A STUDY OF THE
MECHANISM OF RIBONUCLEASE-S
USING CNDO/2 ELECTRONIC STRUCTURE CALCULATIONS

by

Michael Arato

****

Submitted in partial fulfillment
of the requirements for
Honors in the Department of Chemistry

UNION COLLEGE

May, 1980

ABSTRACT

ARATO, MICHAEL A Study of The Mechanism of Ribonuclease-S Using CNDO/2 Electronic Structure Calculations. Department of Chemistry, May 1980.

Bovine Pancreatic Ribonuclease-S hydrolyzes RNA molecules in a two step mechanism: 1. transphosphorylation, and 2. hydrolysis.

The mechanism of the enzyme is studied using a model substrate for RNA, 2-hydroxyethyl methyl phosphate anion. Electronic structure calculations are done at the CNDO/2 level, using the CINMIN optimization routine to minimize geometries. Two possible reaction paths are plotted for each step of the reaction, with the lowest activation energy path assumed to be the one taken by RNA. CNDO/2 results, however, yeild negative activation energies, a physically impossible result.

Improved energy calculations for the minimized geometries are then done using GAUSS/70. These results, unfortunately, also yield negative activation energies.

It is concluded that CNDO/2 is unsuitable for work with cyclic phosphate esters.

ACKNOWLEDGEMENTS

I would like to thank the staff of the Union College Computer Center for the numerous hours of computer time and special attention that were needed to complete this study. I would also like to thank Prof. J. Anderson for her help and, most of all, I would like to thank Prof. David Hayes for his guidance, support and patience throughout the development of this paper.

TABLE OF CONTENTS

CHAPTER I

THE ACTION OF RIBONUCLEASE-A

Bovine pancreatic Ribonuclease-A, like all enzymes, is a protein catalyst which serves to lower the activation energy of a biochemical reaction. The reaction, in this case, is the hydrolytic cleavage of an RNA molecule at a non-endpoint position of the polynucleotide. Ribonuclease-A is an extracellular, monomeric enzyme of 124 amino acid residues and molecular weight 13,700. It is believed to function in the digestion of RNA. The enzyme has received a lot of attention and has been studied in both its native form and as Ribonuclease-S, a more easily crystallized variant in which the polypeptide chain is cleaved into two inactive subunits. The cleavage is effected by subtilisn and the enzyme is reactivated in equimolar ratio.

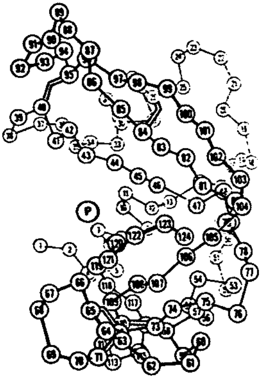


FIG. I. Ribonuclease-A. The phosphate group in crevice opening (P) marks the active site. Disulfide bridges are represented by a bent double line.

The ribonuclease reaction occurs in two steps:

1. TRANSPHOSPHORYLATION: addition of the 2'OH group (on the 2' ribose) to the phosphate group cleaving the ribonucleic acid chain at the 5' end, yielding a 2'-3' cyclic phosphate and a free 5' OH group.

2. HYDROLYSIS: addition of water to the cyclic intermediate, yielding a terminal 3' phosphate monoester.

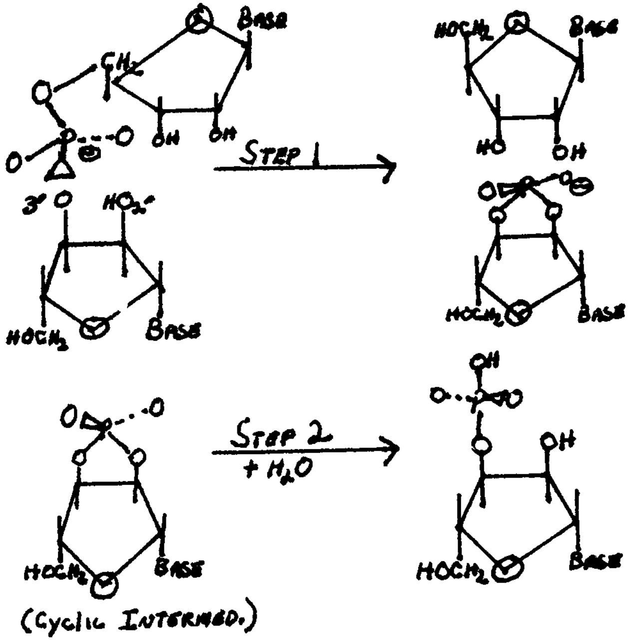


FIG. II. The reaction catalyzed by Ribonuclease-A.

In the early 1960's, His 12 and His 119 were unambiguously implicated in catalysis. Furthermore, from the fact that, at pH = 5.5 the His 119 is approximately eight times more reactive with iodoacetate than His 12, one can conclude that His 119 is active in theacidic form ($pK_a = 5.5$), while residue 12 is active in the basic form ($pK_a = 6.8$).

In 1971, Glick and Barnard showed Lys 41 to be involved in the active site of Ribonuclease-A. The enzyme was inactivated by reaction with dinitrofluorobenzene. Lys 41 was then shown to be the residue responsible for inactivation by selective addition of guanine. These and other experimental results, along with X-ray crystallography of the structure of ribonuclease and theoretical calculations have allowed researchers to propose some possible mechanisms for the action of ribonuclease.

Mathias and Rabin, et. al. have proposed that, for step 1, removal of the proton on the 2' OH by an imidazole residue in the basic form (His 12) and protonation of the 5' O of the leaving nucleoside by the acidic imidazole (His 119) is the basis for the mechanism. This is followed by an attack by the 2' alkoxide on the phosphorous atom to yield the cyclic phosphate. Step 2 follows the reverse of these steps with His 12 acting as the acidic group and His 119 as the basic group. Lysine 41 was also implicated in stabilization of the negative charge on the phosphate oxygens and water was included in hydrogen bonding the acidic His and the carbonyl oxygen of the pyrimidine base.

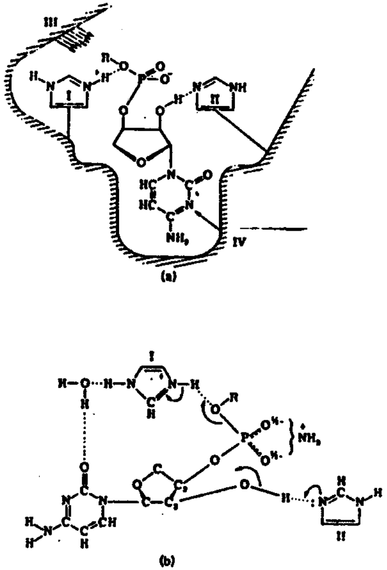


FIG. III. Mechanism of action of Ribonuclease according to Mathias and Rabin, *et. al.* a. role of Histidines. b. with Lysine and H_2O

Another possible mechanism for the action of ribonuclease is the Witzel mechanism. He assumed that the 2' carbonyl oxygen of the pyrimidine base plays a major role and that the enzyme acts only to alter the electrophillic character of the phosphorous by single or double protonation of the transition state. His 119 and Lys 41 are assumed to be the positive sites on the enzyme and His 12 is assumed to interact with His 119 by sharing a proton to maintain proper enzyme geometry.

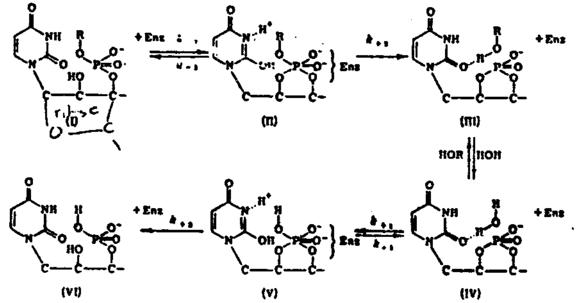


FIG. IV. Mechanism of action of Ribonuclease according to Witzel.

Wang used the concept of facilitated proton transfer to explain Ribonuclease-A action. A basic imidazole group plucks the hydrogen from the 2' OH group of the ribose to form the transition state. The same proton is then donated to the 5' oxygen of the leaving nucleotide at the end of the cyclization of Step 1. For Step 2, the path is reversed. This process would require a lack of movement by any of the atoms involved if it is to occur efficiently. The 2' oxygen atom must, however, be close to the imidazole to donate the proton and the 5' oxygen atom must also be close to the imidazole group to accept it. Pseudorotation is, therefore, a necessary part of the mechanism. Lys 41, or the second His, is believed to be involved in the protonation of a free equatorial oxygen to facilitate attack and deprotonation.

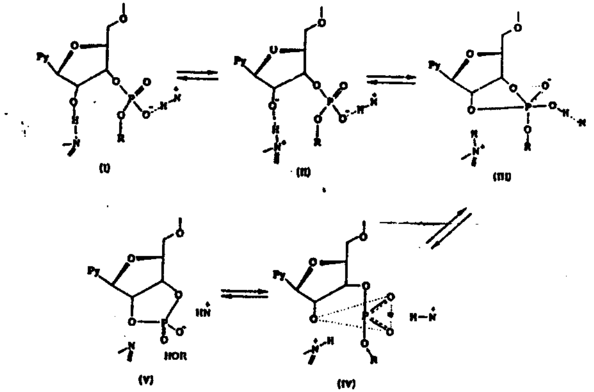


FIG. V. Mechanism of action of Ribonuclease according to Wang.

Intrinsically related to the mechanism of Ribonuclease action is the stereochemistry of the transesterification and hydrolysis of the substrate. For each step, there is the possibility of an "in-line" attack or an "adjacent" attack with respect to the leaving group. In the case of adjacent attack, pseudorotation is necessary. For in-line mechanisms it is possible but not obligatory.

Roberts, in 1961, showed by using a dinucleotide enzyme complex that a linear mechanism is probably correct. Similarly, in 1972, Eckstein, et. al. reacted uridine 2', 3'-O, O-cyclo-thiophosphate with pancreatic Ribonuclease-A in aqueous methanol. The ensuing X-ray analysis of the absolute configuration of the resulting uridine 3'-o-thiophosphate methyl ester showed that Step 2 is also an in-line mechanism. Both of these results have helped support the mechanism proposed by Mathias and Rabin, as the mechanisms of Witzel and Wang both imply adjacent attack. Evidence is, however, still inconclusive.

X-ray studies have also shown the aromatic rings of PHE 120 to be important in the binding of the pyrimidine bases to the enzyme. In 1975, however, Sigman and Mooser tested Ribonuclease-A substituted at position 120. While activity was greatly reduced for substituted amino acids (other than PHE), the loss was attributed to reduced re-binding of the two enzyme fragments, since K_m values for the reaction were unaffected. They therefore concluded that PHE 120 is not essential in binding the substrate.

Hodges and Merrifield, in 1975, used base analogs and Ribonuclease-S to show that the hydroxyl group of SER 123 and the C^4 carbonyl oxygen of uridine contributed significantly to the substrate binding, along with THR 45. Cytidine and formycin did not follow a similar pattern. This evidence shows that pyrimidine bases may bind to the enzyme differently than the purines.

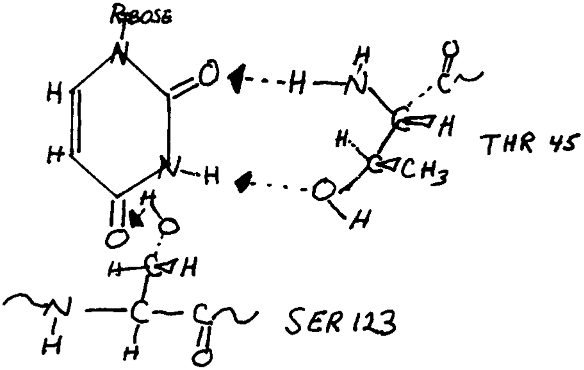


FIG. VI. Binding of Uridine to SER 123 and THR 45.

The specific role that lysine plays in catalysis has been clarified somewhat by Walter and Wold in 1976 and by Gutte in 1977. By studying a 63 amino acid analog of Ribonuclease-A, the researchers showed that LYS 41 plays no role in the binding of the substrate (in the formation of the Michalis-Menton complex) but is essential for catalysis. It is still not clear just what this role is, however.

Deakyne and Allen, in 1978, combined some of the previously accumulated experimental data with CNDO/2 electronic structure theory calculations in examining the mechanism of Step 1. The results favor a modified version of the mechanism proposed by Mathias and Rabin. They concluded that, for the attachment and partial protonation segment of transphosphorylation:

1. HIS 119 partially protonates the the leaving group, weakening the $P-O_5$ bond and facilitating in-line addition of 2' oxygen.

2. The 119-120 backbone N-H aids 2' oxygen attack by increasing the electrophilicity of Phosphorous. GLN 11 ... H_2O serves the same purpose.

3. The substrate must be hydrogen bonded to HIS 119 and the N-H backbone to activate both Phosphorous and the leaving group.

4. The probable function of ASP 121 is to position the adenine ring of the substrate by means of a water mediated hydrogen bond.

For the deprotonation segment of transphorylation:

1. HIS 12 is more likely to deprotonate 2' oxygen, but in its absence ASN 44 may fulfill this role.

For the nucleophilic addition segment of transphosphorylation:

1. HIS 119 rotates with the phosphate group and protonates the leaving group, facilitating breakdown of the trigonal bipyramid.

2. LYS 41 increases the electrophilicity of Phosphorous and stabilizes the TBP during the addition of 2' oxygen.

3. ASP 121 is not involved in this segment of the reaction.

One can, therefore, see that the Mathias and Rabin mechanism has received the most support, although there are many intricacies still being debated. It is, for example, often assumed that Step 1 and Step 2 proceed by reverse mechanisms. This is not necessarily true. More work is needed before a definitive answer to the question of Ribonuclease-A activity can be obtained.

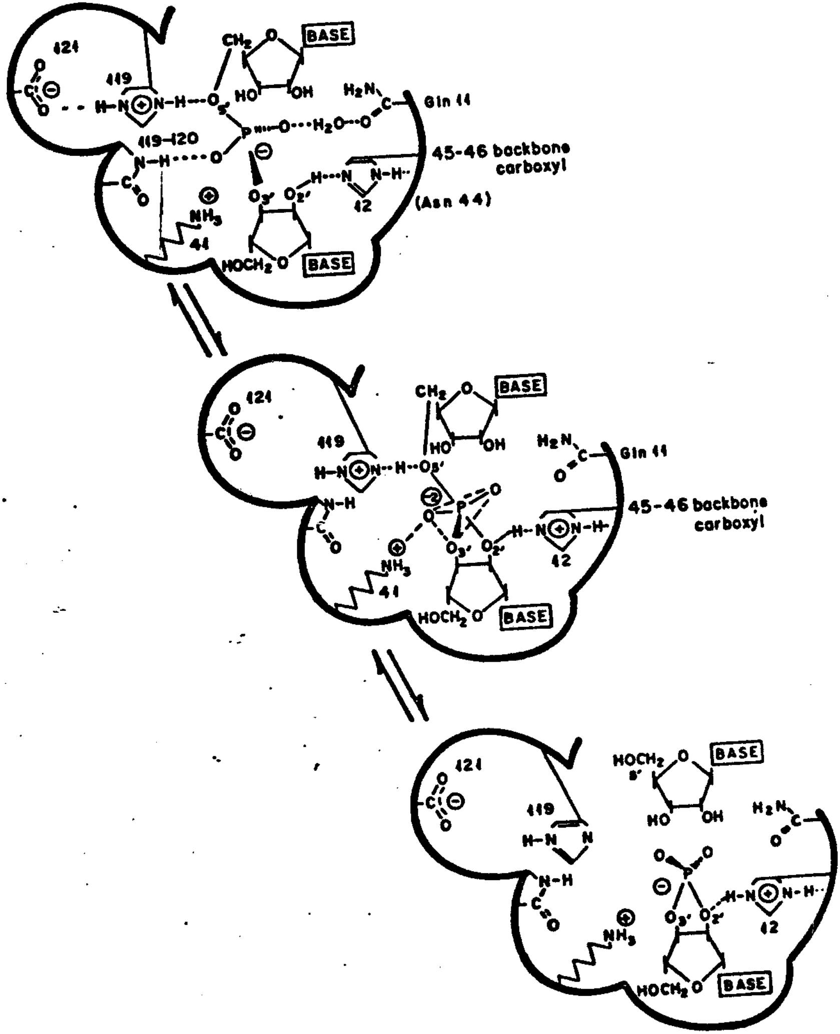


FIG. VII. The mechanism of Ribonuclease action proposed by Deakyne and Allen.

BIBLIOGRAPHY

Barnard. "Ribonucleases". ANNUAL REVIEW OF BIOCHEMISTRY. Vol. 38, 1969.

Bender. MECHANISMS OF HOMOGENEOUS CATALYSIS FROM PROTONS TO PROTEINS. New York: John Wiley and Sons. 1971.

Deakyne and Allen. "The role of active site residues in the Catalytic Mechanism of Ribonuclease-A". unpublished paper received by JACS, Aug. 17, 1978, No. 82047.

Dickerson and Geiss. THE STRUCTURE AND ACTION OF PROTEINS. Mass: W.A. Benjamin, Inc. 1969.

Eckstein, et. al. "Stereochemistry of the Transesterfication step of Pancreatic Ribonuclease." BIOCHM BIOPHYS RES COMMUN, Vol 46, pg. 964-71.

Glick, David, et. al. "Some Relations of Lysine Residues to Reactivities in Pancreatic Ribonuclease." BIOCHM BIOPHYS ACTA, Vol. 214, pg. 320-342. Aug. 21, 1970

Gutfreund. AN INTRODUCTION TO THE STUDY OF ENZYMES. London: Blackwell Science Publications. 1965.

Gutte, B. "Study of RNase A Mechanism and Folding by Means of Synthetic 63-residue Analogs." J. BIOL CHEM, Vol. 252, pg. 663-70. Jan. 25, 1977.

Hodges and Merrifield. "The Role of Serine-123 in the Activity and Specificity of Ribonuclease." BIOCHM BIOPHYS RES COMMUN, Vol. 46, pg. 964-71, Jan. 31, 1972.

Richards and Wyckoff. "Bovine Pancreatic Nuclease" in THE ENZYMES, edited by Paul D. Boyer. (Vol IV). New York: Academic Press. 1971.

Sigman and Mooser. "Chemical Studies of Enzyme Active Sites." ANNUAL REVIEW OF BIOCHEMISTRY. Vol. 38, 1969.

Walter, B., et. al. "The Role of Lysine in the Action of Bovine Pancreatic Ribonuclease". BIOCHEMISTRY, Vol. 15, pg. 304-10, Jan. 27, 1976.

CHAPTER II

HYDROLYSIS OF CYCLIC PHOSPHATE ESTERS

Ribonuclease-A, as mentioned in the previous chapter, works to cleave RNA in a two stage process: 1. transphosphorylation to form a five membered cyclic phosphate ester intermediate and 2. hydrolysis of the intermediate. Using the model substrate, 2-hydroxyethyl methyl phosphate anion, these stages are illustrated as follows.

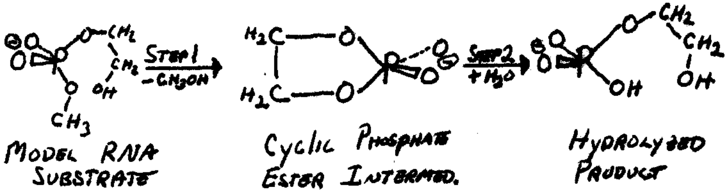


FIG. I. The action of Ribonuclease on a model RNA substrate molecule.

This chapter will outline some of the recent research that has been conducted in the area of phosphate ester hydrolysis. Emphasis will be placed on kinetic acceleration and the mechanism of hydrolysis.

One of the most unique characteristics of cyclic phosphate ester hydrolysis is the tremendous kinetic acceleration of the five membered ring in comparison with corresponding open chain analogs. For example, ethylene phosphate hydrolyzes in both alkali and acid approximately 10^7 times faster than dimethyl phosphate. The six membered cyclic phosphate, trimethyl phosphate, however, and the seven membered cyclic phosphate ester is only slightly more stable. This acceleration of the five membered ring has been attributed to ring strain, as the heat of hydrolysis of methyl ethyl phosphate exceeds that for dimethyl hydroxyethyl phosphate by about 7-9 kcal/mole.

In 1960, Haake ane Westheimer used ^{18}O tracers to investigate the position of cleavage of ethylene and dimethyl phosphate in both acid and base. The hydrolysis result and the accompanying exchange reaction for ethylene phosphate are shown below.

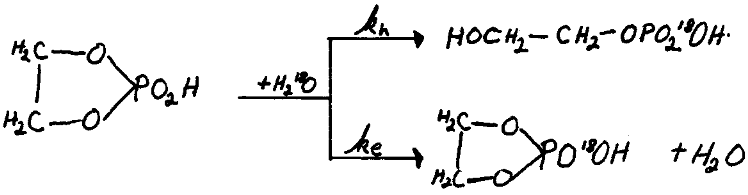


FIG. II. Hydrolysis and Exchange for Ethylene Phosphate.

This exclusive P-O bond cleavage contrasted with the result for dimethyl phosphate, which showed a great deal of C-O bond cleavage in addition to the P-O cleavage.

The large percentage of molecules (approx. 20%) that undergo exchange implies that this rate, like the hydrolysis rate, is enhanced. Since the strain that presumably causes the rapid hydrolysis and exchange is also present in the exchange product, one must assume that the strain is relieved in the transition states for both reactions. Three possible pairs of transition states for the hydrolysis and exchange reaction are shown below.

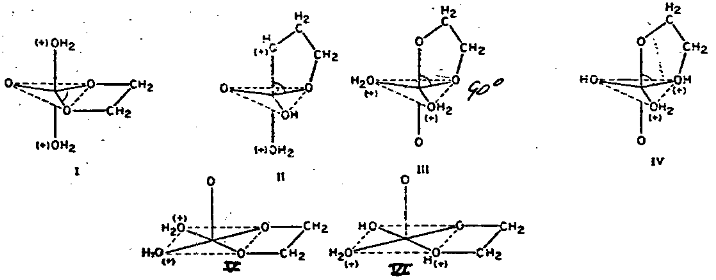


FIG. III. Possible Transition State Geometries for the acid catalyzed Hydrolysis and Exchange of Ethylene Hydrogen Phosphate.

One would expect that, for each pair, strain is relieved in the transition state to explain the fact that both hydrolysis and exchange occur, with the strain still present in the exchange product.

Structures I and II are transition states for an SN2 mechanism. They do not, however, meet the equal strain relief requirement, as geometry I has the O-P-O angle expanded to 120° and geometry II has the O-P-O angle contracted to 90°. The normal angle is approximately 110°. It is unlikely that both expansion and contraction would relieve ring strain and yield both the hydrolysis and exchange product; this set is, therefore, disqualified.

Formulas III and IV are similar to I and II in that they arise from an SN2 reaction but, in this set, both the entering and leaving groups are at the basal positions. Both III and IV have contracted O-P-O angles of 90° and would, therefore, be expected to have comparable energies relative to the reactants, allowing both hydrolysis and exchange to occur. This set, therefore, remains a likely candidate.

Formulas V and VI have the geometries of square pyramids and, similar to III and IV, have equivalent O-P-O angles of 90°. They too remain likely candidates. One must keep in mind, though, that the preceeding analysis assumes that ring strain is caused by angle strain and that the ring strain causes the accelerated kinetics. This was shown to be an erroneous assumption by later workers.

In connection with the phenomenon of kinetic acceleration found in cyclic phosphates, it is of interest to consider studies done on cyclic sulfates and sulfites. In 1961, Davis did studies similar to Westheimer's, using sulfites in lieu of phosphates. His results showed that, while there was a kinetic acceleration of 360 for the cyclic dimethyl ester relative to the open chain analog, the heats of hydrolysis were, however, approximately the same for for the ring and open chain structures. This does not support Westheimer's earlier conclusion of ring strain being the ultimate cause of the acceleration. Similar results were later obtained with phosphite esters.

In 1962, Kaiser, Panar and Westheimer did similar studies with cyclic sulfate esters.

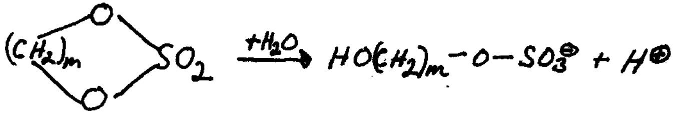


FIG. IV. Sulfate Ester Hydrolysis Reaction.

Kinetic acceleration for cyclic esters with respect to acyclic analogs was discovered and, like the cyclic phosphate esters, a thermodynamic difference of 5-6 kcal was noted. Westheimer, <u>et. al.</u> used these results along with the phosphite and sulfite data to conclude that partial p-d double bonding between the ester oxygen and the heteroatom (respectively) is essential to the kinetic acceleration. This double bond would be suppressed in the sulfites and phosphites by the presence of an unshaired pair of electrons in the heteroatom. X-ray and N.M.R. data have helped to support this theory.

In 1976, Gorenstein, <u>et. al.</u> , used CNDO/2 LCAO-MO calculations to show that some of the extra heat of hydrolysis may be attributed to torsional strains in the cyclic esters. The results of this thesis, however, question the validity of results obtained using CNDO/2 for cyclic phosphate ester calculations. Gorenstein's results will, therefore, be disregarded.

More light was shed on the subject of mechanism when experiments revealed that enhanced rates (10^5–10^8 times) of hydrolysis for phosphates are also manifest in reactions that occur external to the ring. In 1966, Dennis and Westheimer found that, while an enhanced rate external to the ring was evident for the cyclic phosphate esters, molecules such as the methyl ester of propyl-phostonic acid, (see Fig. V.), undergo hydrolysis at an enhanced rate, but with ring cleavage only. This result was explained by the following mechanism.

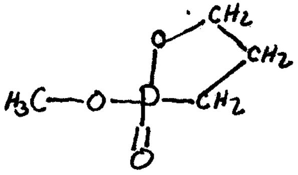


Fig. V. Methyl ester of Propyl-phostonic acid.

Water adds to occupy an apical position in a triginal bipyramid; the shift of the proton produces an equitorial water molecule that can serve as a point for pseudorotation. The pseudorotation process places a methoxy group in the apical position of a new triginal bipyramid; after a proton shift, a methanol molecule can leave from the apical position. In both triginal bipyramids, the five membered ring spans one apical and one equitorial position, with reduced ring strain. This explains why hydrolysis external to the ring is also accelerated.

The corresponding hydrolysis with ring opening can occur from the original triginal bipyramid with a proton shift, with or without pseudorotation. This process would, however, be energetically unfavorable for the molecule in Fig. V., as it would demand that an alkyl group be placed in the apical position. This is undesirable because orbital overlaps favor the more electronegative species in the apical positions. In the apical positions, p_z and d orbitals are directed in space and so can provide good overlap with polar atoms that pull electrons away from the phosphorous. The s character in the equitorial sp^2 orbitals, however, will bond best with electron donating groups. Earlier NMR studies on TBP's of fluorine compounds had shown this to be true.

The intermediates leading to the external hydrolysis for the molecule in Fig. V., while unfavorable, are not forbidden. One may then predict that such reactions will occur when an especially large reduction in ring strain accompanies the formation of the intermediate. Such strained compounds have been tested by Kluger, et. al. in 1967 and the results support this theory.

In 1968, Donald Boyd used the LCAO method of computing molecular orbitals on cyclic phosphate esters. His results verify Westheimer's mechanism and allow us to make the following conclusions regarding cyclic phosphate ester hydrolysis:

1. the hydrolyses proceed via a nucleophilic attack on phosphorous leading to a trigonal bipyramid intermediate,

2. the five membered ring spans one apical and one basal position,

3. groups enter and leave from apical positions,

4. pseudorotation of the trigonal bipyramid can take place leading to exchange of the apical and basal positions,

5. the pseudorotation occurs subject to the preference of less electronegative substituents to occupy basal positions and of the more electronegative substituents to occupy apical positions,

6. pseudorotation is the rate limiting step in hydrolysis external to the ring. (This was verified by Kluger, *et. al.* in 1969.)

BIBLIOGRAPHY

Boyd, Donald. "Mechanisms of Cyclic Phosphate Esters." JACS, Vol. 91(5), pp. 1200-1205. 1969

Bunton, C.A., et. al. "Nucleophilic Displacement in Organic Sulphites. Part III. Comparison of Alkaline and Acidic Hydrolyses of Trimethylene Sulfite and Related Compounds." 1959.

Davis, Robert Earl. "Hydrolysis of Ethylene and Dimethyl Sulfite and the Origin of Ring Strain in Cyclic Esters." JACS, Vol. 84. 1962.

Dennis, E.A. and F.H. Westheimer. "The Geometry of the Transition State in the Hydrolysis of Phosphate Esters." JACS, Vol. 88, pp. 3432. 1966.

Eberhard, A., and F.H. Westheimer. "Hydrolysisis of Phostonates. JACS, Vol. 87(2), p. 253-260. 1965.

Gerlt, J. and F.H. Westheimer and J. Sturtevant. "The Enthalpies of Hydrolysis of Acyclic, Monocyclic and Glycoside Cyclic Phosphate Diesters." J. Biol. Chm., Vol. 250(3), pp.5059-5067. 1975.

Gorenstein, David. "The Barrier to Pseudorotation in Cyclic Alkoxyphosphoranes." JACS, Vol. 92(3), pp. 644-50. 1970.

Gorenstein, David. "pH-rate profiles for the Oxygen-18 Exchange and Epimerization of a Phosphatene Oxide." JACS, Vol. 94(8), pp. 2808-14. 1976.

Gorenstein, David and F.H. Westheimer. "NMR evidence for the Pathways of Pseudorotation in Alkyloxyphosphoranes." JACS, Vol. 92(3), pp. 634-44. 1970.

Gorenstein, David, et. al. "Conformational Study of Cyclic and Acyclic Phosphate Esters. CNDO/2 Calculation of Angle Strain and Torsional Strain. JACS, Vol. 98(7), pp. 1668-73. 1976.

Haake, P. and F.H. Westheimer. "Hydrolysis and Exchange in Esters of Phosphoric Acid." JACS, Vol. 83. 1961

Kaiser, E.T. and M. Panar and F.H. Westheimer. "The Hydrolysis of Some Cyclic Esters of Sulfuric Acid." JACS, Vol. 85. 1962.

Kluger, R. and F.H. Westheimer. "The pH rate profile for the Hydrolysis of some Esters of a Bicyclic Phosphinic Acid. Evidence for Rate Limiting Pseudorotation." JACS, Vol. 91(15), pp. 4143. 1969.

Kluger, R. et. al. "Hydrolysis of Esters of Bicycloheptyl-and-heptenyl Phosphonic Acids." JACS, Vol. 89, pp. 3918-19. 1967.

Kluger, R., *et. al.* "pH Product and and pH Rate Limiting Profiles for the Hydrolysis of Methyl Ethylene Phosphate. Rate Limiting Pseudorotation." JACS, 91(22), pp. 6066-72. 1969.

Szele, I., S. Kubisen and F.H. Westheimer. "Pseudorotation in a Sterically Hindered Phosphorene."" JACS, Vol. 98(12), pg. 3533-6.

Westheimer, F.H. "Pseudorotation in the Hydrolysis of Phosphate Esters." ACCTS CHM RSCH., Vol 1, pp. 70-8. 1968.

Westheimer, F.H. "The Rate of Hydrolysis of Esters of Cyclic Phosphinic Acids." JACS, Vol. 88(14), pp. 3431-2. 1966.

CHAPTER III

THE CNDO/2 METHOD

This section explains the basic workings of the method employed in calculating molecular geometries and energies. First, an outline of quantum mechanical theory is presented and then the Schroedinger equation is used as the basis for the SCF-LCAO method.

QUANTUM MECHANICS

Experiments done in the late 1800's and early 1900's led many scientists to believe that Newtonian (classical) mechanics could not explain observations made at the atomic level. A new model had to be constructed and, in 1926, quantum mechanics was born. Before the resulting quantum mechanical theory is discussed, however, a review of classical physics is in order.

Classical mechanics describes the behavior of particles with two equations. The first is

$$E = \tfrac{1}{2}mv^2 + V(x),$$

where x and v are functions of t. This may also be written in terms of linear momentum

$$E = p^2/2m + V(x).$$

This equation implies that the whole future behavior of a particle can be predicted if its present position and velocity are known. There appeared to be no limit to the accuracy with which these values could be known.

The second equation is a statement of Newton's second law

$$\dot{p} = F(x),$$

which relates the acceleration of a particle and the force it experiences.

This implies that, if the force acting on a particle in every region of space is known, we can find $\not{p}(t)$ and from this, find position.

When the two equations are considered together, one may conclude that the force and time may be varied arbitrarily and the energy of a particle is, therefore, random, being controlled by the impulse of forces and torques acting on it. While these conclusions are in accord with our everyday experiences, experiments have shown that they fail for small masses and small transfers of energy.

For example, in the late 1800's, physicists found that Maxwell's equations fail to predict the observed behavior of thermal radiation for short wavelengths. In 1897, however, Max Planck accidentally found that an accurate equation could be derived if one assumed that the thermal energy is released in discreet packets. This amazing result implies that Maxwell's equations do not fully describe energy behavior and that radiation has particular properties.

There was also evidence that matter, previously assumed to be particular in nature, has wave properties. It wasn't until 1925, though, that Davison and Germer's experiment with Nickel crystals silenced all opposition. They found that the behavior of electron scattering in their sample could only be explained if one assumes that matter has wave properties. This had been suggested by deBroglie in 1924. These and other experiments led scientists to conclude that matter has wave characteristics and that energy has particular properties.

Thus, it seemed that the basic tenents of classical mechanics needed to be modified. The old laws of Newton and Maxwell had to be replaced by quantum mechanical theory. Some of the interesting implications of this theory are that it is impossible to know position and momentum to arbitrary precision, that it is not possible to think in terms of a definite trajectory and that it is not possible to inject arbitrary amounts of energy into a system. Quantum mechanical theory also has philosophical, as well as chemical implications. These will now be discussed.

One of the basic tenents of the doctrine of fatalism, the law of the excluded middle, is voided by quantum mechanical theory and by experimental verifications of the theory. Fatalism is based on the idea that everything that happens in the universe is caused by prior actions and forces. Whether we have a "soul" or are composed merely of atoms and molecules does not concern us, as our souls would still be controlled by the same cause and effect relationship. While a rigorous treatment of fatalist doctrine will not be presented here, we need to know that proponents of this metaphysical perspective argue that any statement may be judged true or false, only. This is the law of the excluded middle.

One could say, for instance, that President John F. Kennedy was assassinated in Texas. This statement is true, regardless of whether or not a person has knowledge of the incident. One could also make the statement that Kennedy was thinking about his mother when the incident occurred. This statement is also true or false, although no person on earth could ever say, with certainty, what Kennedy was thinking of. It follows then, that all statements, actions or thoughts are either true or false. Those events that have already passed differ from those that will happen in the future only in that we have some limited knowledge of the past but know virtually nothing about the future. I could say that

I will complete this thesis on time. This statement is either true or false; we must simply wait to see the result because our limited knowledge prevents us from making a certain prediction.

If the doctrine of the excluded middle is true, as our everyday experiences lead us to believe, then one may conclude that, since all actions are true or false and since all past events have a causal effect on future events (which are also either true or false) then the outcome of future events is inevitable. That is, our fate is already predetermined.

It is fortunate for those of us who feel uncomfortable with this conclusion, however, that the law of the excluded middle is disproved by events that have been observed at the molecular level. To cite the simplest case, the lone electron that orbits a hydrogen atom does not have an exact position. If one were to examine an infinitesimal unit of volume, dv, and tried to answer the question, "Is the electron there?", the reply would be neither yes nor no; the question would have to be answered with an intermediate response. This follows from the Born interpretation of ψ, namely, that $\psi^*\psi$ is a density function for the electron.

Similarly, the Heisenberg uncertainty principle attests to the inexact nature of our universe. While our daily experiences imply that the universe has exact properties, regardless of whether they are known or not, Heisenberg showed this to be untrue at the molecular level. Momentum and position may not be assigned values of arbitrary precision. This is expressed mathematically by

$$\delta p \, \delta q \geq \tfrac{1}{2}\hbar,$$

where p is momentum and q is position.

While other examples could be listed, the point has already been made; quantum mechanical theory disproves the law of the excluded middle and shatters the fatalist arguement.

THE SCHROEDINGER EQUATION

The central feature of quantum mechanics is the association of wave character with particles. We abandon the classical concept of localized particles and replace it with the idea that the position of a particle is distributed like the amplitude of a wave.

In 1926, Erwin Schroedinger modified wave equations that had been derived in the early 1800's by William Hamilton to describe the standing waves that might occur on a planet flooded with water. For example, for a single particle free to move in one direction, the equation reads

$$(-\hbar^2/2m)\,(d^2/dx^2)\,\psi(x) + V(x)\,\psi(x) = E\psi(x)$$

where ψ is a wavefunction, $V(x)$ is potential energy and E is total energy.

The Born interpretation of the wavefunction states that the square of ψ, $(\psi^*\psi)$ is proportional to the probability of finding the electron in an infinitesimal region. This destroys the classical idea of a precise trajectory, as we cannot even define an exact position at a time, t. Another way of describing this anomoly is the Heisenberg uncertainty principle, which reads

$$\delta p\,\delta q \geq \tfrac{1}{2}\hbar,$$

where p is momentum and q is position. Thus, one cannot know the position and momentum of a particle to arbitrary precision.

The imposition of boundary conditions on a particle (such as a particle in a box or an electron in an atom) restricts the energies which the particle may assume. Some energies are possible, while others are not. For large masses and large transfers of energy, though, the quantization becomes so slight that the energy appears to be a continuous function. It is only at the atomic level that the effect is significant.

In work with atomic and molecular structure, one is concerned with the potential of an electron moving in the field of a nucleus. The Schroedinger equation for one particle in three dimensions describes such a system and is written

$$(-\hbar^2/2m)\nabla^2\psi + V(x,y,z)\psi = E\psi,$$

where $\psi = \psi(x,y,z)$ and $\nabla^2 = (\partial^2/\partial x^2) + (\partial^2/\partial y^2) + (\partial^2/\partial z^2)$. V may be replaced by the classical expression for the potential of an electron moving about a nucleus of charge Ze,

$$v = Ze^2/r.$$

The Schroedinger equation then becomes

$$-\hbar^2/2m\,\nabla^2\psi + Ze^2/r\,\psi = E\psi.$$

The sum of the potential and kinetic energy expressions in the above equations is called the Hamiltonian Operator, and is denoted by H. Thus, for a one electron system (eg. the Hydrogen atom) the Schroedinger eqation is often written

$$H(1)\psi(1) = \mathcal{E}\psi(1).$$

(One electron energies are usually written as $\mathcal{E}$ rather than E. The (1) labelling reflects the functional dependence of H, ψ and $\mathcal{E}$ on one electron.

For many particle systems one must consider repulsive and attractive interactions in kinetic and potential energies when constructing the Hamiltonian operator. For N nuclei and n electrons, the Hamiltonian is

$$H(1,2,\ldots N;\ 1,2,\ldots n) = -\hbar/2\pi\sum_{A}^{N} M_A^{-1}\nabla_A^2 + \sum_{A<B} e^2 Z_A Z_B r_{AB}^{-1}\,(-\hbar/2\pi m) \times \sum_{p}^{n}\nabla_p^2 - \sum_{A}\sum_{p} e^2 Z_A r_{Ap}^{-1} + \sum_{p<q} e^2 r_{pq}^{-1}.$$

where M_A is the mass of nucleus A, m is the mass of an electron, Z_Ae is the charge of nucleus A, r_{ij} is the distance between particles i and j, A and B are variables that represent nuclei, and p and q are variables that represent electrons.

The abbreviated form of the Schroedinger equation for a many particle system is thus written

$$H^{total}(1,2,\ldots N;\ 1,2,\ldots n)\Psi(1,2,\ldots N;\ 1,2,\ldots n) = E\Psi(1,2,\ldots N;\ 1,2,\ldots n).$$

This equation will yield an infinite number of solutions, but only some of the solutions will meet the restrictions imposed by boundary conditions discussed earlier.

The acceptable energies and corresponding wavefunctions will each be assigned a numerical subscript and, in general, we may write

$$H^{total}\Psi_i = E_i\Psi_i.$$

Using differential equation nomenclature, the wavefunctions are called eigenfunctions, the energies are called eigenvalues and two or more wavefunctions which result in equal energies are termed degenerate.

A very common approximation for many particle systems is known as the Born-Oppenheimer approximation. It assumes that the nuclei are fixed. This approximation is reasonable because the nuclei are so much heavier than the electrons. The result is that the part of the Hamiltonian that considers nuclear momenta can be factored out and considered as a separate problem. We are then left with the electronic Hamiltonian, H^{el}

$$H^{el} = -\hbar^2/2m^2\sum_p \nabla_p^2 - \sum_A\sum_p e^2 Z_A r_{Ap}^{-1} + \sum_{p<q} e^2 r_{pq}^{-1},$$

and

$$H^{el}(1,2,\ldots n)\Psi_{el}(1,2,\ldots n) = \mathcal{E}\Psi_{el}(1,2,\ldots n).$$

(Since M.O. theory is concerned with electronic energy only, the superscript will be dropped.)

The total energy of this purely electronic system must include the electronic energy, $\mathcal{E}$, calculated from the Schroedinger equation above and the fixed nuclei repulsion terms. The result is

$$E = \mathcal{E} + \sum_{A<B} e^2 Z_A Z_B r_{AB}^{-1}.$$

THE VARIATION METHOD

The obvious disadvantage of the Schroedinger equation is that solutions for many electron systems are quite time consuming. What is used instead is the variational method, which, in its full form, is equivalent to the full differential equation method but is adaptable to approximate wavefunctions.

The variation method approximation involves attempting to find the wavefunction which minimizes the total electronic energy. The initial guess, or trial eigenfunction, is made and the theorem is employed. The theorem states that "the expectation value of H for any arbitrary well-behaved function, Ψ, is not less than the lowest eigenvalue, E_0 of H." Thus, the lowest value of E may be considered the closest to the true value and

$$\partial E/\partial c_1 = \partial E/\partial c_2 = \ldots = \partial E/\partial c_n = 0.$$

A curve fit of these simultaneous equations will afford one the c values which minimize the energy. (A more specific interpretation of c will be presented later.)

If one denotes the resulting function as ϕ_i, called basis functions, then the variation method results in

$$\Psi(c_1, c_2, \ldots c_n) = c_1\phi_1 + c_2\phi_2 + \ldots + c_n\phi_n.$$

These wavefunctions will be used to define orbitals, as will be described in the next section.

THE ORBITAL APPROXIMATION

In constructing an approximate many-electron wavefunction, a combination of one electron orbitals is used, with electron correlational effects ignored. Mathematically stated

$$\Psi(1,2,\ldots n) = \psi_1(1)\psi_2(2)\ldots\psi_n(n).$$

The resulting expression is called a Hartree product and is comprised of a product of one-electron orbitals, ψ_i.

The Hamiltonian is also modified to neglect the electron repulsion term, r_{ij}^{-1}. This is done by writing the Hamiltonian as a sum of one-electron operators, F(i), where

$$F(1,2,\ldots n) = \sum_p F(p) = \sum_p(-\tfrac{1}{2}\nabla_p^2 + V(p)).$$

The one-electron potential, V(p), is calculated by considering the potential of the nuclei and the instantaneous position of all of the electrons. Thus, one can compute the energies for each of the electrons seperately using the one-electron wavefunction and the one-electron Hamiltonian as

$$F(1)\psi_i(1) = \varepsilon_i\psi_i(1).$$

In the formulation of F(p), it is necessary to know the molecular orbitals, ψ_i, which are the basis for the computation of V(p). This is done by the Variational method mentioned above. If the molecular orbitals obtained by solving the one-electron equation above are identical to those used in formulating F(p), then these orbitals are termed self-consistent.

ELECTRON SPIN

Thus far nothing has been said about electron spin states. Electrons behave as if they were spinning about an axis and, since they have a charge, they behave like tiny magnets. The angular momentum of the electrons is quantized into states of $\pm\hbar/2$ and is represented by S. If s_x, s_y and s_z are taken as the three dimensional vector components of S, then one can say that the electron spin states are quantized about an imaginary z axis which represents the alignment of the applied magnetic field. If ξ is the spin coordinate, then these two spin states, $(\pm\hbar/2)$, may be represented by $\alpha(\xi)$ and $\beta(\xi)$, and

$$S_z\alpha(\xi) = \tfrac{1}{2}\hbar\alpha(\xi)$$

$$S_z\beta(\xi) = -\tfrac{1}{2}\hbar\beta(\xi),$$

or

$$S^2\eta(\xi) = s(s+1)\eta(\xi)$$

$$S_z\eta(\xi) = m_s\eta(\xi),$$

where $\eta(\xi)$ may be $\alpha(\xi)$ or $\beta(\xi)$ and m_s may be $\pm\frac{1}{2}$.

Combining the previously considered wavefunction, ψ_i, with the now considered property of electron spin results in the complete wavefunction. This is called a spin orbital

$$\psi_i(r)\eta(\xi).$$

Since there are two spin states, α or β, for a given radial orbital, $\psi_i(r)$, two spin orbitals result,

$$\psi_i(r)\alpha(\xi) \quad \text{and}$$

$$\psi_i(r)\beta(\xi).$$

Our new Hartree product of spin orbitals becomes

$$\psi(1,2,\ldots,n) = \psi_1(1)\alpha(1)\psi_2(2)\beta(2)\ \ldots\psi_n(n)\ (n).$$

ANTISYMMETRY AND DETERMINAL WAVEFUNCTIONS

Since all electrons possess the exact same physical properties, the system will not be affected if they are renumbered. Since the density function is written in terms of $\psi^*\psi$ from Born's interpretation of ψ

$$\rho(1,2,\ldots n) = \psi^2(1,2,3,\ldots n),$$

and since the expression must be unaffected by renumbering, the following result is obtained

$$\psi(1,2,\ldots i,j,\ldots n) = \pm\psi(1,2,\ldots j,i,\ldots n).$$

There are thus two possible results. Either a factor of +1 preceeds ψ or a factor of −1 preceeds ψ. In the former case we designate the wavefunction as being symmetric; in the latter case the wavefunction is designated as antisymmetric.

This "antisymmetry principle" is often stated in terms of a permutation operator, P_{ij}, which interchanges all of the coordinates of the electrons i and j,

$$P_{ij}\,\psi(1,2,\ldots n) = -\psi(1,2,\ldots n)$$

Our original Hartree product equation,

$$\psi(1,2,\ldots n) = \psi_1(1)\alpha(1)\,\psi_2(2)\beta(2)\ldots\psi_n(n)\ (n),$$

must be modified to satisfy the antisymmetry principle, which is the basis for the Pauli exclusion principle. The modification is written, (for a two electron system),

$$\psi(1,2) = \sum_{k}^{2}(-1)^{p}P_k\,\psi_1(1)\alpha(1)\psi_1(2)\beta(2).$$

To illustrate the relationship of the antisymmetry principle to the Pauli exclusion principle, the equations are written in determinant form. These determinants are called Slater determinants. A single Slater determinant is the simplest wavefunction which satisfies the antisymmetry principle.

For our two electron equation, the Slater determinant is

$$\psi(1,2) = 7\begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2) \end{vmatrix} .$$

In the general case of 2n electrons, the Slater determinant is

$$\psi(1,2,\ldots n) = 7\begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) & \psi_2(1)\alpha(1)\ldots\ldots\psi_n(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2)\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ \psi_1(2n)\alpha(2n)\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\psi_n(2n)\beta(2n) \end{vmatrix} .$$

Slater determinants are also written in the followin abbreviated form

$$\psi(1,2,\ldots n) = |\psi_1\bar{\psi}\ldots\ldots\ldots\ldots\psi_n\bar{\psi}| .$$

Using the Slater determinant notation, the antisymmetry principle results from the mathematical result of the sign of the determinant changing if any two rows are exchanged. The Pauli exclusion principle is tied to the mathematical fact that, in order to construct a non-zero determinant, no two columns can be identical.

The Slater determinant also allows one to let the n spin orbitals undergo any orthogonal change without essentially changing the determinantal product function. This will allow transformation of delocalized molecular orbitals into localized orbitals.

ELECTRONIC CONFIGURATION AND ELECTRONIC STATES

Thus far, some of the general features of how orbitals are used to find approximate solutions to the Schroedinger equation have been outlined. We will now describe the relationship between these orbitals and electronic structure.

For a molecule of 2n electrons, the Schroedinger equation will yield 2n spin orbitals, each of discreet energy. If the orbital wavefunction is spin restricted, each spatial orbital may be occupied by two electrons: one of α spin and one of β spin, with degenerate orbital energies.

The different electronic configurations are represented schematically on the next few pages. Fig. I depicts a ground state configuration in which only the first n orbitals are occupied. The configuration is also termed "closed shell", as each orbital has the maximum of two electrons. The closed shell configuration is also called a "singlet state", since S = 0 due to the identical number of α and β electrons.

If the number of electrons is odd, as is shown in Fig. II, the ground state electronic configuration is

$$\psi_1^2 \psi_2^2 \cdots \psi_n^2 \psi_{n+1}^2 .$$

This ground state configuration is one short of the maximum in the highest energy orbital and is termed "open shell". Chemically, this state is a quantum mechanical representation of a free radical. Also, since there are an odd number of electrons, $M_s = \pm\frac{1}{2}$ and the spin state is termed "doublet".

Ground state configurations are the most stable (lowest energy). If a molecule absorbs energy, however, an electron may be moved to a higher orbital and the configuration is known as an "excited state", as shown in Figs. III and IV.

In Fig. III, the singlet, closed shell configuration has been excited, resulting in two unpaired electrons. Four possible values of M_s are, therefore, possible. Both unpaired electrons may have parallel spins of $\alpha\alpha$ ($M_s = 1$) or $\beta\beta$ ($M_s = -1$). The antiparallel configurations both result in $M_s = 0$, so there are two spin states which are present.

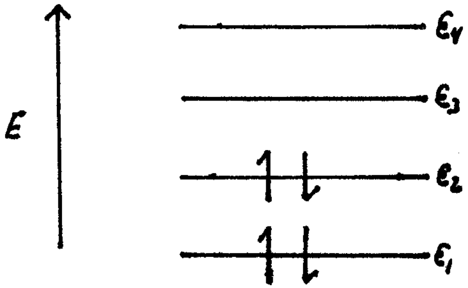


Fig. I. A Ground State, Closed Shell Electronic Configuration.

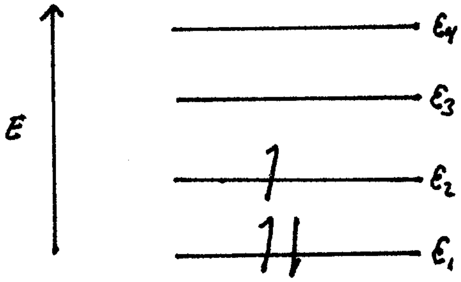


Fig. II. A Ground State, Open Shell Electronic Configuration.

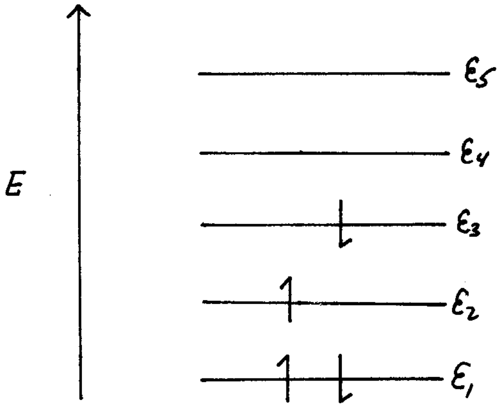


Fig. III. The Excitation of a Closed Shell Electronic Configuration.

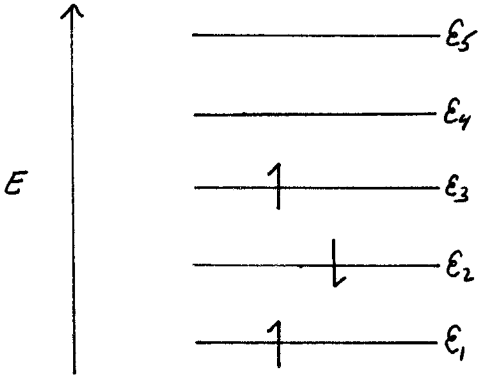


Fig. IV. The Excitation of an Open Shell Electronic Configuration.

A triplet state results from the combination of the two parallel states and one of the antiparallel states, and a singlet state results from the remaining antiparallel state. Surprisingly, the singlet and the triplet state have different energies, even though the orbitals which are occupied are the same. This is due to differing repulsion forces, with the triplet usually resulting in a lower energy.

In Fig. IV., there are three unpaired electrons, resulting from the excitation of the open shell configuration in Fig. II. Using reasoning similar to that employed above, it can be shown that this configuration yields one quartet state of lowest energy and two doublet states of differing energies.

THE USE OF ATOMIC ORBITALS IN MOLECULAR ORBITAL THEORY

Since each atomic orbital of the atoms in a molecule is largely dominant in the area of its own nucleus, a satisfactory approximation to the molecular orbital involves a linear combination of the atomic orbitals, with appropriate weighting coefficients.

If ϕ_u (u = 1, 2, ...) is a set of atomic orbital functions, then ψ_i is represented by

$$\psi_i = c_{1i}\phi_1 + c_{2i}\phi_2 + \dots,$$

where the c's are the weighting coefficients. The problem of finding the molecular wavefunction is thus reduced to the problem of defining the coefficients.

Spherical polar coordinates simplify the math and the one electron solution to the Schroedinger equation is written

$$\phi(r,\theta,\phi) = R_{nl}(r)\, Y_{lm}(\theta,\phi).$$

Slater-type orbitals (STO) are often used to approximate the radial part of the function, $R_{nl}(r)$. The functions are nodeless and are written

$$R_{nl}(r) = (2)^{n+\frac{1}{2}} (2n)^{-\frac{1}{2}} r^{n-1} \exp(-\xi r),$$

where $\xi = (Z-S)/n^*$. Here S is a screening constant and n^* is an effective principle quantum number.

The major drawback to the STO approximation, however, is that the orbitals are not mutually orthogonal, due to neglect of radial nodes. This problem is usually corrected by factoring out the overlap between the orbitals.

SELF-CONSISTENT FIELD MOLECULAR ORBITAL THEORY

Now that a general description of quantum mechanical theory has been outlined, it is necessary to present a means of doing actual calculations. The approach used, called the Self-Consistent Field Molecular Orbital Theory (SCF Theory) , was first explained by Hartree and is based on the Variational method, discussed earlier. A description of some of the aspects of SCF theory follows.

THE ENERGY EXPRESSION FOR A CLOSED SHELL SYSTEM

The Variational method involves manipulating the coefficients of our basis set functions to find a minimum energy. It is,therefore, necessary to have derived an expression for the energy expectation value that is suitable for this type of manipulation.

The orbital wavefunction, Ψ, was earlier shown to be

$$\Psi = \eta \sum_{P} (-1)^{P} \, P(\psi_1(1)\alpha(1) \; \psi_1(2)\beta(2) \; \ldots \; \psi_n(2n)\beta(2n)),$$

where P is a permutation of 1, 2, ... 2n for the coordinates of the electrons.

The normalization constant, η, is found by integrating the following expression over the spin and spatial coordinates of all the electrons.

$$\int \Psi^{*}\Psi \, d\tau_1 \ldots d\tau_{2n} = \eta^2 \sum_{P} \sum_{P'} (-1)^{P}(-1)^{P'} \int \ldots \int P(\psi_1(1)\alpha(1) \ldots \ldots \psi_{2n}(2n)\beta(2n)) \times P'(\psi_1(1)\alpha(1) \ldots \psi_{2n}(2n)\beta(2n)) \, d\tau_1 \ldots \ldots d\tau_{2n}.$$

There is a double summation over all permutations P and P'. The multiple integral will however, vanish unless P = P', since the integration will violate the condition of orthogonality. One can then simplify the expression for η to

$$\eta = ((2n)!)^{-\frac{1}{2}}.$$

We can now evaluate the expectation value for energy

$$\langle \Psi | H | \Psi \rangle.$$

The Hamiltonian is first expressed as the sum of a one electron and a two electron part

$$\langle \Psi | H | \Psi \rangle = \langle \Psi | H_1 | \Psi \rangle + \langle \Psi | H_2 | \Psi \rangle .$$

For the one electron portion of the Hamiltonian, it can be shown that, since all electrons are indistinguishable and since we are working with a restriction of orthonormality, (so all terms in which P $\neq$ P' are eliminated)

$$\langle \Psi | H | \Psi \rangle = ((2n-1)!)^{-1} \times \sum_{P} \int \ldots \int P(\psi_1(1)\alpha(1)\psi_1(2)\beta(2) \ldots \ldots) H^{core}(1) \times P(\psi_1(1)\alpha(1)\psi_1(2)\beta(2) \ldots .) \, d\tau_1 \ldots d\tau_{2n}.$$

Integration of this expression over all electrons gives unity. Also, H^{core} is independent of spin and is, therefore, equal to unity for electron one.

Since there are two electrons in each orbital, the final expression is

$$\langle \Psi | H_1 | \Psi \rangle = 2\sum_{i=1}^{n} H_{ii},$$

where $H_{ii} = \int \psi_i(1)^* H^{core} \psi_i(1) d\tau_1$. This time, however, P' need not equal P. P' can differ from P by interchanging the assignation of electrons 1 and 2. Thus, there are two cases; each will be considered seperately.

<u>Case I</u>: P = P'

There are two possibilities here. Either electrons 1 and 2 are assigned to different spatial molecular orbitals i and j, in which case there are four possible spin permutations and

$$J_{ij} = \iint \psi_i^*(1)\psi_j(2)\, 1/r_{ij}\, \psi_i(1)\psi_j(2)\, d\tau_1\, d\tau_2,$$

or the two electrons may be assigned to the same orbital and they must, therefore, have opposite spins. The contribution is then

$$2\sum_{i}\sum_{j\neq i} J_{ij} + \sum_{i} J_{ii}.$$

<u>Case II</u>: P differs from P' by interchanging the assignation of electrons 1 and 2

As in Case I, the electrons may be placed in different orbitals, in which case the contribution becomes

$$-\tfrac{1}{2}K_{ij},$$

where $K_{ij} = \iint \psi_i^*(1)\psi_j^*(2)\, 1/r_{ij}\, \psi_j(1)\psi_i(2)\, d\tau_1\, d\tau_2$,
or the electrons may be assigned to the same spatial orbital. In this instance, they must have different spin and the integrals vanish.

The final expression for the electronic energy is, therefore,

$$\mathcal{E} = 2\sum_{i}^{n} H_{ii} + \sum_{i}^{n} J_{ii} + \sum_{i}^{n}\sum_{j\neq i}^{n} (2J_{ij} - K_{ij}).$$

Since $K_{ii} = J_{ii}$, this expression may be simplified to

$$\mathcal{E} = 2\sum_{i}^{n} H_{ii} + \sum_{i}^{n}\sum_{j}^{n} (2J_{ij} - K_{ij}).$$

K_{ij} and J_{ij} are known as coulomb integrals and exchange integrals, respectively.

One can relate the terms in the above equations to physical properties as follows. The one-electron integral, H_{ii}, represents the energy of an electron in a molecular orbital, ψ_i, in the field of the bare nuclei. It is multiplied by 2 since there are 2 electrons per orbital. J_{ij}, the two electron integral, represents the interactions of the smoothed out charge distributions, $\psi_i^*\psi_i$ and $\psi_j^*\psi_j$. The exchange integral, K_{ij}, reduces the energy of interaction between electrons with parallel spin in different orbitals, ψ_i and ψ_j.

One may also write an expression for one-electron orbital energies

$$\mathcal{E}_i = H_{ii} + \sum_j (2J_{ij} - K_{ij}),$$

which reflects the energy of an electron in ψ_i interacting with the core and the other electrons. This expression is often called Koopman's ionization potential.

$\mathcal{E}$ is not, however, equal to the sum of $\mathcal{E}_i$ because the sum would include each electron-electron interaction twice. The second term in the following expression corrects for this

$$\mathcal{E} = 2 \sum_i \mathcal{E}_i - \sum_i \sum_j (2J_{ij} - K_{ij}),$$

or

$$\mathcal{E} = \sum_i (\mathcal{E}_i + H_{ii}).$$

THE HARTREE FOCK-EQUATIONS FOR MOLECULAR ORBITALS

Thus far the determinant form for many electron wavefunctions and the electronic energy function have been established. One now needs to write out the details for the actual determination of the orbitals, ψ_i, for a closed shell system. These equations are called Hartree-Fock equations.

According to the variational theorem, we need to find the lowest energy to arrive at the closest approximation to the actual energy. In addition, the orbitals must remain orthonormal or

$$S_{ij} = \int \psi_i(1)\psi_j(1)\, d\tau_1 = \delta_{ij}.$$

This process involves minimization of the function

$$G = \mathcal{E} - 2 \sum_i \sum_j \varepsilon_{ij} S_{ij} = 2 \sum_i H_{ii} + \sum_i \sum_j (2J_{ij} - K_{ij}) - 2 \sum_i \sum_j \varepsilon_{ij} S_{ij} \, ,$$

where ε_{ij} are constants that are yet to be determined. We need to find a point that is stationary, or

$$\delta G = 2 \sum_i \delta H_{ii} + \sum_i \sum_j (2 \delta J_{ij} - \delta K_{ij}) - 2 \sum_i \sum_j \varepsilon_{ij} \delta S_{ij},$$

where

$$\delta H_{ii} = \int \delta\chi_i^*(1) H^{core}(1) \chi_i (1) \delta\tau_1 \quad + \text{ complex conjugate}$$

and the other partial derivative expressions may be written in a similar fashion.

It can also be shown that

$$F\chi_i = \sum_j \varepsilon_{ij} \chi_j,$$

where

$$F = (H^{core} + \sum_j (2J_j - K_j)).$$

F may be interpreted as an effective one electron Hamiltonian for the electron in the molecular environment, with H^{core} representing the one electron-bare nuclei interaction, J_i ($= K_i$) is the potential due to the other electrons in χ_i.

Furthermore, by using matrix algebra, one can show that the equation above can be written as a standard eigenvalue equation

$$F\chi_i = \varepsilon_i \chi_i,$$

where $i = 1, \ldots, n$. These are known as Hartree-Fock equations.

The general procedure involved in solving these equations is a trial and error process. Different guesses are made for $\chi_1, \chi_2, \ldots$, until the orbital no longer changes (within a given energy tolerance) on further iteration. The orbitals are then said to be self-consistent.

The expression for the eigenvalues of the Hartree-Fock equations is

$$\varepsilon_i = H_{ii}^{core} + \sum_j (2J_{ij} - K_{ij}).$$

LCAO MOLECULAR ORBITALS FOR CLOSED SHELL SYSTEMS

Hartree-Fock orbitals are best approximated by a linear combination of atomic orbitals (LCAO) as was outlined earlier. In this method each orbital is defined by

$$\psi_i = \sum_u c_{ui}\phi_u,$$

where ϕ_u are real atomic functions. Since the orbitals, ψ_i, are required to be orthonormal, it is necessary that the number of atomic orbitals in the basis set be greater than or equal to the number of occupied molecular orbitals and

$$\sum_{uv} c_{ui}^{*} \; c_{vj} S_{uv} = \delta_{ij},$$

where δ_{ij} is the Kroneker Delta and

$$S_{uv} = \int \phi_u \,(1)\, \phi_v(1)\, \delta\tau_{ij},$$

is the overlap integral.

The molecular orbitals may be derived within any accuracy, depending on the number of basis functions employed in the function.

The matrix of elements, P_{uv}, is known as the density matris, as $P_{uv}S_{uv}$ is an indication of the extent of chemical bonding between two atoms in a molecule, with

$$\sum_{uv} P_{uv} S_{uv} = \sum_{uv} P_{uv} \int \phi_u(R)\phi_v(R)\; dR = \int \rho(R)\; dR = 2n.$$

We may also write our energy functions in terms of the atomic orbitals

$$H_{ii} = \sum_{uv} c_{ui}^{*} \; c_{vi} \; H_{uv},$$

where

$$H_{uv} = \int \phi_u(1)\, H^{core}\, \phi_v(1)\, d\tau_1,$$

and

$$J_{ij} = \sum_{uv\lambda\sigma} c^*_{ui}\, c^*_{\lambda j}\, c_{vi}\, c_{\sigma j}\, (uv/\lambda\sigma),$$

$$K_{ij} = \sum_{uv\lambda\sigma} c^*_{ui}\, c^*_{\lambda j}\, c_{vi}\, c_{\sigma j}\, (u\lambda/v\sigma),$$

where $(uv/\lambda\sigma)$ is the general two-electron interaction integral over atomic orbitals

$$(uv/\lambda\sigma) = \iint \phi_u(1)\phi_v(1)\; 1/r_{12}\; \phi_\lambda(2)\phi_\sigma(2)\, d\tau_1\, d\tau_2.$$

Finally, the total electronic energy expression may be written

$$\mathcal{E} = \sum_{uv} P_{uv}\, H_{uv} + \tfrac{1}{2} \sum_{uv\lambda\sigma} P_{uv}\, P_{\lambda\sigma} \times \left((uv/\lambda\sigma) - \tfrac{1}{2}(u\lambda/v\sigma) \right).$$

We now need to find the optimum values of c_{ui} which lead to a set of SCF LCAO's where

$$\delta\psi_i = \sum_u \delta c_{ui}\, \phi_u$$

and it can be shown that the equations can take the final algebraic form

$$\sum_v \left(F_{uv} - \mathcal{E}_i S_{uv}\right) c_{vi} = 0,$$

with

$$F_{uv} = H_{uv} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left((uv/\lambda\sigma) - \tfrac{1}{2}(u\lambda/v\sigma) \right).$$

These are called Roothaan's equations.

Since the Roothaan equations for the LCAO-SCF are cubic, the equations have to be solved by an iterative procedure. The coefficients are determined by matrix manipulation, with

$$C = S^{-\frac{1}{2}}\, C^{\lambda}.$$

First an initial set of coefficients, c, are assumed. The corresponding density matrix, P_{uv}, is generated and a first guess at F_{uv} is computed. The process is repeated until a given energy tolerance is reached.

MOLECULAR ORBITALS FOR OPEN-SHELL SYSTEMS

For open shell systems, the same basic SCF-LCAO iterative procedure is employed, with some slight variations. A rigorous treatment of these systems will not be presented here.

APPROXIMATE MOLECULAR ORBITAL THEORIES

Thus far we have outlined what is known as an "*ab initio*" method for computing wavefunctions and energies. This process is, however, quite time consuming and restrictions on computer time have necessitated further approximation. These approximations avoid the evaluation of some of the integrals by replacing them with experimental data. Thus, we will be describing what have become known as "semi-empirical" methods.

In general, any approximate LCAO-MO method must adhere to the following requirements:

1. it must allow one to do calculations on relatively large molecules, without excessive use of computer time;
2. the major physical and chemical properties of the system must be retained;
3. there should be no bias towards any conventional bonding or structural theory;
4. the results must be easily interpreted;
5. the theory should be general enough to take into account all of the chemically effective electrons.

While many different approximation methods have been and are being developed, each with its own inherent strengths and weaknesses, we will restrict our discussion to the complete neglect of differential overlap (CNDO) method, which was used in these calculations.

THE CNDO METHOD

The CNDO method was introduced by Pople, Santry and Segal in 1965. Only valence electrons are treated explicitly, with other electrons and the nucleus considered as part of a rigid core.

The major approximation in CNDO theory is the systematic neglect of electron repulsion integrals having uniformelly small values. That is

$$(uv/\lambda\sigma) = (uu/\lambda\lambda)\,\delta_{uv}\,\delta_{\lambda\sigma}$$

and

$$S_{uv} = \int \phi_u(1)\,\phi_v(1)\,d\tau_1$$

are neglected.

The core integrals

$$H_{uv} = \int \phi_u(1)\, H^{core}\, \phi_v(1)\, d\tau_1$$

are not neglected but are treated semi-emperically.

If the Zero Differential Overlap (ZDO) approximation were applied for all the integrals, the following equations would hold

$$\sum_v F_{uv}\, c_{vi} = \epsilon_i\, c_{ui},$$

with the elements of the Fock matrix given by

$$F_{uu} = H_{uu} - \tfrac{1}{2}P_{uu}\,(uu/uu) + \sum_\lambda P_{\lambda\lambda}(uu/\lambda\lambda)$$

and

$$F_{uv} = H_{uv} - \tfrac{1}{2}P_{uv}\,(uu/vv),$$

with $u \neq v$.

Rotational invariance must be restored, however. This is accomplished by making the remaining two-electron integrals depend only on the nature of atoms A and B, to which ϕ_u and ϕ_v belong, and not on the actual orbital. Thus,

$$(uu/\lambda\lambda) = \gamma_{AB},$$

with all u on atom A and all λ on atom B.

Using the above equation, the CNDO expressions for the Fock Hamiltonian matrix elements now simplify to

$$F_{uu} = H_{uu} - \tfrac{1}{2}\,P_{uu}\,\gamma_{AA} + \sum_B P_{BB}\,\gamma_{AB},$$

ϕ_u on A, and

$$F_{uv} = H_{uv} - \tfrac{1}{2}P_{uv}\gamma_{AB},$$

ϕ_u on A and ϕ_v on B, where $P_{BB} = \sum_{\lambda}^{B} P_{\lambda\lambda}$.

A series of related approximations is now applied to the one-electron Hamiltonian

$$H = -\tfrac{1}{2}\nabla^2 - \sum_B V_B,$$

where $-V_B$ is the core potential. One center terms are treated semi-empirically, two center terms are approximated by neglect of differential overlap and resonance integrals are treated semiemperically.

To summarize the approximations made in CNDO:

1. the overlap matrix is replaced by the unit matrix in the Roothaan equations and the overlap integrals, S_{uv}, are neglected in normalization of the MO's;

2. differential overlap is neglected in all two electron integrals, so

$$(uu/\lambda\sigma) = \delta_{uv}\delta_{\lambda\sigma}(uu/\lambda\lambda);$$

3. the remaining set of coulomb-type integrals is reduced to one value per atom

$$(uu/\lambda\lambda) = \gamma_{AB},$$

with ϕ_u on A and ϕ_λ on B;

4. monatomic differential overlap is neglected in the interaction integrals involving the cores of other atoms

$$(u \mid V_B \mid v) = \delta_{uv} \quad V_{AB};$$

5. diatomic, off-diagonal core matrix elements are assumed to be proportional to the corresponding overlap integrals

$$H_{uv} = \beta_{AB}^{0} S_{uv} \quad \phi_u \text{ on A}, \phi_v \text{ on B}.$$

These approximations result in the matrix elements of the Fock Hamiltonian being reduced to

$$F_{uu} = U_{uu} + (P_{AA} - \tfrac{1}{2}P_{uu})\gamma_{AA} + \sum_{B \neq A} (P_{BB}\gamma_{AB} - V_{AB})$$

and

$$F_{uv} = \beta_{AB}^{o} \int_{uv} - \tfrac{1}{2}P_{uv}\gamma_{AB}, \; u \neq v.$$

This allows a set of CNDO coefficients, c_{ui}, and a corresponding density matrix to be calculated. The total energy can then be derived from

$$\mathcal{E}_{TOTAL} = \tfrac{1}{2}\sum_{uv} P_{uv}\,(H_{uv} + F_{uv}) + \sum_{A<B} Z_A Z_B R_{AB}^{-1} \; ,$$

which can be expressed as the sum of monatomic and diatomic contributions

$$\mathcal{E}_{TOTAL} = \sum_{A} \epsilon_A + \sum_{A<B} \mathcal{E}_{AB}.$$

Two different versions of CNDO are available: CNDO/1 and CNDO/2. These differ in the procedures employed in acquiring values for $\int_{uv}$, U_{uu}, V_{AB}, γ_{AB} and β_{AB}^{o}. The latter method has been shown to be more successful, and is the method that was employed in our calculations.

BIBLIOGRAPHY

Atkins, P.W. PHYSICAL CHEMISTRY. San Francisco: W.H. Freeman and Company. 1978.

Dewar, M.J. THE MOLECULAR ORBITAL THEORY OF ORGANIC CHEMISTRY. New York: McGraw-Hill Book Company. 1969.

Murrell, J.N. and A.J. Harget. SEMI-EMPIRICAL SELF-CONSISTENT-FIELD MOLECULAR ORBITAL THEORY OF MOLECULES. New York: John Wiley and Sons. 1972.

Pople, J.A. and D.L. Bevaridge. APPROXIMATE MOLECULAR ORBITAL THEORY. New York: McGraw-Hill Book Company. 1970.

Taylor, R. METAPHYSICS. New Jersey: Prentice-Hall, Inc. 1974.

CHAPTER IV

RESULTS AND DISCUSSION

Some possible reaction pathways for a model RNA substrate for Ribonuclease were traced at the CNDO/2 level, using the CINMIN optimization routine. (Listings or taped copies of these programs may be obtained from QCPE).

The CINMIN optimization routine automatically changes bond lengths, bond angles and dihedral angles to effect the minimization of a molecular geometry. Initial r, θ, and ϕ coordinates for each atom are read into the program, with all atoms defined in terms of other atoms. Dummy atoms, (atomic number = 9999), are often used to help define an atom. If a parameter is to be kept constant or is linked, symmetrically, to another parameter, a "0" is placed in the column.

For example, Fig. I shows a sample input for a hydrogen atom. The atomic number, "1", is listed and the hydrogen is assigned a number, "8", in defining the molecule. This 8th atom is defined in terms of atoms numbered 1, 3, and 6, respectively. The initial bond length between atoms 6 and 8 is listed as 1.03A; the initial bond angle, 3-6-8, is listed as 106^{0} and the initial dihedral angle, 1-3-6-8, is listed as 0.00^{0}. Since all three parameters are to be optimized, there is a "1" listed in each IA column.

	AN	NA	NB	NC	ND	BND LEN	IA1	BND ANG	IA2	DIH ANG	IA3
H	1	1	3	6	8	1.03	1	106	1	0.00	1

FIG. I. Sample Input for CINMIN.

Each variable is then considered separately and "steps" are taken from the initial value, at preset increments. (These increments are automatically adjusted by the program to speed the optimization process). Each step results in the calculation of a new electronic energy, until a step results in an increase in electronic energy. When this occurs, a quadratic fit is made of all of the energy calculations for the prior steps and the minimum energy, (from the base of the parabola) is calculated at that geometry. This process of minimizing one variable is called a "line-search".

One line-search is done for each variable in what is termed a "cycle". At the beginning of each new cycle the directions of the steps for each variable are improved, based on the results obtained in the preceeding cycle. This process is called a Gram-Schmidt orthogonalization. The program terminates when the energy calculations for successive cycles are within a specified energy tolerance.

Geometries were optimized for the structures shown in Fig. II and improved energy calculations were done using GAUSS/70. The lowest activation energy path would be taken as the path actually followed for each step of the reaction. Work on partial optimization of near-transition state geometries had been started but the project was aborted when anomalous results were obtained.

The results obtained from the CNDO/2 calculations are presented on the following pages.

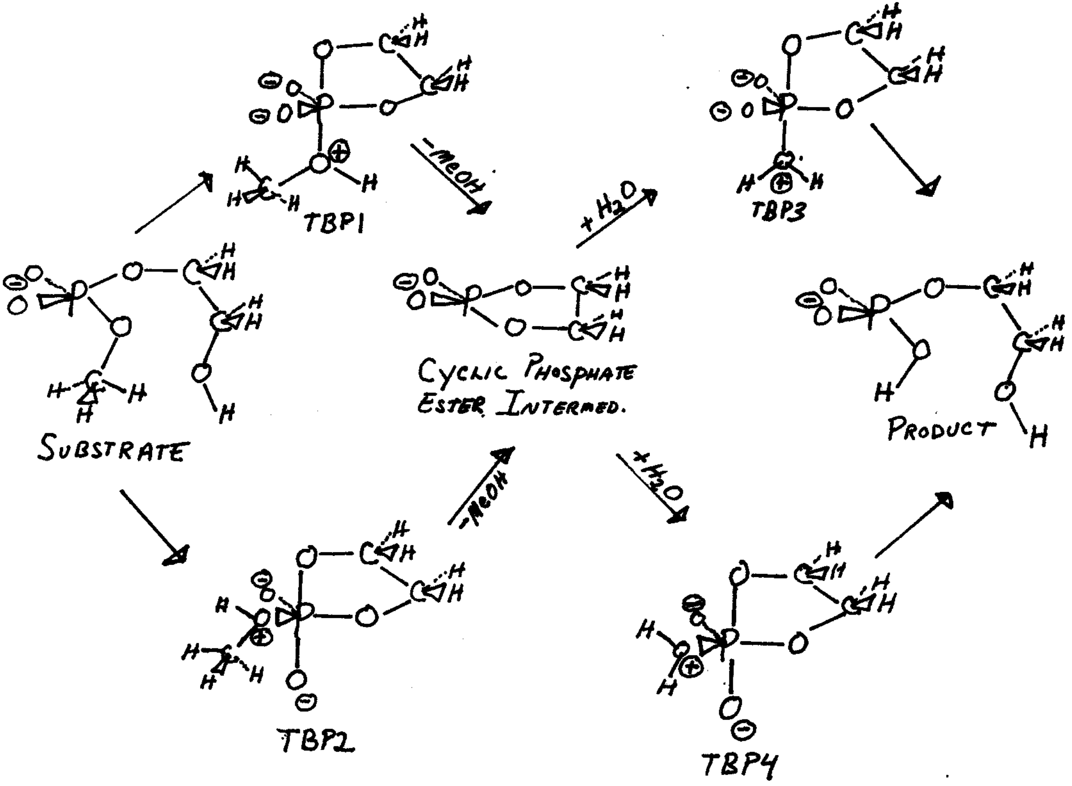


FIG. II. The Reaction Pathways Traced with CNDO/2 and CINMIN.

```
FINAL GEOMETRY

      AN   NA   NB   NC   ND    - -- -   BOND LENGTH     IA1     BOND ANGLE     IA2     DIH ANGLE     IA3

 P    15    0    0    0    1               0.000000      0           0.000     0          0.000      0
 D9999      0    0    1    2               1.000000      0           0.000     0          0.000      0
 D9999      0    1    2    3               1.000000      0          90.000     0          0.000      0
 D9999      3    2    1    4               1.000000      0          90.000     0         90.000      0
 O     8    3    2    1    5               1.698447      1         125.731     1          0.000      0
 O     8    3    2    1    6               1.693020      1         125.731     0        180.000      0
 O     8    3    2    1    7               1.783504      1          64.258     1         99.364      1
 O     8    3    2    1    8               1.790805      1          42.020     1        243.558      1
 C     6    2    1    8    9               1.384365      1          87.512     1        153.895      1
 H     1    1    8    9   10               1.090000      0         109.500     0        -13.576      1
 H     1    1    8    9   11               1.090000      0         109.500     0        106.424      0
 H     1    1    8    9   12               1.090000      0         109.500     0        226.424      0
 C     6    2    1    7   13               1.387173      1          86.786     1        299.983      1
 C     6    1    7   13   14               1.496141      1         122.867     1         33.406      1
 H     1    1    7   13   15               1.090000      0         108.530     1        153.406      0
 H     1    1    7   13   16               1.090000      0         108.530     0        273.406      0
 H     1    7   13   14   17               1.090000      0         116.641     1         63.857      1
 H     1    7   13   14   18               1.090000      0         116.641     0        303.657      0
 O     8    7   13   14   19               1.383675      1         104.640     1        184.518      1
 H     1   13   14   19   20               1.033261      1         107.014     1        178.557      1
```

TABLE I. Optimized Geometry for the Substrate.

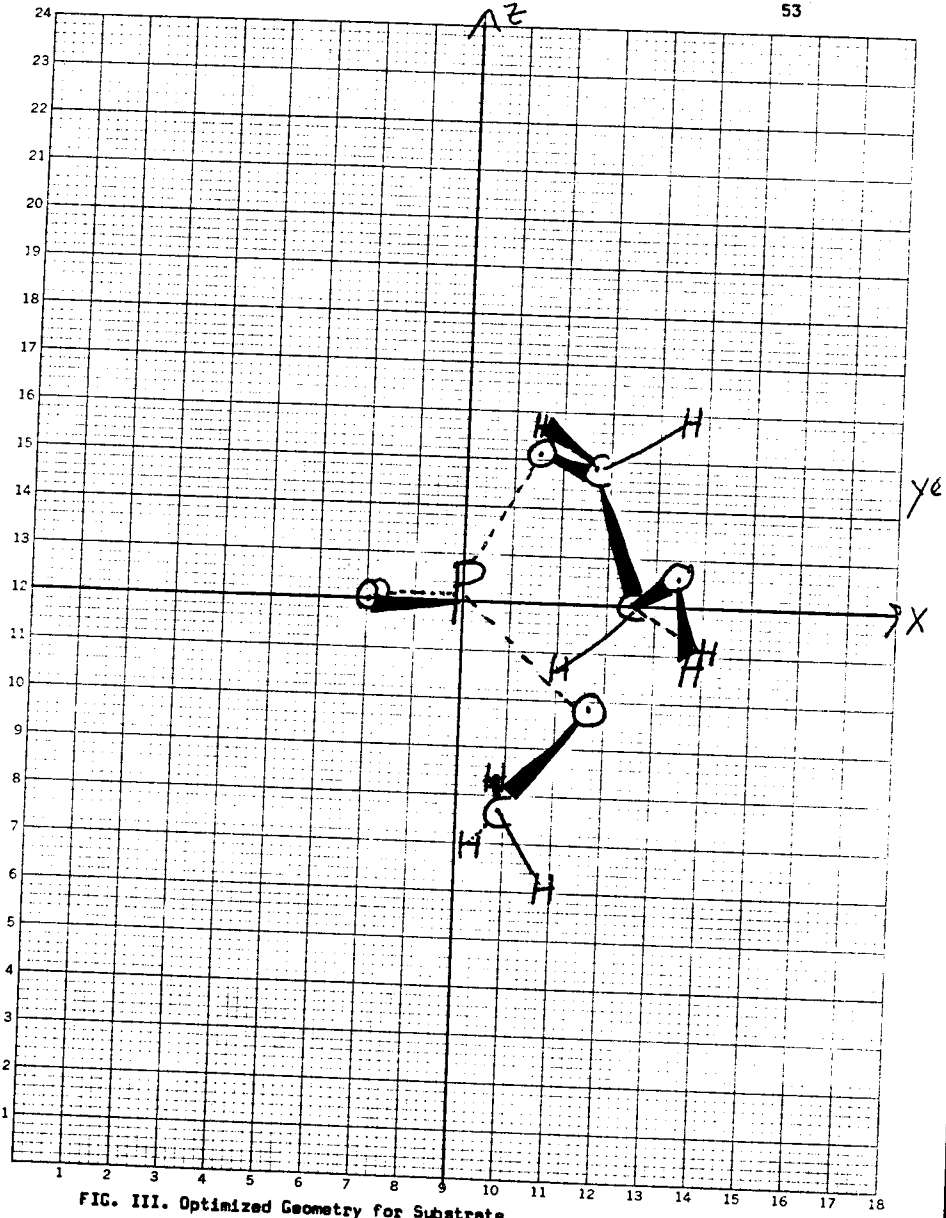


FIG. III. Optimized Geometry for Substrate

```
 FINAL GEOMETRY

        AN   NA   NB   NC   ND             BOND LENGTH      IA1      BOND ANGLE      IA2      DIH ANGLE      IA3

  P     15    0    0    0    1              0.000000        0          0.000         0          0.000        0
  C9999       0    0    1    2              1.000000        0          0.000         0          0.000        0
  O9999       0    1    2    3              1.000000        0         90.000         0          0.000        0
  C9999       3    2    1    4              1.000000        0         90.000         0         90.000        0
  O      8    3    2    1    5              1.695779        1        111.040         1          0.000        0
  O      8    3    2    1    6              1.697153        1        111.040         0        180.000        0
  O      8    3    2    1    7              1.811980        1         90.000         0         90.000        0
  O      8    3    2    1    8              1.931002        1         90.000         0        270.000        0
  C      6    2    1    8    9              1.386770        1         90.435         1        186.467        1
  H      1    1    8    9   10              1.090000        0        109.500         0        -15.392        1
  H      1    1    8    9   11              1.090000        0        109.500         0        104.608        0
  H      1    1    8    9   12              1.090000        0        109.500         0        224.608        0
  C      6    2    1    7   13              1.366859        1        112.780         1         -0.079        1
  C      6    1    7   13   14              1.461710        1        111.266         1          0.052        1
  H      1    1    7   13   15              1.090000        0        110.374         1        120.052        0
  H      1    1    7   13   16              1.090000        0        110.374         0        240.052        0
  H      1    7   13   14   17              1.090000        0        111.602         1        119.976        1
  H      1    7   13   14   18              1.090000        0        111.602         0        239.976        0
  O      8    4    1    2   19              0.779914        1        180.000         0          0.000        0
  H      1    2    1    8   20              1.041048        1         88.095         1        293.298        1
```

Table II. Optimized Geometry of TBP1.

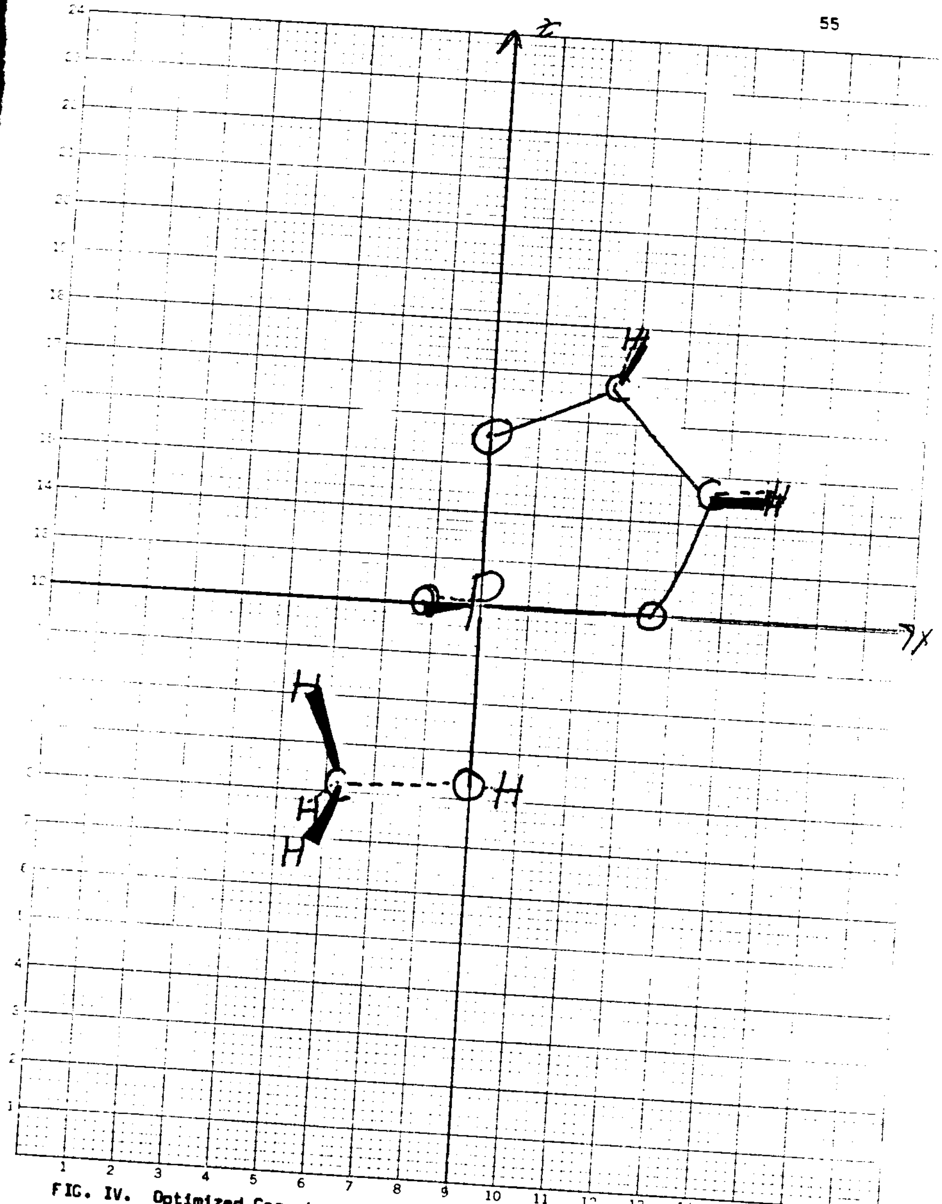


FIG. IV. Optimized Geometry for TBP1.

```
FINAL GEOMETRY

       AN    NA    NB    NC    ND               BOND LENGTH        IA1       BOND ANGLE        IA2       DIH ANGLE        IA3

 P     15    0     0     0     1                  0.000000          0           0.000           0           0.000          0
 D9999       0     0     1     2                  1.000000          0           0.000           0           0.000          0
 D9999       0     1     2     3                  1.000000          0          90.000           0           0.000          0
 D9999       3     2     1     4                  1.000000          0          90.000           0          90.000          0
 O      8    3     2     1     5                  1.723213          1         106.820           1           0.000          0
 O      8    3     2     1     6                  1.715343          1          90.000           0         270.000          0
 O      8    3     2     1     7                  1.820241          1          90.000           0          90.000          0
 O      8    3     2     1     8                  1.946787          1         106.820           0         180.000          0
 C      6    2     1     8     9                  1.392530          1          79.231           1         179.433          1
 H      1    1     8     9    10                  1.090000          0         109.500           0          36.927          1
 H      1    1     8     9    11                  1.090000          0         109.500           0         156.927          0
 H      1    1     8     9    12                  1.090000          0         109.500           0         276.927          0
 C      6    2     1     7    13                  1.370410          1         110.888           1          -1.352          1
 C      6    1     7    13    14                  1.481305          1         113.903           1           4.161          1
 H      1    1     7    13    15                  1.090000          0         109.310           1         124.161          0
 H      1    1     7    13    16                  1.090000          0         109.310           0         244.161          0
 H      1    7    13    14    17                  1.090000          0         111.580           1         114.190          1
 H      1    7    13    14    18                  1.090000          0         111.580           0         234.190          0
 O      8    4     1     2    19                  0.788559          1         180.000           0           0.000          0
 H      1    2     1     8    20                  1.055484          1          76.527           1         289.540          1
```

Table III. Optimized Geometry for TBP2.

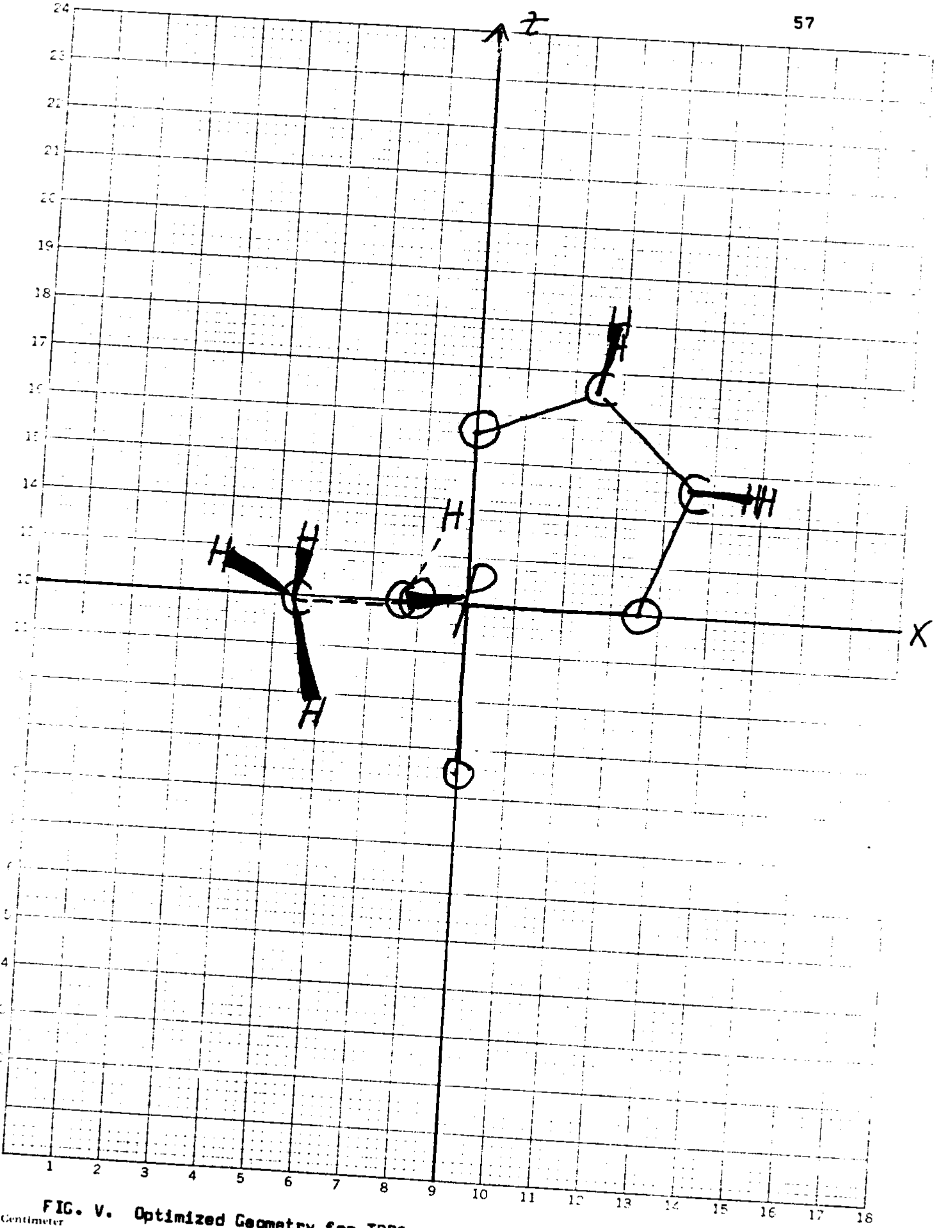


FIG. V. Optimized Geometry for TBP2.

```
FINAL GEOMETRY

       AN  NA  NB  NC  ND             BOND LENGTH     IA1     BOND ANGLE     IA2      DIH ANGLE     IA3

  P    15   0   0   0   1              0.000000        0         0.000        0         0.000        0
  D9999     0   0   1   2              1.000000        0         0.000        0         0.000        0
  D9999     0   1   2   3              1.000000        0        90.000        0         0.000        0
  D9999     3   2   1   4              1.000000        0        90.000        0        90.000        0
  O     8   3   2   1   5              1.670463        1       120.218        1         0.000        0
  O     8   3   2   1   6              1.671501        1       120.218        0       180.000        0
  O     8   3   2   1   7              1.788915        1        60.234        1        92.721        1
  O     8   3   2   1   8              1.792081        1        61.147        1       266.790        1
  C     6   2   1   7   9              1.397676        1        80.830        1       -13.696        1
  C     6   1   7   9  10              1.523329        1       125.838        1        -0.932        1
  H     1   1   7   9  11              1.090000        0       107.200        1       119.068        0
  H     1   1   7   9  12              1.090000        0       107.200        0       239.068        0
  H     1   1   8  10  13              1.090000        0       108.616        1       215.178        1
  H     1   1   8  10  14              1.090000        0       108.616        0        95.178        0
```

TABLE IV. Optimized Geometry for the Cyclic Phosphate Ester Intermediate.

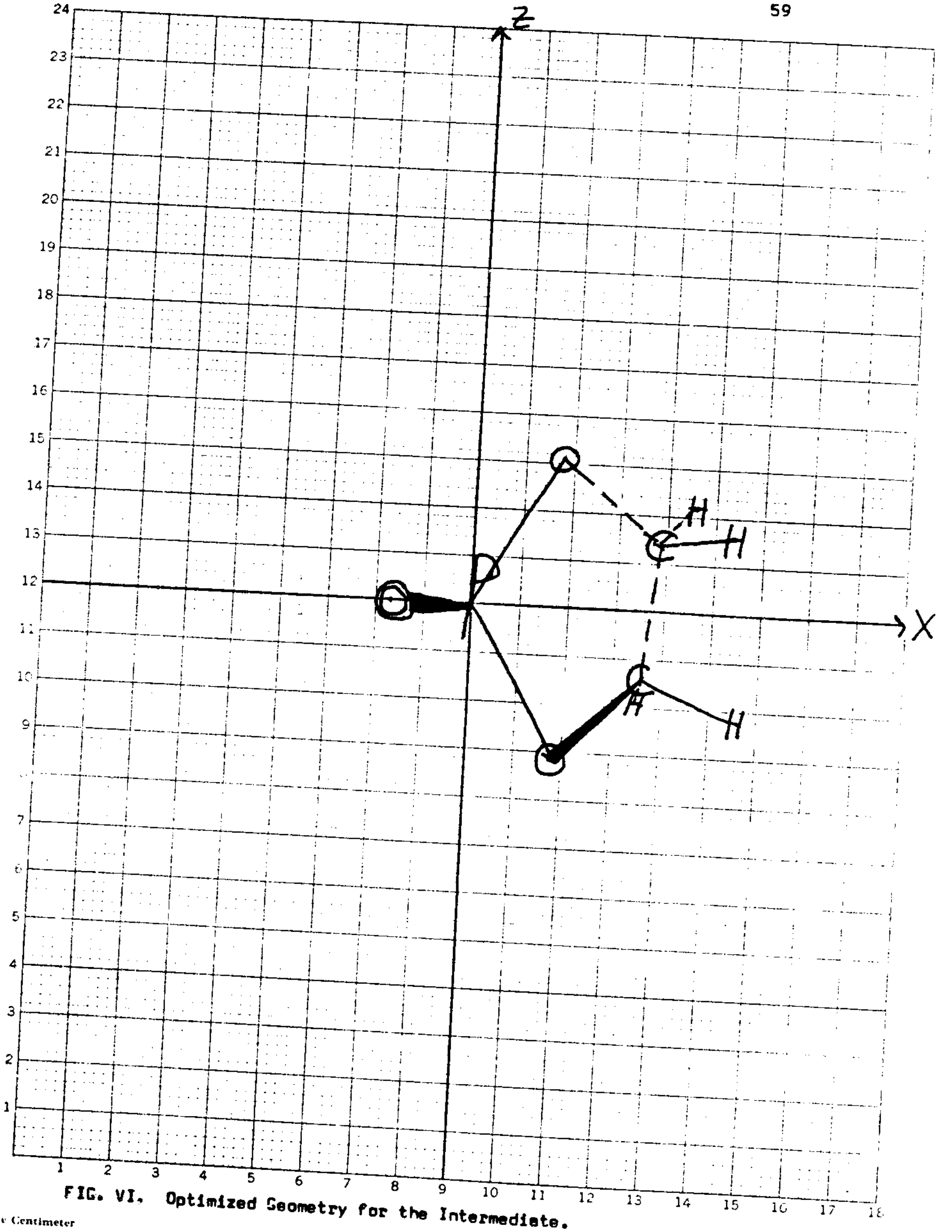


FIG. VI. Optimized Geometry for the Intermediate.

```
FINAL GEOMETRY

       AN   NA   NB   NC   ND                    BOND LENGTH     IA1      BOND ANGLE     IA2     DIH ANGLE      IA3

 P    15    0    0    0    1                       0.000000       0          0.000       0         0.000        0
 D9999      0    0    1    2                       1.000000       0          0.000       0         0.000        0
 D9999      0    1    2    3                       1.000000       0         90.000       0         0.000        0
 D9999      3    2    1    4                       1.000000       0         90.000       0        90.000        0
 O     8    3    2    1    5                       1.685495       1        116.390       1         0.000        0
 O     8    3    2    1    6                       1.687645       1        116.390       0       180.000        0
 O     8    3    2    1    7                       1.809877       1         90.000       0        90.000        0
 O     8    3    2    1    8                       1.934968       1         90.000       0       270.000        0
 H     1    2    1    8    9                       1.038170       1         87.830       1       182.531        1
 D9999      1    8    9   10                       1.090000       0        109.500       0         0.000        0
 D9999      1    8    9   11                       1.090000       0        109.500       0       120.000        0
 D9999      1    8    9   12                       1.090000       0        109.500       0       240.000        0
 C     6    2    1    7   13                       1.368241       1        113.328       1        -1.420        1
 C     6    1    7   13   14                       1.482364       1        110.334       1         0.697        1
 H     1    1    7   13   15                       1.090000       0        110.536       1       120.697        0
 H     1    1    7   13   16                       1.090000       0        110.536       0       240.697        0
 H     1    7   13   14   17                       1.090000       0        111.521       1       120.535        1
 H     1    7   13   14   18                       1.090000       0        111.521       0       240.535        0
 O     8    4    1    2   19                       0.774255       1        180.000       0         0.000        0
 H     1    2    1    8   20                       1.039005       1         87.280       1       288.829        1
```

TABLE V. Optimized Geometry for TBP3.

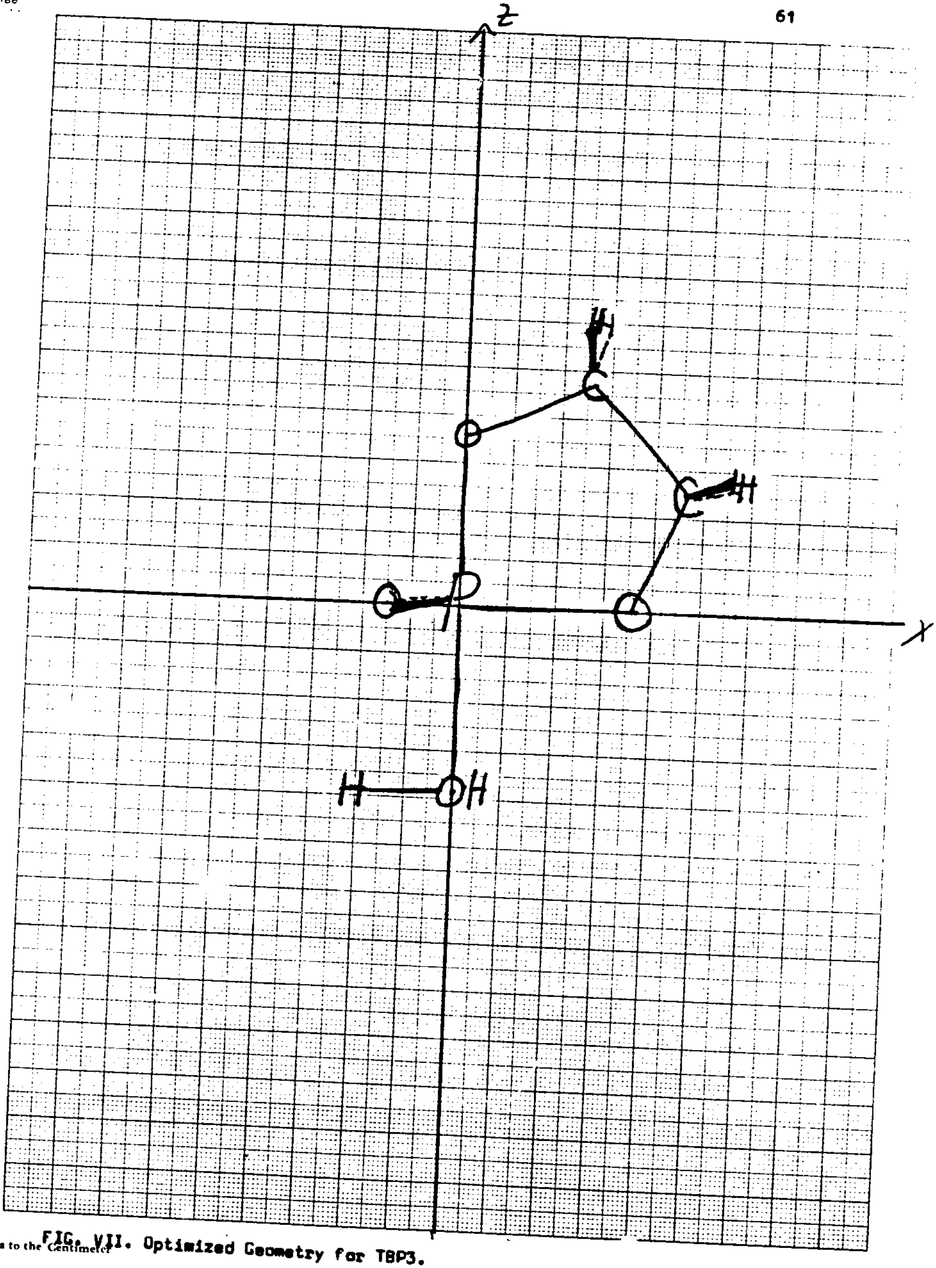


FIG. VII. Optimized Geometry for TBP3.

```
FINAL GEOMETRY

     AN   NA   NB   NC   ND              BOND LENGTH      IA1      BOND ANGLE      IA2      DIH ANGLE      IA3

 P   15    0    0    0    1                0.000000        0          0.000        0          0.000        0
 D9999     0    0    1    2                1.000000        0          0.000        0          0.000        0
 D9999     0    1    2    3                1.000000        0         90.000        0          0.000        0
 D9999     3    2    1    4                1.000000        0         90.000        0         90.000        0
 O    8    3    2    1    5                1.709912        1        119.891        1          0.000        0
 O    8    3    2    1    6                1.785740        1         90.000        0        270.000        0
 O    8    3    2    1    7                1.791116        1         90.000        0         90.000        0
 O    8    3    2    1    8                1.755788        1        119.891        0        180.000        0
 H    1    2    1    9    9                2.191901        1         68.950        1         97.617        1
 D9999     1    8    9   10                1.090000        0        109.500        0          0.000        0
 D9999     1    8    9   11                1.090000        0        109.500        0        120.000        0
 D9999     1    8    9   12                1.090000        0        109.500        0        240.000        0
 C    6    2    1    7   13                1.370493        1        107.717        1        -34.972        1
 C    6    1    7   13   14                1.511770        1        104.253        1          9.091        1
 H    1    1    7   13   15                1.090000        0        111.132        1        129.091        0
 H    1    1    7   13   16                1.090000        0        111.132        0        249.091        0
 H    1    7   13   14   17                1.090000        0        111.178        1        145.143        1
 H    1    7   13   14   18                1.090000        0        111.178        0        265.143        0
 O    8    4    1    2   19                0.762829        1        180.000        0          0.000        0
 H    1    2    1    8   20                1.033983        1        102.288        1        267.085        1
```

TABLE VI. Optimized Geometry for TBP4.

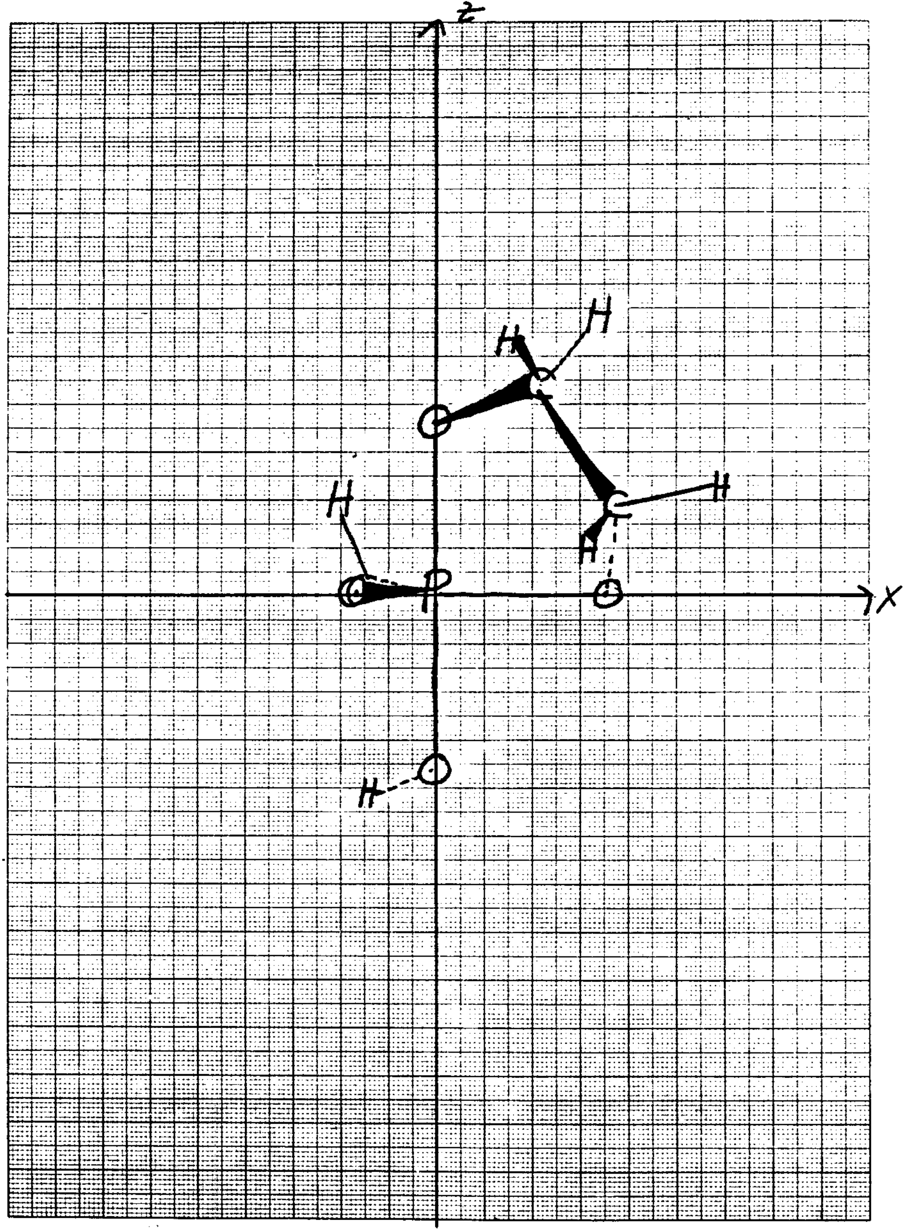


FIG. VIII. Optimized Geometry for TBP4.

```
FINAL GEOMETRY
                                                     BOND LENGTH      IA1      BOND ANGLE     IA2      DIH ANGLE     IA3
         AN     NA     NB     NC     ND
P        15      0      0      0      1               0.000000        0          0.000        0         0.000         0
D9999    99      0      0      1      2               1.000000        0          0.000        0         0.000         0
D9999    99      0      1      2      3               1.000000        0         90.000        0         0.000         0
D9999    99      3      2      1      4               1.000000        0         90.000        0        90.000         0
O         8      3      2      1      5               1.682186        1        126.096        1         0.000         0
O         8      3      2      1      6               1.689261        1        126.096        0       180.000         0
O         8      3      2      1      7               1.775582        1         58.204        1       103.660         1
O         8      3      2      1      8               1.773849        1         55.164        1       255.774         1
H         1      2      1      8      9               1.038111        1         98.775        1       214.053         1
D9999    99      1      8      9     10               1.090000        0        109.500        0         0.000         0
D9999    99      1      8      9     11               1.090000        0        109.500        0       120.000         0
D9999    99      1      8      9     12               1.090000        0        109.500        0       240.000         0
C         6      2      1      7     13               1.387807        1         86.297        1       308.379         1
C         6      1      7     13     14               1.498375        1        124.577        1        26.409         1
H         1      1      7     13     15               1.090000        0        107.969        1       146.409         0
H         1      1      7     13     16               1.090000        0        107.969        0       266.409         0
H         1      7     13     14     17               1.090000        0        116.337        1        64.415         1
H         1      7     13     14     18               1.090000        0        116.337        0       304.415         0
O         8      7     13     14     19               1.384063        1        104.347        1       185.123         1
H         1     13     14     19     20               1.033275        1        106.600        1       179.266         1
```

TABLE VII. Optimized Geometry for the Product.

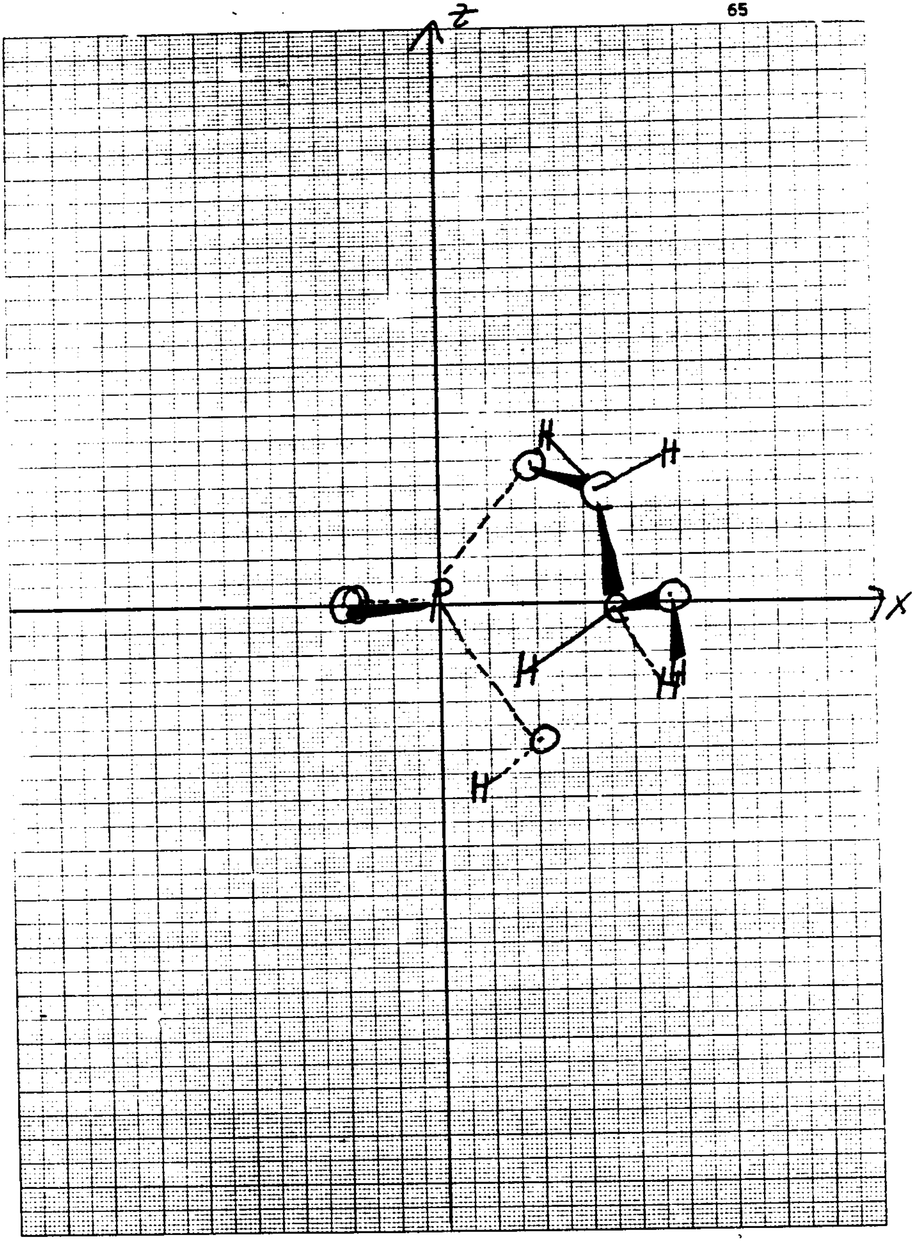


FIG. IX. Optimized Geometry for the Product.

```
FINAL GEOMETRY

     AN   NA   NB   NC   ND                BOND LENGTH    IA1     BOND ANGLE     IA2     DIH ANGLE      IA3

C    6    0    0    0    1                  0.000000      0         0.000      0        0.000        0
O    8    0    0    1    2                  1.371884      1         0.000      0        0.000        0
H    1    0    1    2    3                  1.033105      1       105.356      1        0.000        0
H    1    3    2    1    4                  1.090000      0       109.500      0      180.192        1
H    1    3    2    1    5                  1.090000      0       109.500      0      300.192        0
H    1    3    2    1    6                  1.090000      0       109.500      0       60.192        0
```

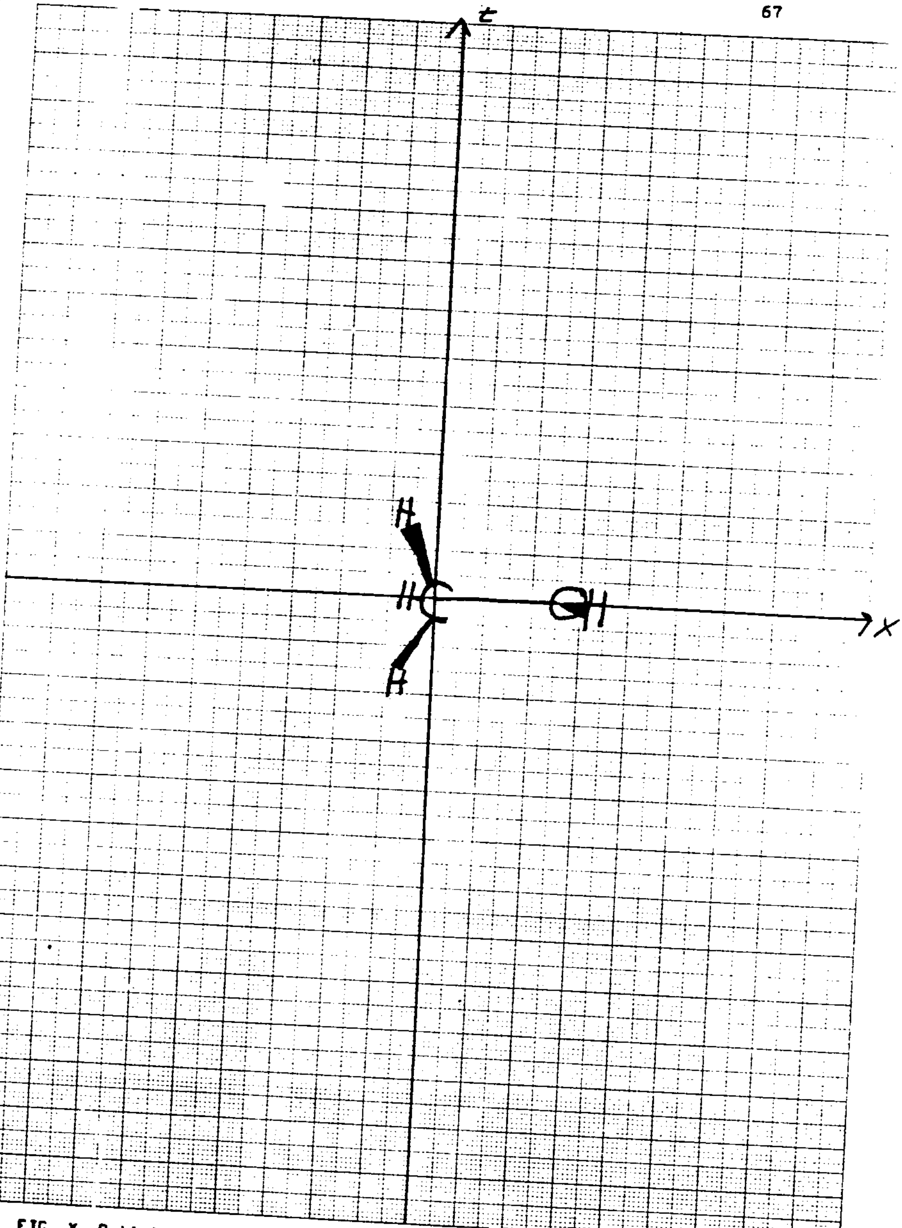
TABLE VIII. Optimized Geometry for Methanol

```
FINAL GEOMETRY

     AN   NA   NB   NC   ND                BOND LENGTH    IA1     BOND ANGLE     IA2     DIH ANGLE      IA3

H    1    0    0    0    1                  0.000000      0         0.000      0        0.000        0
O    8    0    0    1    2                  1.029553      1         0.000      0        0.000        0
H    1    0    1    2    3                  1.029476      1       102.565      1        0.000        0
```

TABLE IX . Optimized Geometry for Water.



10 Millimeters to the Centimeter

FIG. X. Optimized Geometry for Methanol.

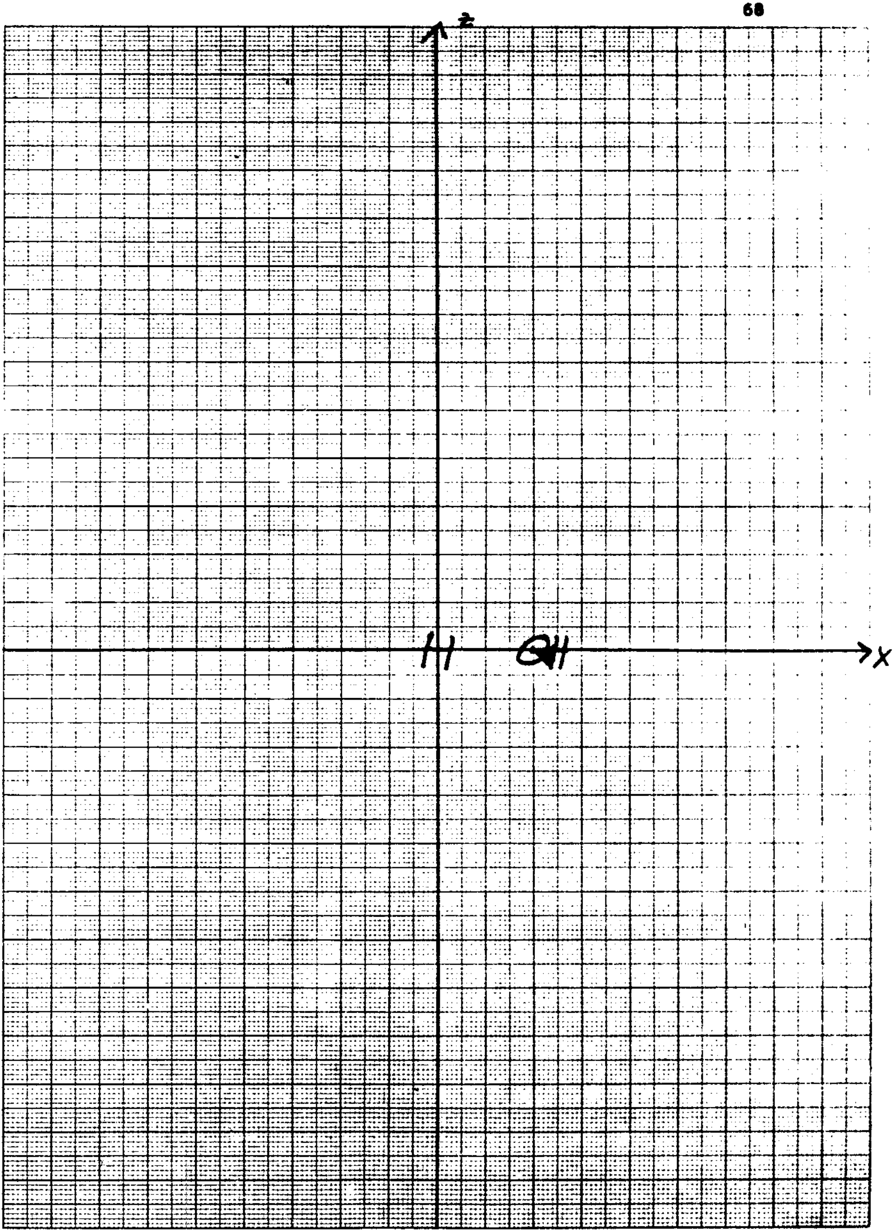


FIG. XI. Optimized Geometry for Water.

The figures appear reasonable but, as the energy calculations will show, CNDO/2 was either not an accurate predictor of optimized geometries or CNDO/2 failed to calculate accurate energies for the minimum geometries. If the latter is true then one would expect the GAUSS/70 calculations of the energies to improve the activation energy plots. This was not the case and it must be concluded that the optimized geometries are erroneous.

OPTIMIZED GEOMETRY		ENERGY (Hartrees)
Substrate	(+ water)	- 146.55570
TBP1	(+ water)	- 146.53453
TBP2	(+ water)	- 146.50494
Intermed.	(+ water and MeOH)	- 146.51692
TBP3	(+ MeOH)	- 146.52023
TBP4	(+ MeOH)	- 146.63277
Product	(+ MeOH)	- 146.51531

TABLE X. Calculated CNDO/2 Energies for Several Optimized Phosphates

These energies are plotted on the following page.

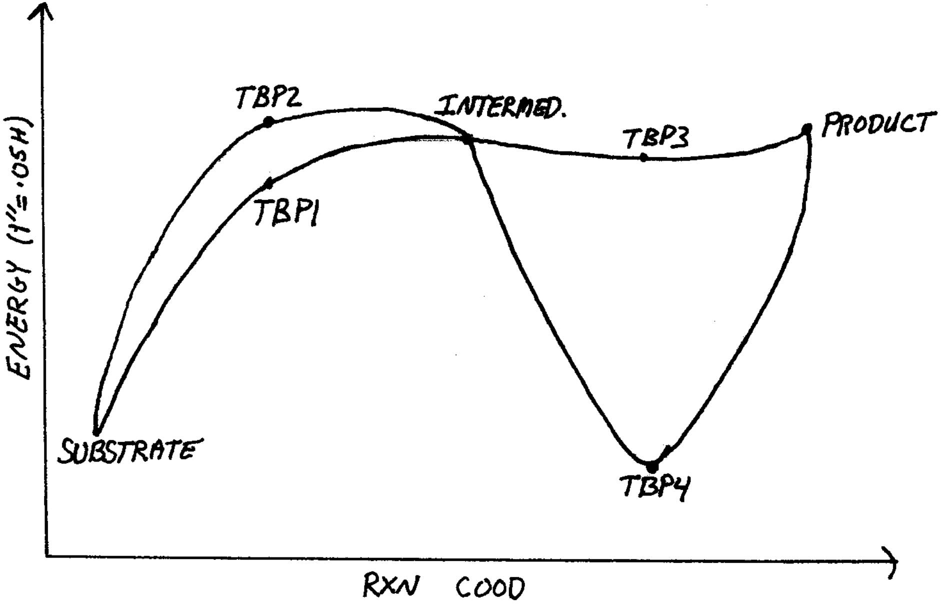


FIG. XII. Energy of Activation Plot for Optimized Geometries Using CNDO/2 Energies

As FIG. XII shows, the normal activation energy plot expected for each step of the reaction was not obtained. It was hoped, however, that GAUSS/70 calculations (done at Berkeley) for the optimized geometries would result in improved plots.

The GAUSS/70 approach differs from CNDO/2 in that GAUSS/70 does not make the semi-empirical estimations that are used in CNDO/2. All one and two electron integrals are evaluated explicitly, with semi-empirical approximations and neglect of integrals used in CNDO/2 being treated in an *ab-initio* manner in GAUSS/70. A rigorous treatment of CNDO/2 approximations may be found at the end of Chapter III.

The results obtained using GAUSS/70 are listed below.

OPTIMIZED GEOMETRY		ENERGY (Hartrees)
Substrate	(+ water)	- 897.47885
TBP1	(+ water)	- 897.55407
TBP2	(+ water)	(Did Not Converge)
Intermed.	(+ water and MeOH)	- 897.52154
TBP3	(+ MeOH)	-897. 58503
TBP4	(+ MeOH)	- 897.66853
Product	(+ MeOH)	- 897.52538

TABLE XI. Calculated GAUSS/70 Energies for Several Optimized Geometries.

These energies are plotted on the following page.

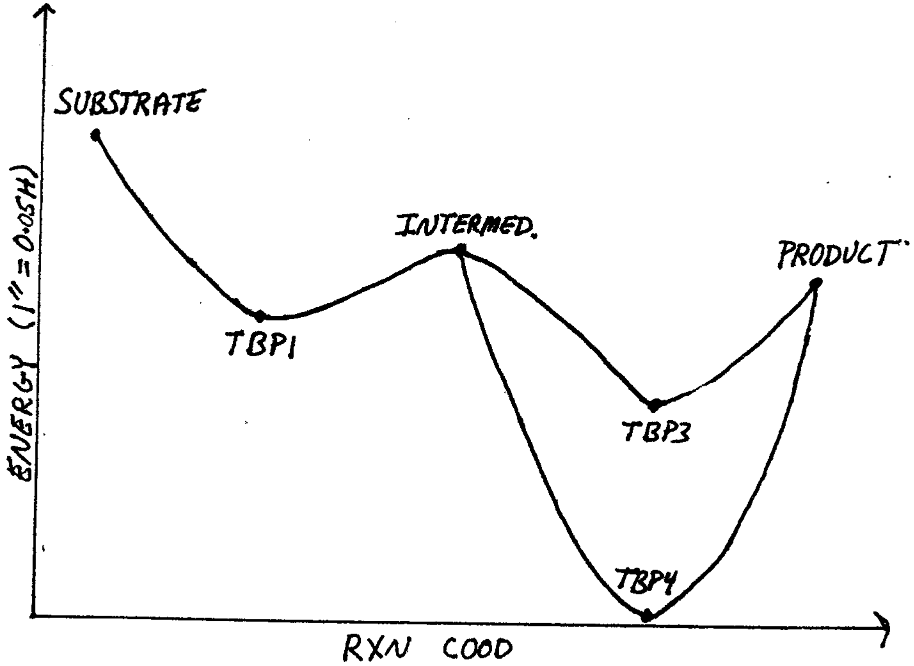


FIG. XIII. Energy of Activation Plot for Optimized Geometries Using GAUSS/70 Energies.

FIG. XIII shows negative activation energies (a physical impossibility) for each step of the reaction. This anomolous result may be attributed to one (or more) of the following:

1. since ΔE's were calculated and ΔE is approximately equal to ΔH, the neglected ΔS term would compensate and the resulting ΔG values would yield positive activation energies;

2. TBP1, TBP2, TBP3 and TBP4 are saddle point geometries, with some of these geometries lying in a large potential energy well in comparison with neighboring geometries;

3. CNDO/2 is not an accurate geometry predictor for phosphate ester molecules.

The first possibility may be ruled out, as the large energy difference between the actual result and what would be expected in a normal activation plot could not be compensated for by the small ΔS for this reaction.

The second possibility may also be discarded, as these saddle point geometries are obviously more strained and unstable than the substrate, intermediate and product geometries and would, therefore, be expected to have higher energies.

The third possibility, that CNDO/2 is not an accurate geometry predictor for the molecules tested, is, unfortunately, the likely cause of the results.

Mark Fraser (unpublished thesis, Union College, 1979), had used CNDO/2 to do calculations on small cyclic phosphate ester molecules. Afte comparing his results with experimental values, he concluded that, although CNDO/2 predicted conformations of acyclic systems with moderate success, it failed for cyclic phosphates. For example, the O-P-O

UN82 ARATO, M. A STUDY OF THE MECHANISM OF, ETC.
A663s/1980 CHEMISTRY HRS. 5/80 SHT. 2 OF 2

END

bond angles predicted for MEP were 10-15 degrees in error of experiment.

It was hoped that these errors would cancel (as we were only interested in relative energy differences of similar molecules) and that final energy calculations done with the GAUSS/70 method would improve results. These expectations were, however, not met.

CHAPTER V

CONCLUSION

Our physically impossible result of negative activation energies for each transition state prevented any prediction of reaction pathway. CNDO/2 must be disregarded for cyclic phosphate ester work. For the future, the program may be applied to more suitable molecules, such as amino acids, which contain only N, C, O and H atoms. A different quantum mechanical program, such as GAUSS/76 (an improved version of GAUSS/70), may also be coupled with the CINMIN optimization routine for work on cyclic phosphate esters.