Stereoelectronic Control in Carbon–Oxygen and Phosphorus–Oxygen Bond Breaking Processes. Ab Initio Calculations and Speculations on the Mechanism of Action of Ribonuclease A, Staphylococcal Nuclease, and Lysozyme

David G. Gorenstein,*[1] John B. Findlay, Bruce A. Luxon, and Debojyoti Kar

Contribution from the Department of Chemistry, University of Illinois, Chicago Circle, Chicago, Illinois 60680. *Received October 29, 1976*

Abstract: Ab initio molecular orbital calculations on dimethyl phosphate monoanion and dimethoxymethane demonstrate that antiperiplanar ester oxygen lone pairs significantly perturb the X–O (X = P or C) bond overlap populations. Specifically molecules in a gauche,trans conformation are "activated" for cleavage of the trans X–O bond. Other consequences of this torsional activation are discussed and presented in the context of a mechanism of action of ribonuclease A, staphylococcal nuclease, and lysozyme.

The role of orbital orientation in organic and enzymatic reactions has been of considerable current interest.[2-4] Deslongchamps and coworkers[2] in studying tetracovalent carbon species have recently demonstrated selective cleavage of bonds which are trans-antiperiplanar (app) to lone pairs on directly bonded oxygen and nitrogen atoms. Molecular orbital calculations have provided theoretical justification for these stereoelectronic effects in tetracovalent carbon species and neutral phosphate molecules.[4-6]

We wish to demonstrate that torsional activation of specific bonds occurs in phosphate diester monoanions as well and to examine the geometric consequences of this stereoelectronic control in the mechanism of action of the RNA hydrolyzing enzyme, ribonuclease A (RNase A), and the DNA/RNA hydrolyzing enzyme, staphylococcal nuclease (Nase).

In addition we wish to present ab initio calculations on torsional effects in acetals and discuss the significance of these results in terms of a mechanism of action for lysozyme.

Methods of Calculation

The SCF LCAO-MO ab initio calculations utilized the GAUSSIAN 70 series of programs with a STO-3G minimal basis set[7] (no 3d polarization functions on phosphorus).

Initial structures for dimethyl phosphate monoanion (DMP) and dimethoxymethane (DMM) were modeled on the basis of x-ray crystallographic[8-10] structures of appropriate molecules. For each conformation of DMM, the O–C–O bond angle θ and the H–C–H bond angle, ϕ, were optimized. C_{2v} symmetry was assumed for the carbon tetrahedron and hence the O–C–H bond angle was automatically adjusted. The final geometries for DMM in the three staggered conformations are given in Table I (see Figure 1 for a definition of the geometrical parameters).

The molecular geometry of DMP was optimized by sequentially varying a set of geometrical parameters until the total energy had been effectively minimized. Key parameters were selected from the initial idealized geometry and optimized iteratively until it was clear that only very small (<0.02 kcal/mol) decreases in energy were possible by further calculation. For this series of programs it has been suggested[11] that the computational error in the calculated total energy is on the order of 10^{-5} hartree for molecules about as large as benzene.

It is well known that the values of optimized scale factors (and therefore charge distributions) can vary a great deal for a given atom in different molecules,[11] so before optimizing the geometry it was decided to optimize the scale factors of the atoms in the phosphate tetrahedron. Only the valence shell scale factors were optimized and then sequentially in three groups—phosphorus 3sp, ester oxygen 2sp, and phosphoryl oxygen 2sp. This was also done in cycles as above, so that when reoptimized those scale factors that had been optimized in the first group would reflect the influence of the now optimized scale factors in successive groups. The cycles were repeated until differences between the scale factors as well as the total energy from one cycle to the next were not significant. Scale factor optimization was conducted on the gauche,gauche conformation exclusively once we had determined that the differences between scale factors for a given atom in different conformations at an identical level of optimization were not significant (e.g., the maximum difference encountered between a scale factor on gauche,gauche and gauche,trans after two complete cycles of optimization was <0.01). The optimized scale factors are $\zeta_{2sp}(O(1)) = 2.127$, $\zeta_{2sp}(O(3)) = 2.206$, and $\zeta_{3sp}(P) = 2.033$. It was at this point that the optimization of geometry was initiated. The sequence of geometry optimization was: (1) O–P–O bond angle, θ; (2) ester P–OR bond lengths, r_1; (3) phosphoryl O–P–O bond angle, ϕ; (4) phosphoryl P–O bond lengths, r_3; (5) θ again; (6) C–O–P bond angle, ψ; and finally (7) the RO–POR torsional angles ω,ω'. The final geometries are given in Table II.

It is important to note that the relative energy differences between the conformations of DMM and DMP alter dramatically during the optimization cycles. Thus with geometry optimization (particularly of the θ bond angle) the energy differences between the three staggered conformations of DMP is <1 kcal/mol[10] rather than ~7 kcal/mol previously reported.[12]

All calculations were carried out on a IBM 370/158 computer. The program was compiled using the Fortran H compiler at an optimization level of 2. The program was not overlayed after it was determined that less CPU time was required during a given run if a full 800K of core was allocated for the computation.

Results

Energies and population analysis from the ab initio molecular orbital calculations on DMP and DMM in gauche,gauche (g,g), gauche,trans (g,t), and trans,trans (t,t) conformations are shown in Tables III and IV. The RO–X–OR (X = C or P) bond angle, θ, in DMM and DMP has been optimized, since large bond angle distortions are found to be coupled to the

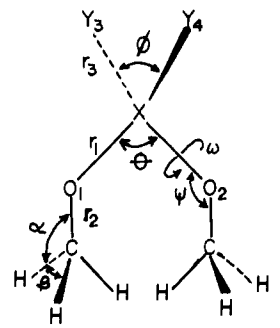


Figure 1. Definition of geometrical parameters for DMP (X = P, Y = O) and DMM (X = C, Y = H).

Table I. Geometries for the Staggered Torsional Conformations of Dimethoxymethane

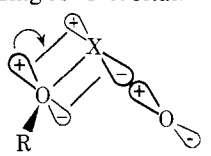
Geometry[a,b]	g,g	g,t	t,t
r_1	1.40	1.40	1.40
r_2	1.43	1.43[c,d]	1.43
r_3[e]	1.090	1.090	1.090
r_{C-H}[e]	1.090	1.090	1.090
θ	114.66	109.36	103.73
ϕ	108.36	107.53	106.85
ψ	114.0	114.0	114.0
ω, ω'	60, 60	60,[c] 180[d]	180, 180
α[e]	109.5	109.5	109.5
β[e]	109.5	109.5	109.5

[a] See Figure 1 for definition of geometrical parameters. C_{2v} symmetry for carbon tetrahedron. [b] Bond lengths in angstroms, angles in degrees. [c] Gauche bond. [d] Trans bond. [e] Assumed.

torsional conformations.[8,10,13] Although a nearly complete geometry optimization has been performed for DMP, the relative P–O overlap populations and conclusions about app lone-pair interactions are largely unaffected by any other geometrical parameter except for the ester torsional conformation.[15]

Discussion

As has been found in previous molecular orbital calculations[4-6] on the $\ddot{X}(1)$–Y–$\ddot{X}(2)$ (Y = C, P; X = O, N) structural fragments, the X(1)–Y bond is strengthened (as indicated by an increase in the Mulliken overlap population[14]), while the Y–X(2) bond is weakened when the X(1) atom lone pair is app to the Y–X(2) bond. Thus, in the g,t conformation of DMP the overlap population for the trans P–O bond is 0.017 e lower than the overlap population for the gauche P–O bond (Table IV). In the g,t conformation of dimethoxymethane, the overlap population of the trans bond is 0.022 e lower than the overlap population of the gauche bond. As shown in Figure 2 for DMP, one lone pair on the gauche bond oxygen is app to the trans bond, while no lone pairs on the trans bond oxygen are app to the gauche bond. Thus, the weakest X(1)–Y bond has one app lone pair and no lone pairs on X(1) app to the Y–X(2) bond. CNDO calculations confirm this conclusion, with the bond order for the trans bond in the g,t conformation of DMP being 1.06 and the bond order for the gauche bond being 1.09. Lehn[4] and Pople[5] and co-workers have shown similar overlap population differences between the gauche and trans bonds in related systems. As suggested by others[4-6,16] the torsional effect is likely attributable to interaction of an app oxygen lone pair with the σ^* antibonding X–O orbital.



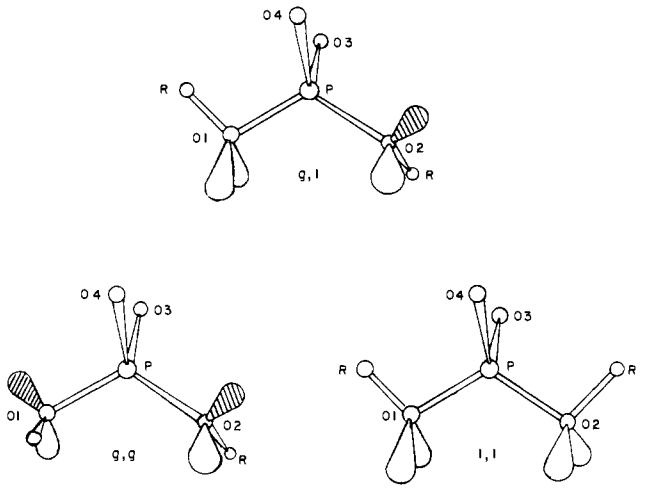


Figure 2. Structure of dimethyl phosphate in various conformations. App lone pairs to ester bonds are shaded.

Table II. Geometries for the Staggered Torsional Conformations of Dimethyl Phosphate Monoanion

Geometry[a,b]	g,g	g,t	t,t
r_1	1.667	1.666,[c] 1.667[d]	1.664
r_2	1.43	1.43	1.43
r_3	1.538	1.538,[c] 1.538[d]	1.539
r_{C-H}[e]	1.09	1.09	1.09
θ	98.81	94.90	91.04
ϕ	125.66	124.30	123.08
ψ	112.41	111.75	110.66
ω, ω'	68.2	74.6,[c] 179.2[d]	180.0
α[e]	109.5	109.5	109.5
β[e]	109.5	109.5	109.5

[a-e] See footnotes from Table I.

Table III. Ab Initio Energies and Population Analysis of Dimethyl Phosphate Monoanion

	Conformation		
	g,g	g,t	t,t
Rel energy,[a] kcal/mol	0.0	0.14	0.88
Bond	Overlap Population		
P–O(1)	0.3348	0.3309[b]	0.3419
P–O(2)	0.3348	0.3479[c]	0.3419
P–O(3)	0.4800	0.4694	0.4718
P–O(4)	0.4800	0.4754	0.4718
	Atomic Charges		
P	1.723	1.708	1.719
O(1)	−0.539	−0.537[b]	−0.526
O(2)	−0.539	−0.530[c]	−0.526
O(3)	−0.920	−0.926	−0.927
O(4)	−0.920	−0.912	−0.927

[a] Total energy g,g conformation, −710.41666 hartree. [b] Trans bond. [c] Gauche bond.

Actually interpretation of the overlap population differences in Tables III and IV is slightly more complicated than the foregoing analysis indicates, since all four lone pairs on the gauche or trans oxygens are *always* app in any of the three staggered conformations to some other bond. Thus, in DMP, a lone pair will be app to either the phosphoryl or ester bond in the g,g, g,t, or t,t conformations.

Table IV. Ab Initio Energies and Population Analysis of Dimethoxymethane

	Conformation		
	g,g	g,t	t,t
Rel energy, kcal/mol[a]	0.0	0.9	2.5
Overlap Population			
Bond			
C(1)-O(4)	0.5402	0.5524[b]	0.5448
C(1)-O(5)	0.5402	0.5304[c]	0.5448
C(1)-H(2)	0.7480	0.7336	0.7346
C(1)-H(3)	0.7480	0.7492	0.7346
Atomic Charges			
C(1)	0.1359	0.1386	0.1417
O(4)	−0.2731	−0.2573	−0.2554
O(5)	−0.2731	−0.2689	−0.2554

[a] Total energy for g,g conformation, −264.5483 hartree. [b] Gauche bond. [c] Trans bond.

Table V. Ab Initio Energies and Population Analysis of Difluorodimethoxymethane[a]

	Conformation		
	g,g	g,t	t,t
Rel energy,[b] kcal/mol	0.0	0.25	3.17
Overlap Population			
Bond			
C(1)-O(4)	0.5112	0.5252[c]	0.5307
C(1)-O(5)	0.5112	0.5145[d]	0.5307
C(1)-F(2)	0.4245	0.4113	0.4093
C(1)-F(3)	0.4245	0.4224	0.4093

[a] Structures used are the same as for the final optimized geometries for dimethoxymethane (Table I). Fluorine bond lengths have been separately optimized, however. [b] Total energy for the gauche conformation, −459.4934 hartree. [c] Gauche bond. [d] Trans bond.

We have devised a system of notation for dimethyl phosphate (easily extendable to DMM or other molecules) which clearly identifies the important antiperiplanar interactions. An "ordered triplet" ($\alpha,\beta;\gamma$) for a particular bond is interpreted as follows: α = the number of lone pairs on the oxygens of a P–O ester bond which are antiperiplanar to the adjacent P–O ester bond; β = the number of lone pairs on the oxygens of a P–O ester bond which are antiperiplanar to an adjacent P–O phosphoryl bond; γ = the number of lone pairs on ester oxygens which are app to the bond in question. The sum ($\alpha + \beta$) represents the total bond-strengthening interactions for the bond in question; by contrast γ represents bond weakening.

Using this convention, we can quickly identify the type and number of app interactions and compare the predicted strengths of the ester bonds in the three staggered comformers of DMP (Table V). Inspection of Tables III and IV indicates that the relative bond strengths follow the pattern:

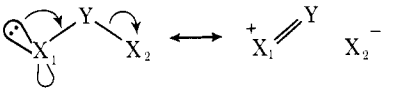
$$\text{g,t (gauche bond)} \gtrsim \text{t,t} \gg \text{g,g} \gtrsim \text{g,t (trans bond)}$$

or in ordered triplet notation:

$$(1,1;0) \gtrsim (0,2;0) \gg (1,1;1) \gtrsim (0,2;1)$$

The gauche bond in the g,t conformation and the trans bonds in the t,t conformation are approximately equally strong, while the trans bond in the g,t conformation and the gauche bonds in the g,g conformation are approximately equally weak. This suggests that the main difference between the four types of bonds is the number of bond-weakening (γ) interactions. The stronger bonds have no γ interactions, while the weaker bonds have one γ interaction. The staggered bonds will always have two bond-strengthening ($\alpha + \beta$) interactions.

Subtle variations in the overlap populations are also informative. Interestingly, for both DMP and DMM, the gauche bond in the g,t conformation is slightly stronger than the trans bonds in the t,t conformation, although the number of γ interactions is the same (0). Similarly the trans bond in the g,t conformation is slightly weaker than the gauche bonds in the g,g conformation, although again the number of γ interactions is the same (1). The slight secondary bond-strengthening (and weakening) effect is attributed to a real distinction between the α and β effects. Apparently an α bond-strengthening effect is slightly more important than a β bond-strengthening effect in these molecules. Assuming that the app effect arises from the interaction of the lone-pair orbital with the adjacent σ^* orbital, then this interaction will be enhanced in those bonds possessing a more electronegative atom. In terms of a simple resonance picture for the app effect, the contribution of resonance structure II is expected to increase with increasing



I II

electronegativity for atom X(2). In DMP the ester oxygen is more electronegative than the phosphoryl oxygen and hence α interactions are greater than β interactions. Similarly, in DMM the acetal oxygen is more electronegative than the hydrogen and again the α interaction is shown to be slightly more important than the β interaction. These conclusions are confirmed by the smaller overlap population differences for the phosphoryl and C–H bonds.

A further confirmation that electronegativity differences are responsible for these secondary effects is provided by an analysis of the overlap populations in difluorodimethoxymethane. As shown in Table V, the order of decreasing bond strength is

$$\text{t,t} > \text{g,t (gauche bond)} \gg \text{g,t (trans bonds)} > \text{g,g}$$

or in ordered triplet notation,

$$(0,2;0) > (1,1;0) \gg (0,2;1) > (1,1;1)$$

Now the β bond-strengthening interaction is *more* important than the α bond-strengthening interaction. The reason is that the app lone-pair interaction with the more electronegative fluorine atoms (β effect) is more pronounced than the app lone-pair interaction with the methoxyl oxygen (α effect). Note that by differentially protonating the phosphoryl oxygens and hence altering the relative electronegativities of these oxygens it is possible to further enhance the reactivity of specific bonds by torsional catalysis.

These calculations suggest that the g,t conformation should be the most hydrolytically reactive conformation, and that most significantly, the trans bond should be specifically weakened at the expense of a strengthened gauche bond.[17]

Mechanism of Action of RNase A. While unfortunately no direct test of these predictions exists for a phosphate diester, it is most significant that pancreatic RNase A preferentially cleaves yeast phenylalanyl transfer RNA (tRNA) at only two positions along the chain.[19] One point of attack is the internucleotide linkage 74–75 which, being at the end of the chain, is relatively flexible, and the other is the linkage between uridine-33 and $O^{2'}$-methylguanosine-34. The x-ray structure[20,21] shows this internucleotide phosphate in a g,t conformation. Although other conformations for phosphate ester bonds with 3′-pyrimidine bases are found at the surface of the tRNA

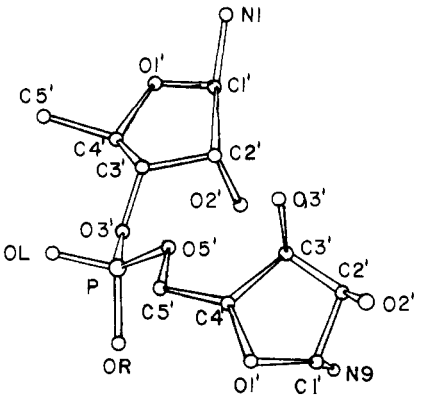


Figure 3. Structure of U^{33}-2′-OMeG^{34} dinucleotide portion of phenylalanyl tRNA. Only the diribose phosphate is shown.

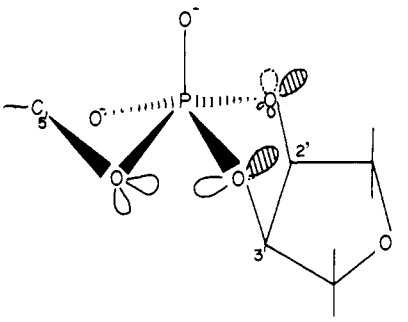
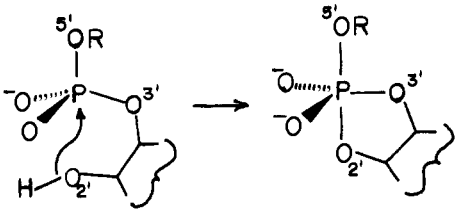


Figure 4. Possible tetragonal pyramid transition state for the transesterification step of RNase A. Lone pairs on O(3′) and O(2′) app to P–O(5′) are shaded. Only slight distortion (≤15°) of the bond angles is required to convert this structure into an "in-line" basal-in, basal-out trigonal bypyramid transition state (apex P–O, O(2′), and O(5′) in the basal plane with O(3′) and the other P–O apical).

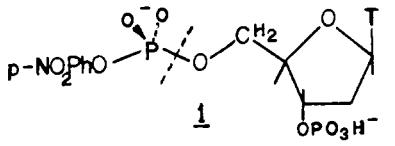
molecule, only this conformationally restricted diester is attacked by RNase A. (For example $C^{56}G^{57}$, $U^{68}U^{69}$, $U^{69}U^{70}$, $C^{70}G^{71}$ internucleotide phosphates in the double helical acceptor stem and T ψ C loop are potentially quite accessable to RNase A, yet none of these g,g diesters are cleaved by the enzyme.) Furthermore, as shown in Figure 3 it is the *5′ bond* of the dinucleotide that is in the trans conformation and the 3′ bond that is in the gauche conformation. Since only the 5′ bond is ever cleaved by RNase A, the enzyme has fully taken advantage of the stereoelectronic consequences of the g,t conformation. Not only is P–O(5′) weakened and P–O(3′) strengthened (thus preventing its breaking), but also the electron denslty on the 5′-oxygen is higher in this conformation than on the 3′-oxygen (Table III). Protonation by one of the active site histidines (presumably His-119)[22] would thus be facilitated. In addition, the g,t conformation is slightly higher in energy than the lowest energy g,g conformation, which should induce strain in the substrate and hence further accelerate the enzymatic reaction. Finally, the optimized ester RO–P–OR bond angle in the g,t conformation is 5–6° smaller than in the g,g conformation and therefore distortion in the direction of the trigonal bipyramid transition-state geometry (RO–P–OR bond angle, 104° → 90°) is achieved. Note that the bond angle is also closer to that of the 2′,3′-cyclic nucleotides (96°),[23] where the reduced bond angle is likely an important factor in the 10^6 times faster hydrolysis of the strained five-membered cyclic phosphates.[24] All of these factors probably contribute to the enormously enhanced hydrolysis of the acyclic diesters by RNase A.

One additional point must be considered in this mechanism since the tRNA, U^{33}-2′-OMeG^{34} dinucleotide fragment is not in the correct conformation for an "in-line" or backside attack by the 2′-hydroxyl group as appears required in the RNase A catalyzed reaction.[25] However, rotation of the uridine-33 ribose



ring by ca. 180° about a rough axis through the C(5′)–O(5′) and C(3′)–O(3′) bonds does line the 2′-hydroxyl group backside to the departing 2′-OMeG^{34} 5′-oxygen. After this rotation of the ribose ring (presumably RNase A mediated) the dinucleotide conformation is comparable to the x-ray structure of a dinucleotide analogue bound to RNase A.[22] However, the geometry of a 2′-O and 5′-O diaxial cyclic five-membered ring pentacovalent transition state diminishes the overlap between the app 3′-O lone pair and the 5′-O bond. Stereoelectronically, the best transition state is one in which the 2′-O attacks basal and the 3′-O and 5′-O take up apical and basal positions, respectively. Now app lone pairs on *both* the 2′-O and 3′-O will further weaken the 5′-O bond *and* strengthen the incipient 2′-O bond. Although this transition state ensures an in-line stereochemistry, present theory suggests that a basal attack and basal displacement mechanism is less likely than apical attack and apical leaving mechanisms.[18,24] Alternatively, with a mechanism involving a pentacovalent square pyramid transition state[24] it is possible to simultaneously observe the requirement of approximate g,t stereochemistry for the 3′ and 5′ bonds, respectively, *and* of the requirement for "in-line" stereochemistry for the attacking 2′-oxygen (Figure 4). An additional advantage implicit in a tetragonal pyramid transition state is the additional nearly app lone pair on the 2′-oxygen, which both helps strengthen the developing 2′-O–P bond and further weaken the 5′-O–P bond. Significantly, recent x-ray studies have revealed that pentacovalent phosphoranes can exist in the tetragonal pyramid as well as the traditional trigonal bipyramid geometries.[26]

Mechanism of Action of Staphylococcal Nuclease. Staphylococcal nuclease (Nase) is both a DNA and RNA hydrolyzing enzyme which, like RNase A, specifically cleaves the 5′-oxygen ester bond.[27] However, unlike RNase A, a cyclic diester intermediate is not a prerequisite (and stereochemically complicating) feature of its mechanism of action. A rather dramatic demonstration of the specificity of Nase for cleaving only the 5′-bond is shown by the Nase-catalyzed hydrolysis of thymidine-3′-phosphate, 5′-*p*-nitrophenyl phosphate, **1**, which yields *only* *p*-nitrophenyl phosphate and 3′-thymidine monophosphate.[29]



1

This contrasts dramatically with the nonenzymatic hydrolysis of this substrate, which because *p*-nitrophenoxide is a much better leaving group than the 5′-alkoxide, yields thymidine 3′,5′-diphosphate and *p*-nitrophenol.[29] While certainly other explanations are possible, one way in which the enzyme can achieve this specificity is to take advantage of the torsional activation *and* stabilization of the different ester bonds of a diester constrained to a g,t conformation.

It is significant, therefore, that based upon the Nase x-ray structural results of Cotton and co-workers,[27,28] we would conclude that, in fact, a phosphate diester substrate likely binds to the enzyme in a g,t conformation, with the 5′ bond in the trans conformation. Our conclusions are based upon the x-ray

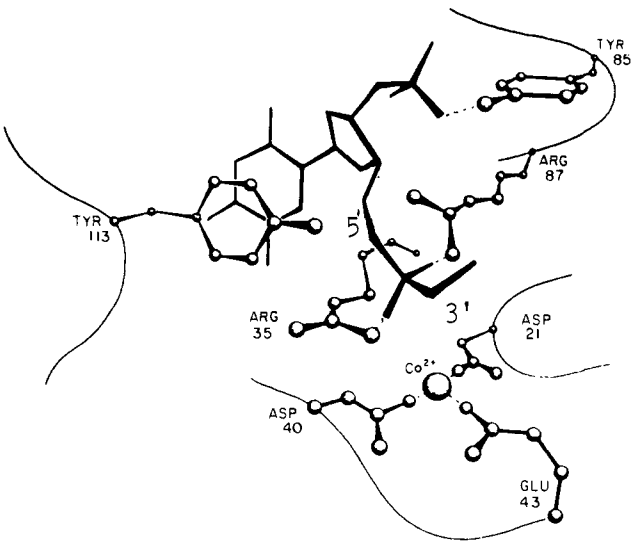


Figure 5. Structure of the active site of staphylococcal nuclease and bound 3',5'-pTp (adapted from ref 27).

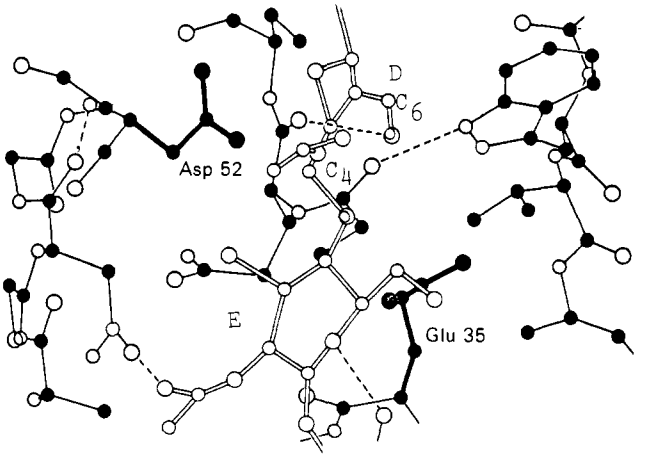
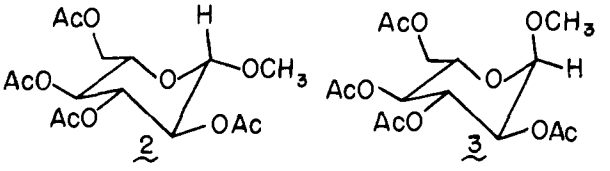


Figure 6. Structure of part of the active site of lysozyme, including rings D and E of a hexasaccharide substrate. Ring D is shown in a sofa-like conformation (adapted from ref 33).

results for the pTp Nase complex shown in Figure 5. As pointed out by Cotton, the rigidity and specificity of binding the 5'-phosphate is rather remarkable. Two arginine residues hydrogen bond to two of the oxygens of the 5'-phosphate. This suggests that the remaining oxygen would normally be in the position of the 3'-ester oxygen of a natural diester substrate. Included in Figure 5 is an estimate of the allowed conformation for a 3'-ester group based upon an attempt to fit the rest of the DNA chain to the remaining open portion of the active site. The x-ray structure shows that with a slight movement of atoms, the 5'-ester bond can be placed in the trans conformation and we feel they additionally support an interpretation placing a 3'-ester bond in the gauche conformation. This conformation ensures that only the 5' bond will be broken. Torsional catalysis capable of specifying which bond is to be broken may be especially important, since there does not appear to be an acid group at the active site which can function as a general acid catalyst to the 5'-alkoxide leaving group.[29] Although Tyr-113 could serve this purpose, the enzyme is still quite active when this phenolic group is deprotonated, ruling out general acid catalysis, at least by this group.

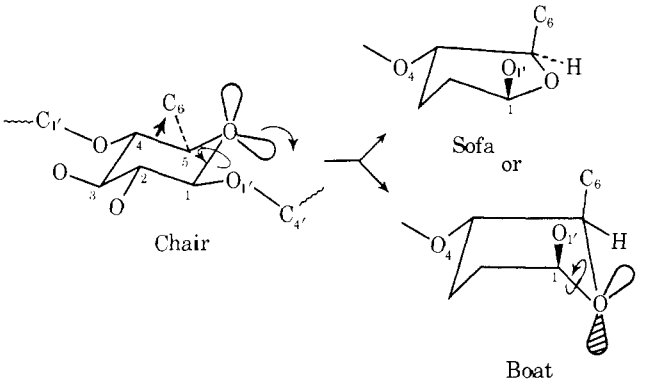
Torsional Catalysis in Acetals. As was true for the phosphate diester, the g,t conformation selectively activates the trans bond for cleavage and strengthens the gauche bond. Thus, as shown in Table IV, the trans bond in the g,t conformation of dimethoxymethane has the smallest overlap population (0.530) of any of the bonds in any of the three staggered conformations. In addition, the gauche bond in the g,t conformation has the largest overlap population (0.552) of any of the C-O bonds. Both Lehn[4,30] and Pople[5] have demonstrated similar effects. In fact, Lehn's ab initio calculations on dihydroxymethane show even larger overlap population differences between the two C-o bonds in the g,t conformation. The trans bond overlap population is reduced by 0.08 e in Lehn's calculation, but only 0.02 e in ours. The differences reflect the known sensitivity of Mulliken overlap populations to the choice of basis sets. Whereas Lehn's calculations utilized a larger basis set and assumed tetrahedral bond angles about carbon, ours have included O-C-O (and H-C-H) bond angle optimization about the anomeric carbon, which we have previously shown has a dramatic effect upon the relative energies of the various conformations and the magnitude of the anomeric effect. Lehn's and Pople's calculations do show that the weaker, trans bond is longer than the gauche bond, thus reflecting the reduced overlap population for the trans bond.

Deslongchamps[2,32] has provided a striking experimental verification of the influence of app lone pairs on reactivity. Thus, the ozone oxidation of the β-anomer of the methyl glycoside, **2**, reacts within minutes at room temperature, while the α-anomer, **3**, is completely unreactive under the same conditions. Deslongchamps proposed, prior to any theoretical calculations, that the difference in reactivity between **2** and **3** is due to the number of lone pairs on the glycosidic oxygens which are app to the anomeric C-H bond being oxidized. Assuming a favorable conformation about the aglycone, C-OCH_3, bond (either −g or t), the axial anomeric C-H bond in **2** is app to two oxygen lone pairs. In **3** only a single lone pair on the aglycone bond can be app to the equatorial C-H bond. The additional app lone pair in **2** must be responsible for the substantially increased reactivity.

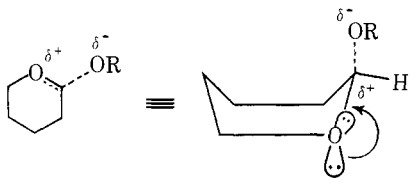


The ab initio calculations in Table IV support these observations, snce the larger C-H overlap populations occur in those conformations with only one oxygen lone pair app to the C-H bond. This exists for both C-H bonds in the g,g conformation and the C(1)-H(3) bond in the g,t conformation. In contrast those anomeric C-H bonds app to two oxygen lone pairs have a smaller overlap population, that is, both C-H's in the t,t conformation and C(1)-H(2) in the g,t conformation.

Mechanism of Action of Lysozyme. This torsional catalysis principle is important to an understanding of the mechanism of action of lysozyme, an enzyme which cleaves the aglycone bond in saccharides. The x-ray studies of Phillips and coworkers have demonstrated that distortion of the reacting sugar residue occurs upon binding[33] (Figure 6). The enzyme cleaves the anomeric bond of the *N*-acetylglucosamine residue that binds to the D subsite (a total of six subsites are thought to comprise the total active site). In the commonly accepted mechanism for lysozyme, the enzyme forces the sugar ring in subsite D into a half-chair, or in later interpretations, into a sofa or boat conformation. Cleavage of the "strained" aglycone bond with general acid catalysis by Glu-35 yields an oxycarbonium ion in the ring which is stabilized by the adjacent negative charge of Asp-52. Much has been made of the strain or distortion catalysis in the enzymatic mechanism, though Warshel and Levitt[34] have recently claimed that their calculations suggest the distortion factor is minimal.



Binding studies[35] and x-ray studies on a tetrasaccharide lactone "transition-state" analogue leave little doubt that distortion does occur, however. All models clearly show that the C(6)-OH group of the sugar ring in subsite D must move out of its equatorial position to reduce severe steric interaction with the enzyme. Phillips argues that this forces the sugar ring into the sofa or boat conformation, with the C(6)-OH group moving into an axial position. In the chair conformation with a β-(1→4)-linkage the torsional conformation about the C(1)-O(5) bond is trans, and hence according to our predictions the wrong bond is disposed to cleavage. (It is the aglycone C(1)-O(4′) bond that must actually be broken by the enzyme). However, by distorting the ring into a boat conformation, the C(1)-O(4′) aglycone bond moves into an axial position along with the C(6)-OH group (the C(6)-OH group serves as the "arm" for this distortion). By pushing C(1)-O(4′) axial, the conformation about the C(1)-O(5) bond moves from trans to gauche. In addition, inspection of the x-ray model of Phillips as reproduced in Figure 6 shows that the C(1)-O(4′) bond is in a trans conformation. Thus, the activated conformation for the anomeric carbon of a lysozyme bound sugar is g,t, with the bond being broken being the weakened trans bond. These conclusions are approximately equally correct even if the sugar takes on a "sofa" conformation. Since these distorted conformations are ≥6 kcal/mol[34] higher in energy than the unstrained, normal chair conformation, significant catalysis is provided if this strain energy is utilized to move the enzyme–sugar complex structure along the reaction coordinate. Only in the boat or a distorted sofa conformation can an app lone pair on the ring oxygen be used to weaken the C(1)-O(4′) bond. In this way the increased overlap between the developing carbonium ion and the app lone pair on ring O(5) stabilizes the oxycarbonium structure.



Note that without proper app orbital orientation, the carbonium-ion-like transition state would not be able to take advantage of the stabilization of the adjacent oxygen atom. In fact, the electron-withdrawing oxygen atom would destabilize a carbonium ion if proper orbital overlap were prevented. This would be the case in the original, undistorted chair conformation and emphasizes the importance of this torsional catalysis in lysozyme.

These considerations would lead us to predict that the mechanism of action of α-glucosidases in which the C(1)-O(5) bond already exists in the proper gauche conformation probably does *not* involve severe distortion of the chair conformation.

Estimate of the Catalytic Advantage in the Stereoelectronic Control of Enzymatic and Nonenzymatic Reactions. It is important to point out that the enzymatic activation energy lowering by stereoelectronic effects is not simply limited to ground-state *enthalpy* differences between different conformers. Thus, if the g,t conformation is the reactive conformation, based upon enthalpy considerations alone, the enzyme-bound g,t conformational isomer should only react faster than the nonenzyme bound, conformationally flexible molecule by a factor representing the energy difference between the lowest energy conformation and the g,t conformation. The basis for this statement derives from the following argument. Assuming again that the g,t conformation is significantly more reactive than the other conformations, then only the fraction of the total freely interconverting molecules possessing this g,t conformation is reacting in solution. This Boltzmann distribution fraction will be determined by the energy difference between the reactive g,t and nonreactive conformations. Any further stereoelectronic acceleration of the enzymatic reaction can only be achieved by increasing the fraction of the reactive conformation. However, for phosphate diesters, we have calculated[10] that the energy difference between the lowest energy conformation and the higher energy g,t conformation is <1 kcal/mol. Thus in solution the fraction of the reactive g,t esters is already quite high and it would appear that the enzyme cannot take much further catalytic advantage of the stereoelectronic effect. (For lysozyme, though, the enthalpic contribution to the stereoelectronic catalysis mechanism is quite large.)

The actual stereoelectronic advantage in reacting via the g,t conformation is likely to be much larger than indicated by these enthalpic arguments. We have ignored entropic considerations, and any stereoelectronic effect may well be masked by the rather substantial *unfavorable* entropic effects. Assuming again that only the g,t conformation is reactive, then during the course of the reaction, two rotational degrees of freedom about the ester bonds must be lost. The rate of the nonenzymatic reaction will reflect this internal entropic disadvantage, which may be as large as 8 eu (2.4 kcal/mol at 25 °C) per lost degree of freedom.[36] The total 16-eu difference may represent a rate difference of 10^3–10^4. In order for the nonenzymatic reaction to be stereoelectronically controlled, then the enthalpic, stereoelectronic advantage must be large enough to at the very least cancel the 4–5-kcal/mol entropic disadvantage of freezing two rotational degrees of freedom. Presumably an enzyme, however, utilizes intrinsic binding energy to freeze these two bonds and restrict the ester to the stereoelectronically reactive conformation. Hence the full stereoelectronic enthalpic gain is realized without any entropic loss.

How large is this enthalpic stereoelectronic effect? We have recently suggested[37] that a significant portion of the 10^6–10^8-fold rate acceleration of five-membered cyclic phosphates relative to acyclic phosphates derives from the stereoelectronic effect.[38,39] The conformation of the locked cyclic esters is apparently nearly as stereoelectronically productive as the g,t conformation and most significantly no loss of internal rotational entropy is required for reaction. The 12–18 eu less favorable entropy of activation for acyclic vs. cyclic phosphates may well arise from the immobilization of the freely rotating ester bonds. To balance this entropic disadvantage, the enthalpic stereoelectronic advantage must be quite substantial.

It is quite likely that many other enzymes take advantage of the potential stereoelectronic *catalysis* and *specificity* to be gained by binding substrates in the g,t conformation. Mock has recently described other applications of the app lone pair effect to several proteases,[40] and there is little doubt that the principle will have (or at least should have) wide application in the near future.

Acknowledgment. Support of this research by the National Science Foundation (GP 38763), the National Institutes of Health (GM-17575), and the Alfred P. Sloan Foundation is gratefully acknowledged. We also thank the Computer Center of the University of Illinois at Chicago Circle for generous allocation of computing time and T. J. O'Donnell for help with display of the dinucleotide fragment on a Chemistry Department Vector display system. The help of Robert Momii with preliminary CNDO calculations is also appreciated.

References and Notes

(1) Fellow of the Alfred P. Sloan Foundation, 1975–1977.
(2) P. Deslongchamps and R. J. Taillerfer, *Can. J. Chem.*, **53**, 3029 (1975), and references cited therein.
(3) D. R. Storm and D. E. Koshland Jr., *J. Am. Chem. Soc.*, **94**, 5815 (1972).
(4) J. M. Lehn and G. Wipfe, *J. Chem. Soc., Chem. Commun.*, 800 (1975).
(5) (a) L. Radom, W. J. Hehre, and J. A. Pople, *J. Am. Chem. Soc.*, **94**, 2371 (1972); (b) G. A. Jeffrey, J. A. Pople, and L. Radom, *Carbohydr. Res.*, **25**, 117 (1972), have provided further support for these anomeric-type effects by noting that the calculated bond lengths agree with the expected overlap population changes.
(6) S. David, O. Eisenstein, W. J. Hehre, L. Salem, and R. Hoffmann, *J. Am. Chem. Soc.*, **95**, 3806 (1973), and references cited therein.
(7) W. J. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton, and J. A. Pople, Quantum Chemistry Exchange Program, No. 236.
(8) See D. G. Gorenstein and D. Kar, *J. Am. Chem. Soc.*, **98**, 1668 (1976).
(9) D. G. Gorenstein and D. Kar, *J. Am. Chem. Soc.*, **99**, 672 (1977).
(10) D. G. Gorenstein, B. A. Luxon, and J. B. Findlay, *Biochim. Biophys. Acta*, **475**, 184 (1977).
(11) W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.*, **51**, 2659 (1969).
(12) M. D. Newton, *J. Am. Chem. Soc.*, **95**, 256 (1973).
(13) D. Perahia and B. Pullman, *Biochim. Biophys. Acta*, **435**, 282 (1976).
(14) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833 (1955).
(15) Unpublished observations.
(16) C. Romers, C. Altona, H. R. Buys, and E. Havinga, *Top. Stereochem.*, **4**, 39 (1969).
(17) Strictly interpreted, this conclusion should only apply to unimolecular bond-breaking processes. Since the hydrolysis of phosphate diesters probably proceeds via an S_N2-type mechanism, the question of orbital stereoelectronic control in pentacoordinate trigonal bipyramidal species actually must also be considered. However, recent calculations indicate that the app lone pair effect in a trigonal bipyramid transition state is similar to that in the tetracoordinate ground state (work in progress). See also ref 18 and references therein for discussion of related topics in molecular orbital calculations on pentacoordinate phosphorus.
(18) R. Hoffmann, J. M. Howell, and E. L. Muetterties, *J. Am. Chem. Soc.*, **94**, 3047 (1972).
(19) S. H. Chang and V. L. Rajbhandary, *J. Biol. Chem.*, **243**, 592 (1968).
(20) (a) G. J. Quigley, N. C. Seeman, A. H-J. Wang, F. L. Suddath, and A. Rich, *Nucleic Acid Res.*, **2**, 2329 (1975); (b) J. L. Sussman and S.-H. Kim, *Biochem. Biophys. Res. Commun.*, **68**, 89 (1976).
(21) In another crystal modification the conformation is −102°, 167° rather than −52°, −171°[20a] and −33°, 181°.[20b] J. E. Ladner, A. Jack, J. D. Robertus, R. S. Brown, D. Rhodes, B. F. C. Clark, and A. Klug, *Nucleic Acid Res.*, **2**, 1629 (1975).
(22) R. M. Richards and H. W. Wyckoff, "The Enzymes", Vol. IV, P. D. Boyer, Ed., Academic Press, New York, N.Y., 1971, Chapter 24.
(23) C. L. Coulter, *J. Am. Chem. Soc.*, **95**, 570 (1973).
(24) F. H. Westheimer, *Acc. Chem. Res.*, **1**, 70 (1968).
(25) D. A. Usher, D. I. Richardson, and F. Eckstein, *Nature (London)*, **228**, 663 (1970).
(26) J. A. Howard, D. R. Russell, and S. Trippett, *J. Chem. Soc., Chem. Commun.*, 856 (1973).
(27) (a) F. A. Cotton and E. E. Hazen Jr., "The Enzymes", Vol. IV, P. D. Boyer, Ed., Academic Press, New York, N.Y., 1971, Chapter 7; (b) C. B. Anfinsen, P. Cuatrecasas, and H. Taniuchi, *ibid.*, Chapter 8.
(28) F. A. Cotton, C. J. Bier, V. W. Day, E. E. Hazen, and S. Larsen, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 243–249 (1972).
(29) B. M. Dunn, C. D. Bello, and C. B. Anfinsen, *J. Biol. Chem.*, **248**, 4769 (1973).
(30) J. M. Lehn, G. Wipfe, and H.-B. Burgi, *Helv. Chim. Acta*, **57**, 493 (1974).
(31) R. V. Lemieux, *Pure Appl. Chem.*, **25**, 527 (1971), and references cited therein.
(32) P. Deslongchamps, *Tetrahedron*, **31**, 2463 (1975).
(33) L. O. Ford, L. N. Johnson, P. A. Mechin, D. C. Phillips, and R. Tjian, *J. Mol. Biol.*, **88**, 349 (1974).
(34) A. Warshel and M. Levitt, *J. Mol. Biol.*, **193**, 227 (1976).
(35) D. M. Chipman and N. Sharon, *Science*, **165**, 454 (1969).
(36) E. P. Jencks in "Advances in Enzymology", E. Meister, Ed., Wiley, New York, N.Y., 1975, pp 219–410.
(37) D. G. Gorenstein, B. A. Luxon, J. B. Findlay, and R. Momii, submitted for publication.
(38) (a) F. H. Westheimer, *Acc. Chem. Res.*, **1**, 70 (1968); (b) E. A. Dennis and F. H. Westheimer, *J. Am. Chem. Soc.*, **88**, 3431, 3432 (1966).
(39) G. Asknes and K. Bergesen, *Acta. Chem. Scand.*, **20**, 2508 (1966).
(40) W. L. Mock, *Bioorg. Chem.*, **4**, 270 (1970).

Porphyrin–Protein Bond of Cytochrome c_{558} from *Euglena gracilis*

Marvin J. Miller and Henry Rapoport*

Contribution from the Department of Chemistry, University of California, Berkeley, California 94720. *Received October 18, 1976*

Abstract: The atypical cytochrome c_{558} isolated from dark-grown cultures of *Euglena gracilis* contains a single heme to protein thioether bond. Sodium amalgam reductive cleavage of this bond followed by esterification yielded a porphyrin which is identical with synthetic 2-vinyl-4-ethyldeuteroporphyrin IX dimethyl ester and different from its 2-ethyl-4-vinyl isomer. This establishes *Euglena* porphyrin c_{558} as 2-vinyl-4-(α-*S*-cysteinylethyl)deuteroporphyrin IX, and the monothioether linkage in cytochrome c_{558} is at the 4-α-ethyl position of the heme. Implications of this result on the sense of the porphyrin ring relative to the peptide chain in cytochromes *c* with two thioether linkages are considered.

Cytochromes *c* are proteins which contain a covalently bound iron–porphyrin prosthetic group and are active in the electron-transport processes in photosynthesis and respiration.[1] While considerable evidence had suggested that the covalent linkage is through two thioether bonds involving two cysteine residues of the apoprotein, it was not until recently that studies with yeast and horse heart cytochrome *c* firmly established porphyrin *c*, the porphyrinic substance obtained from acid hydrolysis of cytochrome *c*, as 2,4-di(α-*S*-cysteinylethyl)deuteroporphyrin IX.[2] This also established the α-thioether linkages in cytochrome *c*. Consistent with the generality of two thioether linkages to the heme group has been the finding that in the more than 60 cytochromes *c* for which amino acid sequences are known, all contained the partial sequence ···Cys·X·Y·Cys·His···.[3,4]

Thus the two cysteine residues were considered invariant and hence essential until the recent demonstration that in *Crithidia oncopelti* cytochrome c_{557}[5] and *Euglena gracilis* mitochondrial cytochrome c_{558}[6] the heme group is covalently bound through only one thioether linkage; yet the cytochromes are functional. Therefore, it is of considerable interest to determine the exact nature of this monothiother bond since it must now be considered the minimum linkage necessary for the function of at least some of the *c*-type cytochromes. De-