1 over N-3, however. The amino nitrogen is invariably less susceptible to protonation than are the ring nitrogens. This phenomenon of preference for ring N protonation in systems containing a 4-aminopyrimidine moiety (*e.g.*, adenine) is by no means unprecedented.[15]

Participation of the Sulfur in the THZ Ring. Both theoretical methods utilize a similar basis set for S, one 3s, three 3p, and five 3d orbitals accommodating the valence electrons. Participation of S d orbitals has always been a puzzling question to both theoretician and experimentalist. According to CNDO/2 (from the diagonal elements of the density matrix), the electron populations are 1.78 in 3s, 3.81 in 3p, and 0.53 electrons in 3d in the ylide and 1.79 in 3s, 3.98 in 3p, and 0.74 in 3d orbitals in the thiazolium (+) ring. Apparently, the d type orbitals are the ones most influenced by ylide formation. A bond index for each bond in the thiazolium ring was calculated as the sum of the squares of the off-diagonal density matrix elements between the two bonded atoms.[17] The values are 1.483 (0.305 to d orbitals) for S–C-2, 1.419 for C-2–N-3, 1.063 for N-3–C-4, 1.662 for C-4–C-5, and 1.176 (0.259 to d orbitals) for C-5–S in the ylide and 1.429 (0.379d), 1.485, 1.116, 1.531, and 1.314 (0.371d) for the corresponding bonds in the positively charged thiazolium ring. The total numbers add up to 6.88 bond orders in the positively charged and 6.80 bond orders in the ylide contrasted to 7.00 for the Kekule structure. The implication is that significant and not too different double bond character exists in both neutral and protonated rings along most bonds. In both neutral (ylide) and protonated thiazolium rings the two bonds to sulfur have considerable double bond character with sizable d orbital participation. In the protonated ring the N-3–C-2–S–C-5 system has most of the multiple bond character whereas in the ylide the C-4–C-5 double bond is more localized.

Dipole Moments and Ring Charges. Table I summarizes the dipole moments where these are applicable. Both theoretical methods assign significant dipole moments to the ylide THC structure. The only one calculated for TPP (THZ^+, PY^0, P^-) is very large due to the distance of separation between THZ^+ and P^-. Also listed are partial charge sums for the PY ring (including C_{br} and attached hydrogens) and THZ ring with the phosphate moiety where applicable. While the extended Hückel method does not allow for charge transfer between THZ and PY, the CNDO/2 method does.

Quantities Associated with Molecular and Orbital Energies. The uniterated extended Hückel method assigns similar ionization potentials (deduced from the negative of the highest occupied molecular orbital energy according to Koopmans' theorem) to the TPP (THZ^+, PY^+, P^-) and TPP (THZ^+, PY^0, P^-) and somewhat lower figure to the TPP (THZ^0, PY^0, PY^-) species (Table I) EH also gives similar results on THC and on TPP. CNDO/2, on the other hand, behaves in the anticipated manner,[15] the neutral THC (THZ^0, PY^0) (ylide) having the lowest, followed by the THC (THZ^+, PY^0) and by the dipositive THC (THZ^+, PY^+) ionization potential. The dipositive species has nearly twice as high an ionization potential as the neutral one.

The calculated molecular energies *per se* are not meaningful. Differences in energies are of interest in that such quantities can be associated with relative gas phase proton affinities,[18] to which the calculations are applicable. The ylide is indicated to be more basic than the 4-aminopyrimidine according to all calculations (as indicated by the more negative energy change for the protonation process) as is also found in solution,[14] pK_a of thiazolium being 14.[19]

Results of the Conformational Energy Calculation. Figures 10–13 present the conformational energy maps for THC. Calculations were performed both for the neutral (ylide) and protonated states of THZ. Two sets are presented for each, with and without the inclusion of the Coulombic interaction energy. For the Coulombic interaction energy the CNDO/2 charges

(17) This approach was suggested in K. B. Wiberg, *Tetrahedron*, **24**, 1083 (1968), based on L. Salem, "Molecular Orbital Theory of Conjugated Systems," W. A. Benjamin, New York, N. Y., 1966, p 39.

(18) E. M. Arnett, *Accounts Chem. Res.*, **6**, 404 (1973).

(19) P. Haake, L. Bausher, and W. Miller, *J. Amer. Chem. Soc.*, **91**, 1113 (1969).

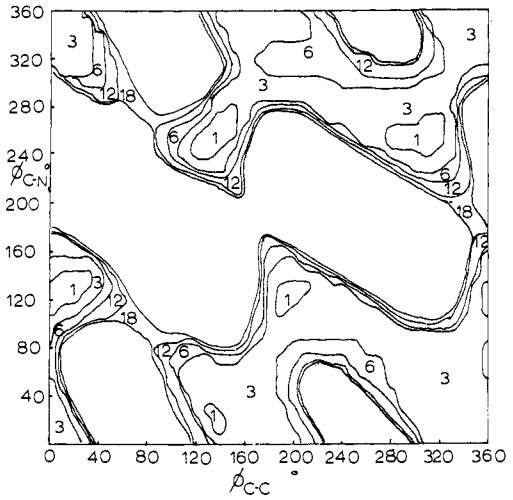


Figure 10. Conformational map on THC including Lennard-Jones energy only. The numbers represent regions with energies within the specified number of kcal/mol from the minimum energy conformer; ϕ_{C-N} and ϕ_{C-C} defined in the text and in Figure 2.

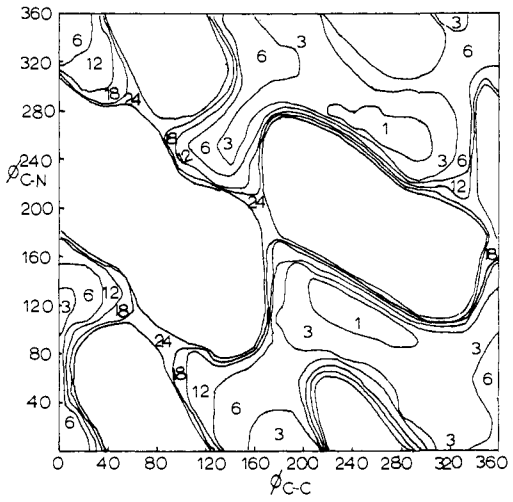


Figure 11. Conformational map on THC including Lennard-Jones and Coulombic energies.

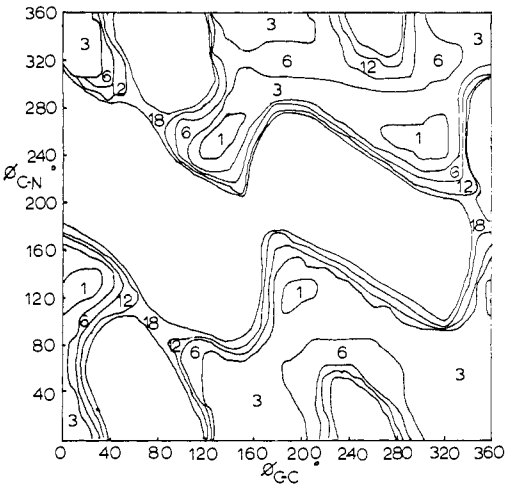


Figure 12. Conformational map on THC ylide including Lennard-Jones energy only.

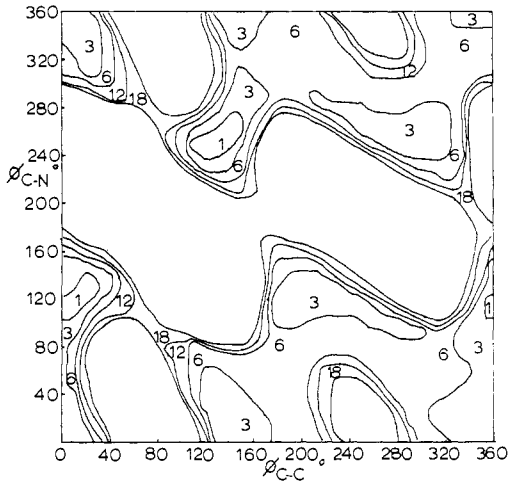


Figure 13. Conformational map on THC ylide including Lennard-Jones and Coulombic energies.

were utilized. Since the dielectric constant was assumed to be one, the actual Coulombic contribution is probably smaller than here suggested. While in other studies on nucleosides and protein configurations an ϵ of 4 is commonly employed, justification of any value is tedious; hence the present study suggests the maximum possible contribution from the Coulombic term.

It should be reemphasized that the torsional barrier is neglected hence an unknown but probably small correction is neglected (*vide supra*).

The total nonbonded interaction energy was computed for each 10° rotation along the ϕ_{C-N} and ϕ_{C-C} rotational angles and the two dimensional energy surface is plotted indicating energies above the lowest energy conformations. In the plots the numbers indicate those regions whose energy is within the given number of kcal/mol above the lowest energy conformation. For example, the regions marked by 1 are all less than 1 kcal/mol less stable than the most stable conformers. The cutoffs of course are arbitrary and subject to the uncertainty due to lack of a torsional barrier. The regions indicated are meaningful in a gross sense. The unmarked regions are at least 18 kcal/mol above the lowest energy conformers. The following conclusions can be drawn.

(1) The crystallographic geometry of THC[7] (ϕ_{C-N} = ϕ_{C-C} = 0°) is not at an absolute minimum, crystal packing forces must be at least partially responsible for the conformation found.

(2) The region containing the experimental TPP[8] conformation (essentially obtained by a near 180° ϕ_{C-C} rotation of the experimentally observed THC conformer) is as stable a region as that containing THC.

(3) The least energy path connecting the experimentally found THC and TPP regions takes less than 6 kcal/mol in any approximation for either substrate examined; hence such interconversion should be facile at room temperature but should become slow at low temperatures. Such a path can be achieved by smaller rotations around both ϕ_{C-N} and ϕ_{C-C} but *not by rotations* along one conformational freedom alone. In fact, *complete 360° rotation around any ϕ_{C-N} or ϕ_{C-C} is subject to a substantial barrier of perhaps at least 18 kcal/mol.*

(4) Inclusion of the Coulombic term in the THC map (normal thiazolium) introduces some changes:

raising the energy of some regions, lowering the energy of some others (contrast Figures 10 and 11). It appears to slightly raise the energy of the region containing the crystallographic conformation.

(5) Inclusion of the Coulombic term in the ylide conformational maps (Figures 12 and 13) reduces the regions of most favorable conformations as shown by comparison of the sizes of the 3-kcal regions.

(6) Contrasting the thiazolium map with the ylide map (Figures 10 and 12) in the L–J approximation shows *no* quantitative differences.

(7) Finally, contrasting the thiazolium and ylide maps (Figures 11 and 13) employing coulombic and L–J terms shows some quantitative changes. In general terms the ylide appears to have larger regions within the 3 kcal/mol boundaries while raising the energy of the lowest energy (1 kcal/mol) regions.

Because of the uncertainty in the magnitude of Coulombic contributions what one can suggest with certainty is that the gross features of the conformational maps are not significantly influenced by ylide formation. If the suggestion of a hydrophobic catalytic site[2b] is correct, the Coulombic contribution could take on significance.

One may also remark that the simple potential function here employed totally neglects any possible one-electron attractive (or repulsive) nonbonded interactions. These have recently been shown to be of importance in determining certain conformational and configurational preferences. For example, the cis configuration of 1-methoxypropene is suggested to be stabilized by such nonbonding (through space) interactions between the π component of the CH_3 and the oxygen lone pair orbital.[20]

As to whether such interaction could be important in the present potential surface cannot be predicted but remains an interesting possibility. Based on the calculated values in ref 20 (subject to the uncertainty in those calculations), such interactions could influence the local minima in the present potential surface but not the gross features here emphasized.

Discussion

Returning to the introductory statement concerning the proposed mechanism of enzymic pyruvate decarboxylations,[22] how can such calculations be helpful?

First of all, the conformational maps can be used to predict three limited regions readily accessible to the coenzyme. It is likely that the ylide exists in the same regions of preferred conformations as does the positive ion. What is the conformation of coenzyme upon substrate binding or upon binding to the enzyme?

First of all, Sable, *et al.*, suggested some years back that the indole nucleus of tryptophan can bind to the THZ ring.[21] In TPP[8] the pyrophosphate side chain is "trans" to the 4′-amino group of the pyrimidine (see Figure 2 and illustration in ref 8). The solution conformations are as of the date of writing unknown. Assuming the solution conformation is basically similar to the solid state one, the tryptophan molecule most probably is bound on the same side of THZ as is the 4-amine of the PY ring to minimize repulsions. Whether such a THZ–tryptophan complex is significant with the enzyme present could be questioned especially since results from this laboratory indicate that the strength of such a complex would greatly diminish in less polar solvents (*e.g.*, more than 30% by volume added ethanol in water). Our results also indicate, however, complex formation of the model 2,3,4-trimethylthiazolium iodide of the ionic association type (as evidenced by a new charge transfer band)[22] suggesting that a negatively charged residue may be involved in the coenzymatic binding mechanism. Such an ionic association process would be greatly enhanced in a hydrophobic environment. Also, the negatively charged residue (*e.g.*, aspartate or glutamate) would be spatially accommodated rather easily on the face of the THZ ring, most likely on the face opposite from the dimethylene pyrophosphate side chain so as to minimize electrostatic repulsions. The calculations confirm that the charge on the thiazolium ring is near +1, on the pyrophosphate near −1 total.

In either case, if the THZ ring is bound to some residue on the enzyme, the conformations found in the solid state on TPP most likely carry over to the enzymatic problem without bound substrate or product. When the substrate and product are coenzyme bound (structures of CH_3–C(TPP)(OH)CO_2^- and CH_3–C(TPP)-(OH)H, respectively), models indicate that the regions near the crystallographically found conformations would become much less favorable. But consider, for example, the map in Figure 11, on which TPP would be located near the $\phi_{C-N} = 0$ and $\phi_{C-C} = 180°$ point in the solid state. In that region rotation around ϕ_{C-C} is already restricted to perhaps 180 ± 40°. Rotation around ϕ_{C-N} by ±100 is feasible with less than 6 kcal/mol barrier. The coenzyme when substrate bound may prefer such distortion along the ϕ_{C-N} bond to reduce repulsions between the PY ring and the substrate or product moieties.

The role of the pyrophosphate and pyrimidine as well as of the essential requirement for divalent ions[1] (specifically Mg(II)) are not simple to account for. Evidence by Schellenberger's group[23] indicates that all aspects of the 4-aminopyrimidine ring are required in the enzyme reaction. As to whether this requirement is for the binding mechanism or the catalytic one is still not known.[14] For many years the amino group was thought to participate in some proton-transfer step (either as donor or as acceptor to enzyme bound substrate and/or product). Based on the fact that N-1′ is always protonated in the solid-state structures,[7,8] this theory is very difficult to accept. Parenthetically, in aqueous medium near pH 5, it is also probably the N-1′ atom which is protonated.[14] The calculations would disfavor such a role for the 4′-amino group based on charge densities. Perhaps this group acts as a hydrogen bond donor in stabilizing the holoenzyme complex.

The pyrophosphate moiety would appear likely to participate in the binding mechanism. It has two strong hydrogen bond donors sites on the terminal

(20) N. D. Epiotis, D. Bjorkquist, L. Bjorkquist, and S. Sarkanen, *J. Amer. Chem. Soc.*, **95**, 7558 (1973). There is a general review of such through space bonding as a possible source of stabilization in R. Hoffmann, *Accounts Chem. Res.*, **4**, 1 (1971); N. D. Epiotis, *J. Amer. Chem. Soc.*, **95**, 3087 (1973).

(21) J. J. Mieyal, J. Suchy, J. E. Biaglow, and H. Sable, *J. Biol. Chem.*, **244**, 4063 (1969).

(22) F. Jordan and B. Farzami, unpublished observations.

(23) A. Schellenberger, K. Wendler, P. Creutzburg, and G. Hubner, *Hoppe-Seyler's Z. Physiol. Chem.*, **348**, 501 (1967), and other articles quoted therein.

phosphate and three nearly equally charged oxygens (O-11, O-12, O-21) which could partake in ionic binding or hydrogen bonding, or perhaps as general acid–base agents for the ylide protonation deprotonation step. The pyrophosphate charge may also be used in maintaining the Mg(II) metal in its required position.

Further theoretical work on the mechanism of binding as well as a conformational analysis of the dimethylene pyrophosphate side chain will be reported later. The need for further experimental work on the definition of the solution conformation of the coenzyme as well as on the details of the enzyme–coenzyme–substrate interactions is obvious.

Acknowledgment. Computer time was generously provided by the Rutgers University Center for Computer and Information Services. The experimental work, on which preliminary results were discussed, was supported in part by the Rutgers University Research Council.

Topography of Cyclodextrin Inclusion Complexes. III. Crystal and Molecular Structure of Cyclohexaamylose Hexahydrate, the $(H_2O)_2$ Inclusion Complex[1]

Philip C. Manor and Wolfram Saenger*

Contribution from Max-Planck-Institut für Experimentelle Medizin, Abteilung Chemie, 34 Göttingen, Germany. *Received September 29, 1973*

Abstract: Cyclohexaamylose, α-cyclodextrin, is a torus-shaped cyclic hexasaccharide consisting of α-(1→4)-linked glucopyranose residues. Due to the annular aperture 4.5–5.5 Å in diameter in the center of the molecule it is able to form inclusion complexes with a variety of substrate molecules even in aqueous solution. Cyclohexaamylose hexahydrate, the $(H_2O)_2$ inclusion complex, crystallized from water in the orthorhombic space group $P2_12_12_1$ with cell dimensions $a = 14.856$ Å, $b = 33.991$ Å, $c = 9.517$ Å, and four formula units per unit cell. The crystal structure has been determined on the basis of 4077 observed diffractometer data and refined by the method of least squares to a residual of $R = 6\%$. Four of the six hydration water molecules are located outside the cyclohexaamylose and are part of an extensive hydrogen bonding network. The two remaining water molecules are located within the aperture and almost on the cyclohexaamylose molecular axis. The two water molecules are hydrogen bonded to one another and the water molecule nearer to the O(6) side of the macrocyclic ring is hydrogen bonded to two O(6) hydroxyl groups. The corresponding two C(6)–O(6) bonds are in gauche,trans orientation with respect to the C(5)–O(5) and C(5)–C(4) bonds while the other four C(6)–O(6) bonds are oriented gauche,gauche. The macrocyclic conformation of the cyclohexaamylose torus is less symmetrical and, according to potential energy calculations, is of higher energy than the conformations found with the iodine, 1-propanol, methanol, and potassium acetate adducts. Thus, when the conformation of the cyclodextrin molecule in solution before adduct formation is the same as that found in the case of the hexahydrate here reported, then a conformational change must necessarily be associated with the inclusion process.

The reaction of cyclodextrin glucanotransferase from *Bacillus macerans* with starch yields the cyclodextrins (Schardinger dextrins) which are cyclic oligosaccharides consisting of six or more α-(1→4)-linked glucopyranose rings. Through the utilization of the 4.5–5.5 Å in diameter void in the center of these torus-shaped molecules, inclusion complexes may be formed both in aqueous solution and in the solid state with the only obvious requirement for complex formation being that the substrate molecule must fit geometrically into the annular void. Eight different nonisomorphous crystalline modifications of α-cyclodextrin (cyclohexaamylose, α-CD) with different guest molecules are known.[2] The structure we report here is that of α-CD hexahydrate which is approximately isostructural with the I_2, 1-propanol, methanol, acetic acid, and butyric acid complexes.

As the members of this series of crystalline inclusion complexes are nearly isostructural with one another, the influence of molecular packing is an almost constant factor. The comparison of the structures of these complexes reveals the influence of the guest molecules on the α-CD conformation. In this regard, the structure of α-CD hexahydrate is particularly interesting as it may represent the structure of the native, "empty" α-CD molecule occurring in aqueous solution and might give a clue as to why α-CD forms inclusion complexes.

This contribution is the third of a series concerned with the structure of cyclodextrin inclusion complexes. A preliminary report of the results of the investigation of the α-CD hexahydrate complex has previously appeared.[3] In this paper the details of the structure determination by X-ray crystallographic means and of the molecular geometry of the α-CD molecule will be discussed.

Experimental Section

α-CD was purchased from Corn Products Development, Englewood Cliffs, N. J., and further purified by recrystallization once from 1-propanol and twice from water. Transparent, colorless,

(1) For part II of this series, see R. K. McMullan, W. Saenger, J. Fayos, and D. Mootz, *Carbohyd. Res.*, 31, 211 (1973).

(2) R. K. McMullan, W. Saenger, J. Fayos, and D. Mootz, *Carbohyd. Res.*, 31, 37 (1973).

(3) P. C. Manor and W. Saenger, *Nature (London)*, 237, 392 (1972).