Structure of Bis(methylguanidinium) Monohydrogen Orthophosphate. A Model for the Arginine–Phosphate Interactions at the Active Site of Staphylococcal Nuclease and Other Phosphohydrolytic Enzymes[1]

F. A. Cotton,*[2a] V. W. Day, E. E. Hazen, Jr.,*[2a] S. Larsen,[2b] and S. T. K. Wong

Contribution from the Departments of Chemistry, Texas A&M University, College Station, Texas 77843, and University of Nebraska, Lincoln, Nebraska 68508. *Received February 6, 1974*

Abstract: A compound has been found which provides an excellent model of certain essential features of the arginyl–phosphate interactions in the complex consisting of *Staphylococcus* nuclease, deoxythymidine 3′,5′-diphosphate, and calcium ion. The crystal structure of this substance, bis(methylguanidinium) monohydrogen orthophosphate, $[CN_3NHC(NH_2)_2]_2PO_3OH$, has been determined. Crystal data: space group, $Fdd2\text{-}C_{2v}{}^{19}$; unit cell dimensions $a = 23.608$ (3), $b = 24.113$ (5), $c = 7.917$ (1) Å; $Z = 16$ for the formula unit $(C_2N_3H_5)_2HPO_4$; $d_{calcd} = 1.436$, $d_{obsd} = 1.430$ g cm^{-3}. Using Zr-filtered Mo Kα radiation a total of 5428 reflections having $\lambda^{-1} \sin\theta < 1.030$ were measured with an automated diffractometer. Using 4465 reflections adjudged to the statistically significant, the structure was solved by Patterson and Fourier methods and refined by full-matrix least squares to final unit-weighted and weighted residuals of 0.049 and 0.048, respectively. The crystallographically independent guanidyl groups are planar and each forms two hydrogen bonds to the $HPO_4{}^{2-}$ ion through separate N–H groups. One phosphate oxygen atom participates in two hydrogen bonds, one from each guanidyl. The overall arrangement is very similar to, though not precisely the same as, that in the enzyme–inhibitor complex and provides an excellent model for the latter. The $HPO_4{}^{2-}$ ions form hydrogen bonded pairs related by a twofold axis. The O–H- -O distances here are relatively short, 2.544 and 2.503 Å, but it would appear that the hydrogens must be considered to be disordered rather than symmetrically located. The P–O distances of 1.514, 1.524, 1.556, and 1.567 Å are values that might be more typically expected for an $H_2PO_4{}^-$ ion and may be a reflection of the hydrogen-bonding effect of the guanidyl ions.

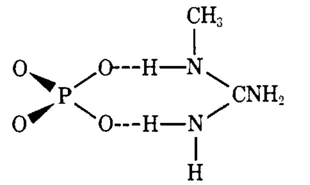
Only in recent years has it become apparent that the guanidyl groups of arginine residues play an important role in binding and possibly even more active roles at the functional sites of both enzymes and noncatalytic proteins. Limiting reference only to those cases where it has been reasonably demonstrated that the chemical modification of arginine does occur at the binding or functional site of the protein molecule, functionally active arginine residues have been found in *E. coli* alkaline phosphatase,[3] yeast inorganic pyrophosphatase,[4] lactate dehydrogenase,[5] D-amino acid oxidase,[6] pepsin,[7] ribonuclease T_1,[8] carboxypeptidases A[9] and B,[10] and antibody combining sites directed against haptens containing such anions as arsonate, phosphonate, and carboxylates.[11–15] Of particular interest are those cases where the chemical modifications in solution can be correlated with the results from crystal structure analyses. Thus, the single arginine which is shown by chemical modification to be at the active site in carboxypeptidase A[9] is very probably Arg-145 which Lipscomb, *et al.*,[16] in their crystallographic studies, have found to bind the terminal carboxylate of peptides. Similarly, the estimation of three essential arginines per subunit of lactate dehydrogenase by Yang and Schwert from their chemical modification studies[5] correlates very neatly with the recent observations from Rossmann's and Kaplan's laboratories. These observations indicate arginines-101, -109, and -171 are located at the active site in the crystal structure of the abortive lactate dehydrogenase–nicotinamide adenine dinucleotide–pyruvate ternary complex. The first of these arginines bridges the pyrophosphate linkage in the coenzyme, and the other two interact with the substrate.[17]

Our direct determination of the high resolution crystal structure of the ternary complex of the Staphylococcal nuclease with its potent competitive inhibitor, thymidine 3′,5′-diphosphate and calcium ion has revealed that the 5′-phosphate of the inhibitor forms two hydrogen bonds each to the guanidinium ions of arginines-

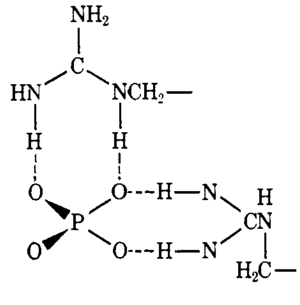
(1) This study was supported by Grant GM13300 from the National Institute of General Medical Sciences, National Institutes of Health. The structure has been briefly described in a preliminary note: F. A. Cotton, E. E. Hazen, Jr., V. W. Day, S. Larsen, J. G. Norman, Jr., S. T. K. Wong, and K. H. Johnson, *J. Amer. Chem. Soc.*, **95**, 2367 (1973).

(2) Address correspondence to these authors at the Department of Chemistry, Texas A&M University, College Station, Texas 77843. (b) On leave, 1970–1971, from the Technical University of Copenhagen.

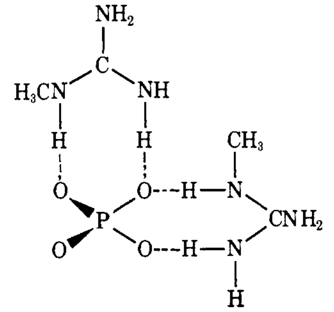
(3) F. J. M. Daemen, *Fed. Proc., Fed. Amer. Soc. Exp. Biol.*, **32**, 473 (1973).
(4) B. S. Cooperman and N. Y. Chiu, *Biochemistry*, **12**, 1676 (1973).
(5) P. C. Yang and G. W. Schwert, *Biochemistry*, **11**, 2218 (1972).
(6) A. Kotaki, M. Harada and K. Yagi, *J. Biochem. (Tokyo)*, **60**, 592 (1966).
(7) W.-Y. Huang and J. Tang, *J. Biol. Chem.*, **247**, 2704 (1972).
(8) A. Takahashi, *J. Biochem. (Tokyo)*, **68**, 659 (1970).
(9) J. F. Riordan, *Biochemistry*, **12**, 3915 (1973).
(10) M. M. Weber and M. Sokolovsky, *Biochem. Biophys. Res. Commun.*, **48**, 384 (1972).
(11) A. L. Grossberg and D. Pressman, *Biochemistry*, **7**, 272 (1968).
(12) M. H. Freedman, A. L. Grossberg, and D. Pressman, *Biochemistry*, **7**, 1941 (1968).
(13) M. H. Freedman, A. L. Grossberg, and D. Pressman, *J. Biol. Chem.*, **243**, 6186 (1968).
(14) G. L. Mayers, A. L. Grossberg, and D. Pressman, *Immunochemistry*, **9**, 196 (1972).
(15) G. L. Mayers, A. L. Grossberg, and D. Pressman, *Immunochemistry*, **10**, 37 (1973).
(16) W. N. Lipscomb, G. N. Reeke, Jr., J. A. Hartsuck, F. A. Quiocho, and P. H. Bethge, *Proc. Roy. Soc., Ser. B*, **257**, 177 (1970).
(17) M. J. Adams, M. Buehner, K. Chandrasekhar, G. C. Ford, M. L. Hackert, A. Liljas, M. G. Rossmann, I. E. Smiley, W. S. Allison, J. Everse, N. O. Kaplan, and S. S. Taylor, *Proc. Nat. Acad. Sci. U. S.*, **70**, 1968 (1973).



1



2



3

35 and -87.[18–20] Additional hydrogen bonds occur between these guanidinium ions and other parts of the enzyme molecule to effectively lock the 5′-phosphate rigidly in the active site. Thus, our structural observations suggest a rather unique and highly specific functional role for these two arginine residues, a notion that is strongly reinforced by the observation of Chaiken and Anfinsen that the replacement of Arg-35 by either lysine or citrulline in semisynthetic variants of the nuclease completely abolishes enzymatic activity.[21,22]

Since the chemical modification studies have indicated the presence of arginines at active sites of other phophohydrolytic enzymes, namely, alkaline phosphatase,[3] inorganic pyrophosphatase,[4] and ribonuclease T_1,[18] there may well be a general subclass of enzymes involved in phosphate metabolism having arginines at their active sites.

These structural and chemical observations as to the specificity and functional importance of the guanidinium–phosphate interactions in the Staphylococcal nuclease lead to a search for simpler model systems.[1b,23] There are two purposes in examining model systems: (1) to confirm the plausibility of the overall interpretation of the enzyme–inhibitor interaction obtained from model fitting to the electron density maps of the nuclease–deoxythymidine 3′,5′-diphosphate–Ca^{2+} complex, and (2) to obtain accurate structure parameters which can provide a more refined understanding of the interactions than that obtainable from the enzyme structure itself.

(18) Historically protein chemists have referred to such interactions between two such groups bearing formal charges as "salt linkages" or "salt bridges." Since this terminology was first introduced, there has been ample evidence from the crystal structures of smaller systems that interactions such as these normally occur *via* hydrogen bonds and since extensive charge delocalization is characteristic of most charged groups found in proteins, the use of this rather artificial terminology when the hydrogen bonding groups both have formal charges, could well be discontinued.

(19) A. Arnone, C. J. Bier, F. A. Cotton, V. W. Day, E. E. Hazen, Jr., D. C. Richardson, J. S. Richardson, and A. Yonath, *J. Biol. Chem.*, **246**, 2302 (1971).

(20) F. A. Cotton, C. J. Bier, V. W. Day, E. E. Hazen, Jr., and S. Larsen, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 243 (1971).

(21) I. M. Chaiken and C. B. Anfinsen, *J. Biol. Chem.*, **246**, 2285 (1971).

(22) In preliminary results from this laboratory, phenyl glyoxal has been found to modify the arginine residues of the nuclease with the concurrent loss of enzymatic activity.

(23) F. A. Cotton, V. W. Day, E. E. Hazen, Jr., and S. Larsen, *J. Amer. Chem. Soc.*, **95**, 4834 (1973).

The search for models was begun with substances obtainable from solutions containing methylguanidinium ion $[CH_3NHC(NH_2)_2]^+$ (MGD) which closely resembles the end of the side chain of an arginyl residue and phosphate ions, $H_nPO_4^{n-3}$. The first substance obtained and examined[23] was $(MGD)H_2PO_4$, which exhibits the type of double hydrogen-bonded interaction that is of interest, **1**, but does not mimic the entire arrangement seen in the enzyme–inhibitor complex, **2**.

When crystals of a second methylguanidinium phosphate, $(MGD)_2HPO_4$, were obtained, it was considered worthwhile to carry out another structure determination with the objective of finding a more complete facsimile of the enzyme–substrate arrangement, **2**. This objective has been accomplished, since, as shown here, $(MGD)_2HPO_4$ contains the arrangement, **3**, which very closely resembles **2**, though it does not precisely duplicate it.

Experimental Section

An aqueous solution containing equimolar quantities of methylguanidinium sulfate (Eastman Organic Chemicals) and $Ba(OH)_2$ was stirred overnight to precipitate $BaSO_4$ which was then filtered off. Half an equivalent amount of phosphoric acid was added to the filtrate to form an aqueous solution of bis(methylguanidinium) monohydrogen phosphate (pH 7.5). After evaporating this solution to near dryness, ethanol was added until the solution became turbid. Large single crystals of bis(methylguanidinium) monohydrogen phosphate grew over a period of several days.

Anal. Calcd for $[(NH_2)_2CNHCH_3]_2(HPO_4)$: C, 19.67; N, 34.42; H, 7.02. Found: C, 19.42; N, 34.19, H, 7.03.

A spherical specimen 0.70 mm in diameter was ground from a larger crystal and glued to the end of a glass fiber with a tip diameter of 0.10 mm.

Precession photographs, used to determine a preliminary set of lattice constants, indicated orthorhombic, *mmm*, symmetry. The systematically absent reflections were those uniquely required by the noncentrosymmetric space group, $Fdd2$-C_{2v}.[19] This choice was fully supported by the positive results of sensitive tests for piezoelectricity and by the subsequent structure determination. The crystal was accurately centered on a Syntex P$\bar{1}$ full-circle goniometer and a total of 15 reflections, chosen to give a good sampling of reciprocal space and diffractometer settings ($2\theta_{\mathrm{MoK}\alpha} > 50°$), were used to align the crystal and calculate angular settings for each reflection. A least-squares refinement of the diffraction geometry for these 15 reflections, recorded at the ambient laboratory temperature of 21 ± 1° with Mo Kα radiation (λ(Mo Kα) 0.71069 Å) gave the lattice constants $a = 23.608 \pm 0.003$, $b = 24.113 \pm 0.005$, and $c = 7.917 \pm 0.001$ Å. A unit cell content of 16 bis(methylguanidinium) monohydrogen phosphate molecules gives a calculated density of 1.436 g/cm³, in good agreement with the observed density of 1.430 g/cm³, measured by flotation in a mixture of dichloromethane and carbon tetrachloride.

Intensity measurements utilized Zr-filtered Mo Kα radiation and the θ–2θ scanning technique with a 3° takeoff angle and a standard-focus X-ray tube on a computer controlled Syntex P$\bar{1}$ diffractometer. A scanning rate of 3°/min was employed for the scan between 2θ settings 1.0° above and below the calculated Kα doublet values (λ(Kα_1) 0.70926 and (Kα_2) 0.71354 Å) of each reflection except for those reflections having $83.3° < 2\theta < 94.1°$ where a 2°/min scanning rate was used. Background counts (each one lasting half the total scan time) were taken at both ends of the scan range. A total of 5428 independent reflections having ($\sin\theta/\lambda$) < 1.030 (four times the number of data in the limiting Cu Kα sphere) were measured in concentric shells of increasing 2θ containing approximately 1400 reflections each. The six standard reflections, measured every 300 reflections as a monitor for possible disalignment and/or deterioration of the crystal, gave no indication of either.

The linear absorption coefficient of the crystal for Mo Kα radiation is 0.26 cm^{-1}, yielding μR of 0.09 for the spherical crystal used. Since the absorption of X-rays by a spherical crystal having $\mu R = 0.09$ is essentially independent of scattering angle, no absorption correction was made, and the intensities were reduced to relative squared amplitudes, $|F_o|^2$, by means of standard Lorentz and polarization corrections.

Of the 5428 reflections examined, 963 were rejected as objectively unobserved by applying the rejection criterion, $I < \sigma(I)$, where $\sigma(I)$ is the counting statistics standard deviation in the observed intensity computed from

$$\sigma(I) = (C_t + k^2B)^{1/2}$$

C_t being the total count from scanning, k the ratio of scanning time to total background time (in this case $k = 1$), and B the total background count. The remaining 4465 observed intensities were used in the determination and refinement of the structure.

Structure determination was achieved through a combination of the heavy-atom technique, difference Fourier syntheses, and least-squares refinement. The wholly straightforward interpretation of the Patterson synthesis of the 697 $|F_o|^2$ data having $(\sin\theta/\lambda) \leq 0.52$ placed the phosphorus atoms in 16-fold general positions $(0,0,0;\ 0,{}^1/_2,{}^1/_2;\ {}^1/_2,0,{}^1/_2;\ {}^1/_2,{}^1/_2,p) + (x,y,z;\ \bar{x},\bar{y},\bar{z};\ {}^1/_4 - x,{}^1/_4 + y,{}^1/_4 + z;\ {}^1/_4 + x,{}^1/_4 - y,{}^1/_4 + z$. These atomic coordinates, an isotropic thermal parameter, and a scale factor were varied in two cycles of isotropic full-matrix least-squares refinement.[24] This resulted in a conventional unweighted residual, $R_1 = 0.499$, for these low angle data.

$$R_1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$$

A difference electron density map at this stage revealed the locations of the four phosphate oxygen atoms. Two cycles of least-squares refinement varying the scale factor, atomic coordinates, and isotopic temperature factors for phosphorus and oxygen atoms, respectively, gave $R_1 = 0.372$. A second electron density difference map clearly revealed the remaining ten non-hydrogen atoms of the asymmetric unit, all of which lie in general positions. Isotropic full-matrix refinement using unit weighting for the 15 non-hydrogen atoms gave $R_1 = 0.072$, for 697 reflections. All of the 4465 reflections were then included in a fully anisotropic least-squares minimization of the function $\Sigma w(|F_o| - k|F_c|)^2$ to give, with unit weighting (*i.e.*, all $w = 1$) $R_1 = 0.063$. This and all subsequent refinement cycles employed an anomalous dispersion correction[25] to the scattering factor of the phosphorus atom and a least-squares refinable extinction coefficient[26] of the form $F(x) = 1/(1 + 2x)^{1/2}$, where $x = gI$ and g refined to a final value of 0.36×10^{-7}. A Fourier difference synthesis based on the refined parameters afforded direct evidence for the placement of all hydrogen atoms. Further unit-weighted full-matrix least-squares cycles were used to refine hydrogen atoms isotropically and all other atoms anisotropically to give $R_1 = 0.049$ and a conventional weighted residual, $R_2 = 0.044$.

$$R_2 = \{\Sigma w(|F_o| - |F_c|)^2/\Sigma w|F_o|^2\}^{1/2}$$

Empirical weights ($w = 1/\sigma^2$) were then calculated from

$$\sigma = \sum_0^3 a_n|F_o|^n = 1.54 - 0.20 \times 10^{-1}F + 0.23 \times 10^{-3}F^2 - 0.35 \times 10^{-6}F^3$$

the a_n being coefficients derived from the least-squares fitting of the curve

$$||F_o| - |F_c|| = \sum_0^3 a_n|F_o|^n$$

The F_o values were calculated from the fully refined model using unit weighting. The final cycles of least-squares refinement utilized these weights and anomalous dispersion corrections for the phosphorus atom to refine hydrogen atoms isotropically and all other atoms anisotropically together with the scale factor and extinction coefficient to give final values of 0.049 and 0.048 for R_1 and R_2, respectively. During this last cycle of refinement no parameter for non-hydrogen atoms shifted by more than $0.2\sigma_1$ with the average shift being 0.03σ.

The following computer programs were employed in this work: MAGTAP and SCTFT2, data reduction programs written by V. Day; FORDAP, Fourier and Patterson synthesis program, a modified version of A. Zalkin's original program; ORFLSE, full-matrix least-squares refinement program, a highly modified version of Busing, Martin, and Levy's original ORFLS; ORFFE, bond lengths and angles with standard deviations by Busing, Martin, and Levy; ORTEP-II, thermal ellipsoid plotting program by C. K. Johnson; MPLANE, least-squares mean plane calculation program from L. Dahl's group.

(24) Scattering factors were taken from D. T. Cromer and J. B. Mann, *Acta Crystallogr., Sect. A*, **24**, 321 (1968).
(25) D. T. Cromer, *Acta Crystallogr.*, **18**, 17 (1965).
(26) W. H. Zachariasen, *Acta Crystallogr.*, **23**, 558 (1967).

Results

The final coordinates and anisotropic thermal parameters for all atoms except hydrogen atoms are listed in Tables I and II, respectively; the refined positions and isotropic thermal parameters of the hydrogen atoms are listed in Table III.[27] The rule used in the atom numbering scheme for bis(methylguanidinium) monohydrogen phosphate is as follows. Atoms of the methylguanidinium ions are grouped according to cation. A numerical subscript is used to differentiate atoms of the same non-hydrogen element. For each hydrogen atom the subscript letter and first subscript number indicate the atom to which it is covalently bonded, while the second numerical subscript distinguishes among hydrogen atoms attached to the same atom.

A projection of one asymmetric unit is presented in Figure 1; each atom is numbered in conformity with Tables I–IX and each non-hydrogen atom is represented by an ellipsoid having shape, orientation, and relative size consistent with the thermal parameters listed in Tables II. Bond lengths and angles in the molecular skeleton are presented in Figure 2 and are listed along with their estimated standard deviations in Tables IV and V; the dimensions of various interionic hydrogen bonds are listed in Table VI. The equations of the mean planes that partially characterize important subgroupings of atoms within the asymmetric unit specified by the coordinates of Tables I and III are given in Table VII,[27] and the displacements from these planes of the atoms constituting the asymmetric unit are listed in Tables VIII and IX.[27]

Table I. Atomic Coordinates in Crystalline Bis(methylguanidinium) Monohydrogen Phosphate[a]

Atom[b] type	Fractional coordinates 10^5x	10^5y	10^4z
	Cation I		
N_1	9959 (9)	10244 (7)	−3507 (2)
N_2	9264 (11)	10208 (8)	−6414 (2)
N_3	6717 (9)	2542 (6)	−4890 (3)
C_1	12657 (16)	15602 (11)	−3417 (3)
C_2	8655 (8)	7711 (6)	−4935 (3)
	Cation II		
N_1	23115 (7)	1538 (7)	1680 (3)
N_2	28807 (7)	8798 (6)	2535 (3)
N_3	20656 (7)	10454 (7)	1008 (3)
C_1	27274 (12)	−2734 (9)	2017 (5)
C_2	24204 (7)	6942 (7)	1758 (2)
	Anion		
P	8675 (1)	−164 (1)	0
O_1	4917 (5)	−2160 (5)	−1514 (2)
O_2	5189 (6)	−1066 (7)	1645 (2)
O_3	14016 (4)	−3632 (4)	73 (2)
O_4	9965 (5)	5994 (4)	−199 (2)

[a] Figures in parentheses are the estimated standard deviations. Coordinate listed without standard deviation is symmetry required. [b] Atoms numbered to agree with Figures 1–5.

(27) See paragraph at end of paper regarding supplementary material.

Table II. Anisotropic Thermal Parameters in Crystalline Bis(methylguanidinium) Monohydrogen Phosphate[a]

Atom[b] type	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}	B,[c] Å^2
				Cation I			
N_1	4.31 (8)	2.05 (5)	1.64 (5)	−0.60 (5)	0.04 (5)	0.13 (4)	2.40
N_2	5.83 (11)	2.22 (6)	1.73 (5)	−0.65 (6)	−0.01 (6)	0.26 (4)	2.77
N_3	5.17 (9)	2.07 (4)	2.09 (5)	−0.86 (5)	−0.10 (7)	0.16 (5)	2.75
C_1	6.81 (16)	2.97 (8)	2.55 (8)	−1.91 (9)	−0.08 (9)	0.13 (7)	3.48
C_2	2.93 (5)	1.80 (4)	1.73 (4)	0.06 (4)	0.11 (5)	0.27 (5)	2.07
				Cation II			
N_1	2.31 (5)	1.99 (4)	4.17 (8)	−0.32 (4)	−1.16 (6)	−0.01 (5)	2.52
N_2	2.74 (5)	2.20 (4)	3.24 (6)	−0.65 (4)	−1.16 (6)	0.19 (5)	2.49
N_3	2.50 (5)	2.10 (5)	3.83 (8)	0.28 (4)	−0.87 (5)	−0.38 (5)	2.62
C_1	4.01 (9)	2.19 (6)	5.62 (14)	0.23 (6)	−2.12 (10)	−0.00 (7)	3.39
C_2	1.98 (5)	2.01 (4)	2.33 (5)	−0.19 (4)	−0.19 (4)	−0.23 (4)	2.08
				Anion			
O_1	1.85 (3)	2.14 (4)	1.49 (3)	−0.12 (3)	−0.19 (3)	−0.19 (3)	1.79
O_2	1.75 (4)	4.73 (7)	1.33 (3)	0.28 (4)	0.07 (3)	0.25 (4)	2.21
O_3	1.57 (3)	1.82 (3)	2.35 (4)	0.25 (2)	−0.00 (3)	0.15 (3)	1.87
O_4	2.76 (4)	1.35 (3)	3.08 (6)	−0.04 (3)	−0.82 (4)	0.06 (3)	2.19
P	1.41 (1)	1.46 (1)	1.38 (1)	0.04 (1)	−0.07 (1)	0.09 (1)	1.41

[a] The number in parentheses that follows each B_{ij} value is the estimated standard deviation in the last significant figure. The B_{ij}'s in Å^2 are related to the dimensionless β_{ij} employed during refinement as $B_{ij} = 4\beta_{ij/a_i^*a_j^*}$. [b] Atoms numbered to agree with Figures 1–5. [c] Isotropic thermal parameter calculated from $B = 4[V^2 \det(\beta_{ij})]^{1/3}$

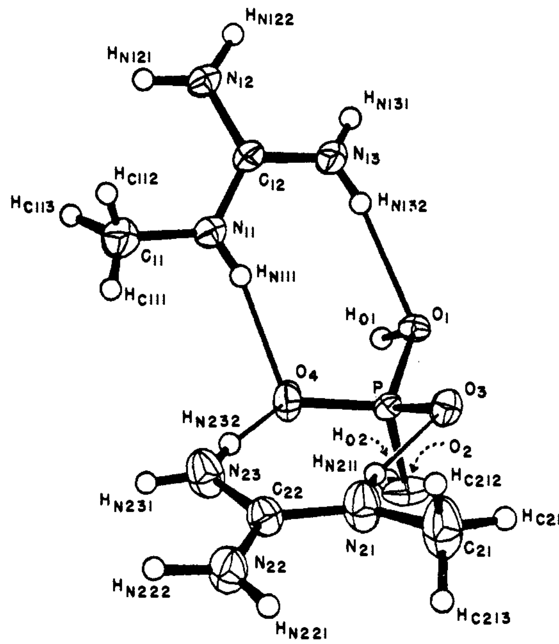


Figure 1. A perspective view of one asymmetric unit. The atom numbering scheme is explained in the text. For clarity the actual thermal ellipsoids of the hydrogen atoms ae not used.

Table III. Refined Parameters for Hydrogen Atoms in Crystalline Bis(methylguanidinium) Monohydrogen Phosphate[a]

Atom[b] type	Fractional coordinates 10^3x	10^3y	10^3z	Isotropic thermal parameter, B, Å^2
		Cation I		
H_{N1}	99 (1)	85 (1)	−262 (4)	0.7 (5)
H_{N21}	101 (1)	139 (1)	−646 (4)	1.5 (6)
H_{N22}	85 (1)	83 (1)	−731 (5)	1.5 (6)
H_{N31}	54 (2)	12 (1)	−570 (5)	2.1 (7)
H_{N32}	61 (1)	8 (1)	−399 (4)	0.2 (4)
H_{C11}	138 (2)	162 (2)	−243 (7)	4.1 (9)
H_{C12}	157 (1)	159 (1)	−415 (5)	2.1 (7)
H_{C13}	102 (2)	189 (2)	−393 (7)	5.3 (11)
		Cation II		
H_{N1}	200 (1)	7 (1)	117 (3)	0.3 (4)
H_{N21}	309 (1)	66 (1)	310 (4)	0.9 (5)
H_{N22}	296 (1)	127 (1)	246 (5)	1.3 (5)
H_{N31}	211 (1)	138 (1)	112 (4)	1.1 (5)
H_{N32}	175 (1)	92 (1)	54 (4)	1.0 (5)
H_{C11}	256 (2)	−62 (3)	196 (8)	7.6 (16)
H_{C12}	309 (2)	−25 (2)	128 (7)	5.8 (13)
H_{C13}	278 (2)	−26 (2)	322 (8)	5.4 (14)
		Anion		
H_{O1}	0.0	0.0	−138 (11)	5.8 (19)
H_{O2}	0.0	0.0	143 (12)	8.8 (25)

[a] Figures in parentheses are the estimated standard deviations of the last significant digit. Coordinates listed without standard deviations are symmetry required. [b] Atoms numbered to agree with Figures 1–5.

Discussion

Overall Structure. As emphasized in Figure 3, the structure can be thought of as centered around a pair of monohydrogen phosphate ions which are held together by two strong hydrogen bonds. The O· · ·O distances in these bonds are 2.544 and 2.503 Å which means that they are close to or perhaps within the range where the hydrogen bonds might be symmetrical. Moreover, there is a crystallographic twofold axis which passes through the midpoints of the two hydrogen-bonded O· · ·O pairs. However, this symmetry requirement could be satisfied either by having the hydrogen atoms on the twofold axis (truly symmetrical hydrogen bonds) or by having them off the axis but disordered. The experimental data did not allow a choice; both the truly symmetric structure or one in which there are nearly symmetric but disordered bonds are consistent with the data. However, Speakman, in a recent review of short hydrogen bonds,[28] has concluded that symmetrical hydrogen bonds are likely only when the O--H--O distance is less than 2.44 Å, and on this basis we believe that the hydrogen bonds in this structure are unsymmetrical and disordered. Should these hydrogen atoms become ordered at lower temperatures, the re-

(28) J. C. Speakman, *Struct. Bonding (Berlin)*, 12, 141 (1972).

Table IV. Bond Lengths in Bis(methylguanidinium) Monohydrogen Phosphate[a]

Type[b]	Bond length, Å: Cation I	Cation II	Av	Type[b]	Bond length, Å: Cation I	Cation II	Av
C_1–N_1	1.442 (3)	1.448 (3)	1.445	C_1–H_{C11}	0.84 (5)	0.93 (6)	
C_2–N_1	1.322 (3)	1.330 (2)		C_1–H_{C12}	0.93 (4)	1.04 (5)	0.96
C_2–N_2	1.324 (3)	1.327 (2)	1.327	C_1–H_{C13}	1.07 (5)	0.96 (6)	
C_2–N_3	1.328 (2)	1.331 (2)			Anion bond length, Å		
C_2–N				P–O_1	1.567 (1)		
Av	1.325	1.329			(1.573)[c]		1.562
N_1–H_{N1}	0.82 (3)	0.86 (3)		P–O_2	1.556 (1)		(1.569)[c]
N_2–H_{N21}	0.91 (3)	0.85 (3)			(1.565)[c]		
N_2–H_{N22}	0.87 (4)	0.96 (3)		P–O_3	1.514 (1)		
N_3–H_{N31}	0.79 (4)	0.82 (3)	0.86		(1.521)[c]		1.519
N_3–H_{N32}	0.84 (3)	0.89 (3)		P–O_4	1.524 (1)		(1.527)[c]
					(1.533)[c]		
				O_1–H_{O1}	1.276 (7)		1.27
				O_2–H_{O2}	1.263 (13)		

[a] The figure in parentheses following each individual distance is the estimated standard deviation. [b] Atoms numbered to agree with Figures 1–5 and Tables I and III. [c] Bond length corrected for libration of HPO_4^{2-} group as a rigid body according to V. Schomaker and K. N. Trueblood, *Acta Crystallogr., Sect. B*, **24**, 63 (1968).

sulting crystals of this compound could prove to be ferroelectric.[29]

Surrounding each central dimer of phosphate ions is a total of 12 methylguanidinium ions, six from each of the two crystallographically distinct cations. These are linked to the phosphate dimer by a total of 18, 20, or 22 hydrogen bonds, namely, ten from cation II and 8, 10, or 12 from cation I. The hydrogen bonding pattern of a single guanidinium ion is illustrated in Figure 4 for cation I and in Figure 5 for cation II. Considering a single phosphate dimer, cation I and II both form a single hydrogen bond to O_3 of each phosphate ion for a total of four. Cation II forms two pairs of hydrogen bonds bridging across the phosphate dimer, giving altogether four of this pattern and a subtotal of eight. Cation I and cation II each form a pair of hydrogen bonds to two oxygen atoms of one phosphate ion, giving eight H bonds of this type for a subtotal of 16. This type of guanidinium–phosphate interaction, which is best illustrated in Figures 1 and 2, has also been observed in the structure of methylguanidinium dihydrogen orthophosphate[1b,23] and in the structure of propylguanidium diethylphosphate.[30] However, the most significant aspect of this type of paired hydrogen bond interaction between a guanidinium ion and a phosphate ion is that it provides an excellent model for the interaction of arginines-35 and -87 of the Staphylococcal nuclease with the 5′-phosphate of its potent inhibitor, thymidine 3′,5′-diphosphate, as illustrated diagramatically by **2**.

The final type of guanidinium–phosphate interaction observed in this structure is the N_{13}--H--O_2 of cation I (Figure 4) with an N–O distance of 2.90 Å which is paired with another N–O interaction to a different oxygen atom of the same phosphate, *i.e.*, N_{12}--H--O_4 with an N–O distance of 3.17 Å. This distance is somewhat long to be considered a real hydrogen bond, and the final N–O interaction shown in Figure 4, that of N_{12}--H--O_2 with an N–O distance of 3.27 Å, is even more doubtful. The grand total of N to O hydrogen bonding interactions to a single dimer of phosphate ions is thus 18, 20, or 22, the latter two numbers including one or both of the pair of N–O interaction over 3 Å. This

(29) W. C. Hamilton and J. A. Ibers, "Hydrogen Bonding in Solids," W. A. Benjamin, New York, N. Y., 1968, pp 238–255.
(30) S. Furberg and J. Solbakk, *Acta Chem. Scand.*, **26**, 3699 (1972).

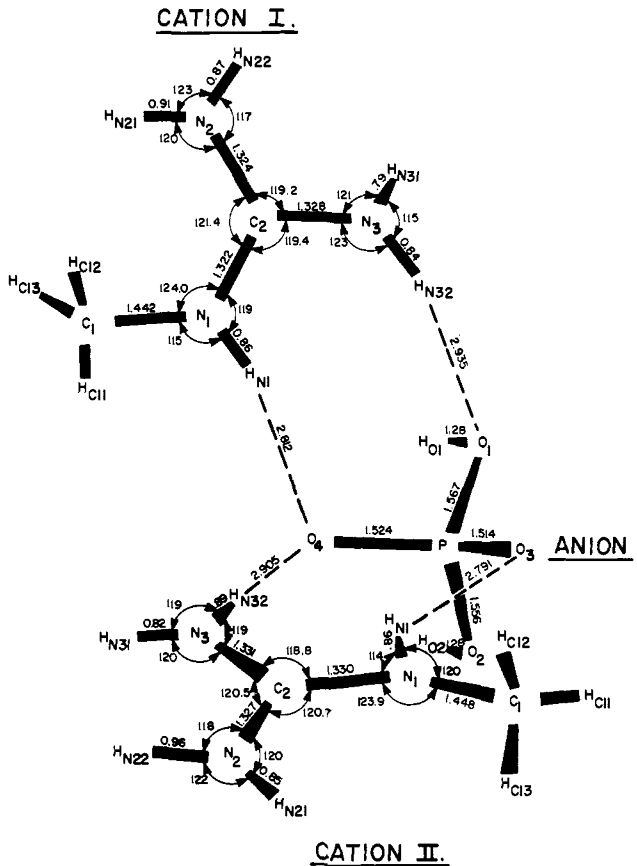


Figure 2. Bond lengths and angles for the asymmetric unit as seen in Figure 1. A complete set of values and standard deviations are listed in Tables IV, V, and VI.

final pattern of guanidinium–phosphate interaction, showing in this structure a third paired cyclic N–O system formed by one strong hydrogen bond and a second rather weak one, may well have some biochemical significance. In the structure of propylguanidinium diethylphosphate[30] one paired interaction with two strong hydrogen bonds is observed, but there is also a second pair with the one strong and one weak pattern that we also observe in this structure. The Staphylococcal nuclease is, of course, a phosphodiesterase with a degree of base specificity. Its initial interaction with

Table V. Bond Angles in Bis(methylguanidinium) Monohydrogen Phosphate[a]

Type[b]	Angle, deg Cation I	Cation II	Av	Type[b]	Angle, deg Cation I	Cation II	Av
N_1–C_2–N_2	121.4 (1)	120.7 (2)	121.1	H_{C11}–C_1–N_1	109 (3)	111 (3)	110
N_1–C_2–N_3	119.4 (2)	118.8 (2)	119.1	H_{C12}–C_1–N_1	112 (2)	114 (3)	113
N_2–C_2–N_3	119.2 (2)	120.5 (2)	119.9	H_{C13}–C_1–N_1	114 (3)	104 (3)	109
C_1–N_1–C_2	124.0 (2)	123.9 (2)	124.0	H_{C11}–C_1–H_{C12}	109 (4)	112 (4)	111
H_{N1}–N_1–C_1	115 (2)	120 (2)	118	H_{C11}–C_1–H_{C13}	114 (4)	98 (4)	106
H_{N1}–N_1–C_2	119 (2)	114 (2)	117	H_{C12}–C_1–H_{C13}	98 (3)	117 (4)	108
				Angles in Anion, deg			
H_{N21}–N_2–H_{N22}	123 (3)	122 (3)	123	O_1–P–O_2	107.32 (6)		
H_{N21}–N_2–C_2	120 (2)	120 (2)	120	O_1–P–O_3	109.33 (7)		
H_{N22}–N_2–C_2	117 (2)	118 (2)	118	O_1–P–O_4	109.47 (7)		
H_{N31}–N_3–H_{N32}	115 (3)	119 (3)	117	O_2–P–O_3	109.35 (8)		
H_{N31}–N_3–C_2	121 (3)	120 (2)	121	O_2–P–O_4	109.19 (9)		
H_{N32}–N_3–C_2	123 (2)	119 (2)	121	O_3–P–O_4	112.07 (7)		
				P–O_1–H_{O1}	109 (3)		
				P–O_2–H_{O2}	112 (4)		

[a] Figures in parentheses are the estimated standard deviations of the last significant digit. [b] Atoms numbered to agree with Figures 1–5 and Tables I and III.

Table VI. Interionic Hydrogen Bonds in Bis(methylguanidinium) Monohydrogen Phosphate

Donor atom (D)	Acceptor atom (A)	Distance, Å[b] D···A	Distance, Å[b] H···A	Angle, deg[b] H–D···A	Asymmetric unit of (A)
		Intracomplex Bonds, Cation I[c]			
N_1–H_{N1}	O_4	2.812 (2)	2.01 (3)	10 (2)	x, y, z
N_3–H_{N32}	O_1	2.935 (2)	2.11 (3)	7 (2)	x, y, z
		Intracomplex Bonds, Cation II[c]			
N_1–H_{N1}	O_3	2.791 (2)	1.96 (3)	14 (2)	x, y, z
N_3–H_{N32}	O_4	2.905 (2)	2.02 (3)	5 (2)	x, y, z
		Intercomplex Bonds, Cation I[c]			
N_2–H_{N21}	O_3	2.837 (2)	1.97 (3)	15 (2)	$^1/_4 - x, ^1/_4 + y, -^3/_4 + z$
N_2–H_{N22}	O_4	3.169 (3)	2.38 (4)	21 (2)	$x, y, -1 + z$
N_2–H_{N22}	O_2	3.268 (3)	2.52 (3)	27 (2)	$x, y, -1 + z$
N_3–H_{N31}	O_2	2.900 (3)	2.17 (4)	19 (3)	$x, y, -1 + z$
		Intercomplex Bonds, Cation II[c]			
N_2–H_{N21}	O_3	2.908 (2)	2.10 (3)	15 (2)	$^1/_2 - x, -y, ^1/_2 + z$
N_2–H_{N22}	O_4	2.865 (2)	1.99 (3)	19 (2)	$^1/_4 + x, ^1/_4 - y, ^1/_4 + z$
N_3–H_{N31}	O_1	2.990 (2)	2.19 (3)	12 (2)	$^1/_4 - x, ^1/_4 + y, ^1/_4 + z$
		Intercomplex Bonds, Anion[c]			
O_1–H_{O1}	O_1	2.544 (3)	1.276 (7)	5 (4)	$-x, -y, z$
O_2–H_{O2}	O_2	2.503 (3)	1.263 (13)	8 (4)	$-x, -y, z$

[a] The hydrogen atom actually involved in the bond is also indicated. [b] Figures in parentheses are the estimated standard deviations of the last significant digit. [c] The complex is considered to be the asymmetric unit of Tables I and III and of Figures 1 and 2. All donor atoms belong to this asymmetric unit.

the substrate through the agency of its two active-site guanidinium groups might fall into the two strong, one weak, one strong pattern of H bonds, found in propylguanidinium diethylphosphate. If, as we have suggested,[20] enzymatic hydrolysis proceeds by the nucleophilic back-side attack of a water molecule or hydroxide ion coordinated to the calcium ion, then, in the process, the guanidinium–phosphate interaction could be transformed easily to the two strong, two strong pattern of interaction, as observed in this structure, with a concomitant increase of the positional and chemical stability of the transition state and consequent reduction of the energy barrier to reaction. In the nuclease, the observed geometrical arrangement of the guanidyl groups around the 5′-phosphate of the inhibitor appears to be consistent with such a notion.[19,20]

Including the hydrogen bonds, the bonding array around phosphate oxygens O_1 and O_3 (four bonds) is essentially tetrahedral, around O_4 midway between a tetrahedron and a trigonal bipyramid. That around O_2 (two or three bonds) is irregular.

The Methylguanidinium Cations. The structures of these monoalkyl guanidinium cations present the same features, *viz.*, the essential planarity of the guanidyl group, the C_2–N distances grouped around 1.32–1.33 Å, and the C_1–N distances of 1.44–1.45 Å found and discussed by us[23] and others[30–32] in earlier publications. The methyl groups of the cations deviate somewhat more (0.36 and 0.14 Å *vs.* 0.07 Å) from the mean plane of the guanidyl moiety than was found in the structure of methylguanidinium dihydrogen phosphate.[23]

The Monohydrogen Phosphate Anion. The structure of guanidinium dihydrogen phosphate[23] can also be viewed as a complex hydrogen-bonded system surrounding a central dimer of phosphate ions. In the dihydrogen phosphate structure, each phosphate oxygen is involved in two hydrogen bonds; in the monohydrogen phosphate structure, three hydrogen bonds

(31) K. Oaki, T. Nagano, and Y. Iitaka, *Acta Crystallogr., Sect. B*, **27**, 11 (1971).
(32) W. Saenger and K. G. Wagner, *Acta Crystallogr., Sect. B*, **28**, 2237 (1972).

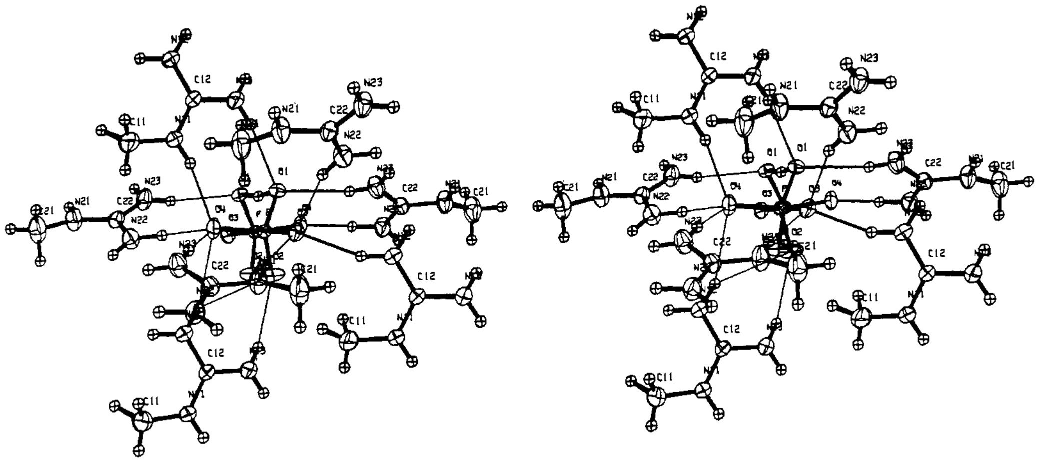


Figure 3. An ORTEP stereo drawing illustrating all the modes of hydrogen bonding interaction of the methylguanidinium cations with the paired monohydrogen phosphate anions.

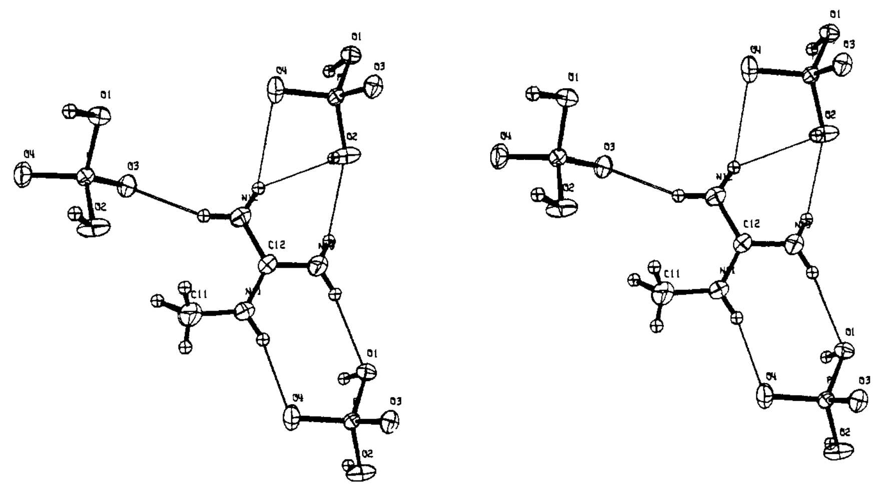


Figure 4. An ORTEP stereo drawing showing the hydrogen bonding pattern for methylguanidinium cation I.

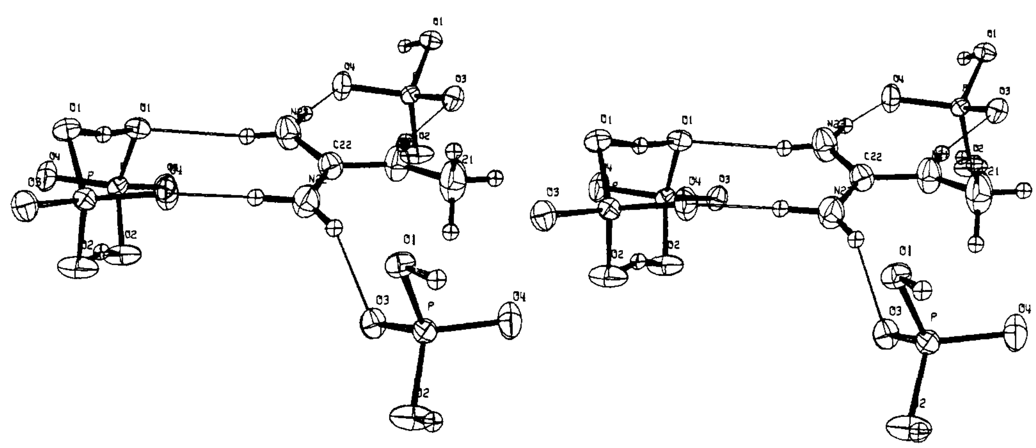


Figure 5. An ORTEP stereo drawing showing the hydrogen bonding pattern for methylguanidinium cation II.

per phosphate oxygen are the norm. Usually a monohydrogen phosphate ion would be expected to show one relatively long P–O distance for the P–O–H bond and a family of three shorter distances for the remaining P–O distances.[33] However, the HPO_4^{2-} ion in this structure shows a P–O bond length pattern of two long and two short P–O bonds, a pattern found in the phosphates of methylguanidinium dihydrogen phosphate and most other $(RO)_2PO_2^-$ compounds where R = H, alkyl, or aryl groups.[23,33] Presumably this mildly anomalous P–O bond-length distribution in the HPO_4^{2-} ions of this structure is due to the very extensive system of guanidinium–phosphate H bonding interactions influencing the distribution of electrons in the phosphate anions. As we have suggested,[1] such effects may

(33) D. E. C. Corbridge, *Bull. Soc. Fr. Mineral. Cristallogr.*, **94**, 271 (1971), is a recent review and extensive tabulation of phosphate structures.

well be of significance in enzymic catalysis where arginines are present at the active site.[34]

Supplementary Material Available. A listing of structure factor amplitudes and Tables VII–IX will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 24× reduction, negatives) containing all of the supplementary material for the papers in this issue may be obtained from the Journals Department, American Chemical Society, 1155 16th St., N.W., Washington, D. C. 20036. Remit check or money order for $3.00 for photocopy or $2.00 for microfiche, referring to code number JACS-74-4471.

(34) M. Legg and T. LaCour in this laboratory have obtained preliminary results which indicate that guanidinium ions affect, but decrease, the rate of hydrolysis of *p*-nitrophenyl phosphate in alkaline ethanol–water solutions. We also have evidence that a two guanidinium–one ester complex forms in these solutions.

Stability–Selectivity Relationships for Solvolytic Displacement Reactions[1]

J. Milton Harris,*[2] Donald C. Clark, Allyn Becker, and James F. Fagan

Contribution from the Department of Chemistry, California State University—Fullerton, Fullerton, California 92634. *Received December 10, 1973*

Abstract: A general theory incorporating concepts of ion pairing is presented for the origin of stability–selectivity relationships for solvolytic displacement reactions on alkyl derivatives. It is concluded from a study of alkyl chloride ethanolysis and examination of previous work that these relationships result from a blending of selectivities for attack on neutral substrate (I), tight ion pair (II), solvent-separated ion pair (III), and free cation (IV); the selectivities for attack on all of these species except solvent-separated ion pair are indicated to be constant despite changes in stability. The effect of a change in intermediate stability is to shift the equilibria between I, II, III, and IV and to change the selectivity of III if it is involved. The possibility of solvent sorting about transient carbocation intermediates is also considered and concluded to be unimportant for alkyl chlorides in aqueous ethanol.

It is generally recognized that an increase in stability of an intermediate produces an increase in selectivity of the intermediate.[3–5] Despite the apparent simplicity of this principle, the origin of stability–selectivity relationships for solvolytic reactions of alkyl esters and halides is poorly understood.[6,7] This difficulty is primarily due to the involvement in solvolysis reactions of more than one type of intermediate. Recent works have shown the solvolysis reaction to be best described as in Scheme I.[5,7–14] The various species I, II, III, and IV represent neutral substrate, tight ion pair, solvent-separated ion pair, and free carbocation, respectively. Actually, there is evidence for the importance of even more ion-pair intermediates than those shown in Scheme I (*e.g.*, nucleophilically solvated ion pairs).[11,14b] Scheme I therefore represents the simplest possible combination. If these intermediates are involved, observed selectivities will be a blend of selectivities for attack on each of the intermediates.

Scheme I

$$\begin{array}{ccccccc}
\underset{\text{I}}{\mathrm{R{-}X}} & \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} & \underset{\text{II}}{\mathrm{R^+X^-}} & \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} & \underset{\text{III}}{\mathrm{R^+\,\|\,X^-}} & \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} & \underset{\text{IV}}{\mathrm{R^+ + X^-}} \\
k_s^{\mathrm{I}}\downarrow\mathrm{Nuc} & & k_s^{\mathrm{II}}\downarrow\mathrm{Nuc} & & k_s^{\mathrm{III}}\downarrow\mathrm{Nuc} & & k_s^{\mathrm{IV}}\downarrow\mathrm{Nuc} \\
\mathrm{R{-}Nuc^+} & & \mathrm{R{-}Nuc^+} & & \mathrm{R{-}Nuc^+} & & \mathrm{R{-}Nuc^+}
\end{array}$$

Sneen[4] and Schleyer[5] have demonstrated the occurrence of stability–selectivity relationships for azide and water attack on intermediates from the solvolysis of alkyl halides.[4,5] Sneen[4] considered the intermediates for his series of compounds to be free carbocations; actually, as we will discuss in the following sections, there is extensive evidence that many of these compounds lead to products by nucleophilic attack on solvent-separated ion pairs. Schleyer and his coworkers[5] did not specify the exact nature of the intermediates for their series, but many of the compounds included were also included on Sneen's plot.

Ritchie has recently made the suggestion that stability–selectivity relationships for solvolytic reactions are determined by a blending of selectivities for solvent-separated ion pair (III) and free carbocation (IV) but details of the hypothesis were not given.[6] Ritchie has also shown the selectivity of triarylmethyl, tropylium, and aryldiazonium cations to be independent of stability.[7] This remarkable observation emphasizes the inadequacies of the present ill-defined theories of the relationship between stability and selectivity.

The purpose of this article is to present a theory for

(1) A preliminary account of this work has appeared in communication form: J. M. Harris, A. Becker, D. C. Clark, J. F. Fagan, and S. L. Kennan, *Tetrahedron Lett.*, 3813 (1973).
(2) Address correspondence to this author at the Department of Chemistry, University of Alabama–Huntsville, Huntsville, Ala. 35807.

(3) J. March, "Advanced Organic Chemistry," McGraw-Hill, New York, N. Y., 1968, pp 164, 394, 526.
(4) R. A. Sneen, J. V. Carter, and P. S. Kay, *J. Amer. Chem. Soc.*, **88**, 2594 (1966).
(5) D. J. Raber, J. M. Harris, R. E. Hall, and P. v. R. Schleyer, *J. Amer. Chem. Soc.*, **93**, 4821 (1971).
(6) C. D. Ritchie, *J. Amer. Chem. Soc.*, **93**, 7324 (1971).
(7) C. D. Ritchie, *Accounts Chem. Res.*, **5**, 348 (1972).
(8) S. Winstein, B. Appel, R. Baker, and A. Diaz, *Chem. Soc., Spec. Publ.*, No. **19**, 109 (1965).
(9) D. J. Raber, J. M. Harris, and P. v. R. Schleyer in "Ion Pairs," M. Szwarc, Ed., Wiley, New York, N. Y., in press.
(10) R. A. Sneen, *Accounts Chem. Res.*, **6**, 46 (1973).
(11) J. M. Harris, *Progr. Phys. Org. Chem.*, **11**, 89 (1974).
(12) V. J. Shiner, Jr., and R. D. Fisher, *J. Amer. Chem. Soc.*, **93**, 2553 (1971).
(13) H. L. Goering and H. Hopf, *J. Amer. Chem. Soc.*, **93**, 1224 (1971).
(14) (a) C. D. Ritchie in "Solute-Solvent Interaction," J. F. Coetzee and C. D. Ritchie, Ed., Marcel Dekker, New York, N. Y., 1969, Chapter 4. (b) F. G. Bordwell, *Accounts Chem. Res.*, **5**, 374 (1972).