Molecular Orbital Investigation of Multiply Hydrogen Bonded Systems. Formic Acid Dimer and DNA Base Pairs

Steve Scheiner[1a] and C. W. Kern*[1b]

Contribution from the Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, Illinois 62901, and the Department of Chemistry, The Ohio State University, Columbus, Ohio 43210. *Received June 26, 1978*

Abstract: Proton transfers within the formic acid dimer are studied using various methods. Minimum atomic basis sets yield single- and double-well potentials for single and double proton transfers, respectively, in agreement with extended basis set calculations. A potential energy surface for the decoupled proton transfers using the 4-31G basis set indicates that a simultaneous and synchronous motion of two protons is an adequate first approximation to the optimal mode of double transfer. All-electron PRDDO calculations are carried out for proton transfers within the base pairs guanine–cytosine and adenine–thymine. Single transfers are again found to be characterized by single-well potentials and double transfers by double-well potentials. Quantum mechanical tunneling rates and equilibrium constants are calculated for the most facile mode which involves a double proton transfer in the guanine–cytosine pair. The values obtained are compared to experimental data and inferences are drawn concerning the importance of the tunneling phenomenon in mutagenesis.

Introduction

The prevalence of hydrogen bonding in chemical and biological systems has led to extensive research into the nature of this interaction by various techniques including molecular orbital methods. In the majority of these studies, the interacting species have been connected by a single hydrogen bond.[2] There are, however, a number of biologically important phenomena in which multiple hydrogen bonding plays an integral role; e.g., the base-pair interactions in DNA which are essential to the integrity of the genetic code.[3] Such multiple bonding has been the subject of a number of previous quantum mechanical investigations.[4-17]

Löwdin[18] suggested that hydrogen-bonded protons can be transferred from one base of DNA to its complement and may play a key role in mutagenesis. These transfers, which could take place either via quantum mechanical "tunneling" or via more classical means, were proposed to lead to a pair of "rare" tautomeric bases that might be "read" incorrectly during the replication process and result in an altered form of the original DNA molecule.

Early molecular orbital studies of DNA base pairs[4-8] assumed separability of the σ- and π-electron systems. In this approximation, single proton transfers between the bases were found to be characterized by double-well potentials[7] as was assumed by Löwdin. More recently, however, an ab initio all-electron calculation of the guanine–cytosine pair[9] yielded a single-well potential for the transfer of each proton taken one at a time. In this calculation, 334 Gaussian-type functions were contracted to a minimum basis set of 105 atomic orbitals and required a total of 8 days of computer time. In an effort to examine multiple simultaneous transfers, a substantially larger basis set was applied to the formic acid dimer. Clementi et al.[9] found that a single transfer, which leads to a pair of charged species, is characterized by a monotonically increasing energy function. The double transfer was examined by synchronously coupling the motions of the two protons along $r_1 \equiv r_2$ (see Figure 1). This preserves electroneutrality in both species and results in a double-well potential.[9-15] Clementi et al.[9] thus concluded that double proton transfers are more likely to produce double-well potentials than are single transfers.

Since there is no a priori reason to believe that such a simultaneous transfer is the optimum mode, we have relaxed this restriction in the $(HCOOH)_2$ study so that each proton moves independently of the other. The results are used as a basis for the investigation of transfers between the base pairs guanine–cytosine (G–C) and adenine–thymine (A–T).

Formic Acid Dimer

The geometry used for our study was taken from a full STO-3G optimization of the formic acid dimer bond lengths and angles by Del Bene and Kochenour[11] except that we set the hydrogen bond length equal to the experimental value of 2.73 Å.[19] Our calculations were carried out with the GAUSSIAN 70 program[20] using a split-valence shell 4-31G basis set which, unlike minimum basis sets, reproduces accurately observed proton dissociation energies of various molecules including formic acid.[21] This basis set is thus particularly well suited to the proton transfers considered here.

The total energy calculated for the dimer (with $r_1 = r_2 = 1.01$ Å) was −376.970 66 au, corresponding to a stabilization energy between the two monomers of 20.2 kcal/mol. Clementi et al.[9] have noted a near degeneracy of the highest occupied σ and π molecular orbitals of the dimer. We find a separation of 0.019 au (11.7 kcal/mol) between these two levels, the σ orbital being the higher in energy. The first ionization potential is hence predicted to be of σ symmetry and the excitation requiring the least energy to be of $\sigma \rightarrow \pi^*$ type.

The protons, H_1 and H_2, were allowed to move independently along the lines joining O_1 and O_3 and joining O_2 and O_4, respectively. All other nuclei were held stationary. The potential energy surface is shown in Figure 2 as a function of the two OH distances r_1 and r_2 (Figure 1). The symmetry of the surface about the diagonal is a consequence of the C_{2h} point group for the dimer when $r_1 = r_2$. A single proton transfer is represented by the path A → B. The diagonal path A → C corresponds to a double transfer in which the motions of the two protons are coupled as was assumed in ref 9-15. The highest energy configuration of the entire surface, D, which is 45.9 kcal/mol higher in energy than A, occurs on the diagonal path at $r_1 = r_2 = 1.41$ Å. It was thought that a large amount of this instability might be a result of repulsion between protons H_1 and H_2. However, no stabilization of D occurred when these protons were allowed to move freely in the planes perpendicular to their transfer paths, i.e., in the yz plane of Figure 1.

The transfer path of lowest energy between points A and C would tend to avoid structure D. A possible candidate for such a path, denoted by the dashed line in Figure 2, represents a mode in which the transfer of one proton, H_1, is further advanced than is H_2 ($r_1 > r_2$); H_2 "catches up" to H_1 only at the conclusion of the transfer. The circumvention of D results, however, in only a small savings in activation energy, lowering the energy barrier for double transfer from 45.9 to about 44

0002-7863/79/1501-4081$01.00/0

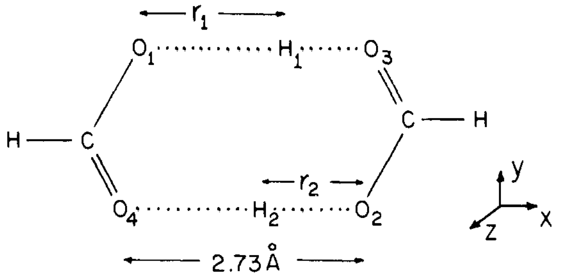


Figure 1. Proton transfers within the formic acid dimer. $r_1 = r_2 = 1.01$ Å for the dimer with zero proton transfer.

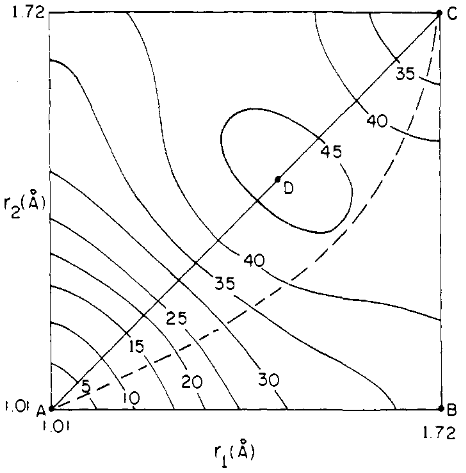


Figure 2. Potential energy (kcal/mol) surface for proton transfers in formic acid dimer. The zero of energy is taken to be that calculated for $r_1 = r_2 = 1.01$ Å (point A).

kcal/mol. For purposes of characterizing the potential for double proton transfers, path ADC would therefore appear to be a good first approximation to the optimum path.

The energetics of the single proton transfer (A → B), as calculated using several basis sets, are depicted in Figure 3. In addition to the 4-31G results, we show also those calculated using the STO-3G minimum basis set, all within the context of the GAUSSIAN 70 program. The PRDDO curve was obtained with a minimum basis set of Slater atomic orbitals.[22] The other results are taken from the paper by Clementi et al.[9] in which a slightly different internal geometry was used for formic acid. The notation (7/3) refers to a minimal basis set of contracted Gaussian functions[23] and (9/5) to a more expanded one.[24] Several points are shown also which were computed using a (9/5) basis set plus polarization functions on all atoms.[9] The potentials obtained using the extended basis sets 4-31G and (9/5) are nearly identical as are the minimal basis set PRDDO and STO-3G results. All basis sets, including (9/5) plus polarization functions, yield a single-well potential for the single proton transfer.

Results for the synchronous transfer of both protons H_1 and H_2 (path ADC) are shown in Figure 4. We note once again the close agreement between the pair of extended basis sets 4-31G and (9/5) as well as that between the minimum basis set PRDDO and STO-3G potentials. All four of these sets, as well as the "(9/5) plus polarization" set, yield double-well potentials for the coupled proton transfers with minima located at $r =$ 1.0 and 1.7 Å. The (7/3) set, on the other hand, predicts a single-well potential for this transfer mode. Although a shoulder appears in the (7/3) curve at about $r_1 = r_2 = 1.4$ Å, a similar shoulder is present in the (7/3) potential for the single transfer (Figure 3) for which all other basis sets produce a single-well potential. We note further that, for both single and double proton transfers, the minimum in the (7/3) potential occurs at $r_1 = 1.1$ Å, in contrast to the value of 1.0 Å found by all other basis sets. We conclude that the STO-3G and PRDDO methods yield synchronous transfer energy paths more representative of those obtained by use of extended basis sets than does the (7/3) set.

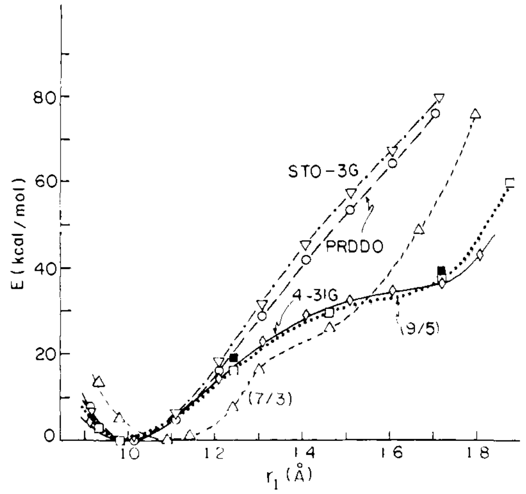


Figure 3. Energetics of single proton transfer in formic acid dimer $r_2 =$ 1.01 Å. Closed squares (■) represent results of calculations using "(9/5) + polarization" basis set.

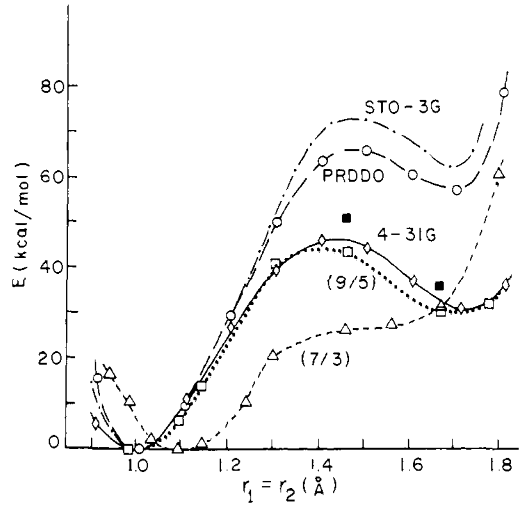


Figure 4. Calculated results for simultaneous double proton transfer in formic acid dimer. STO-3G curve was taken from ref 11.

The modes of proton transfer investigated thus far have allowed no motion of the heavy nuclei which are assumed stationary. Since the much bulkier atomic framework of the dimer is considerably less mobile than the protons being transferred, this is expected to provide a reasonably good basis for studying quantum mechanical tunneling. It is of interest, however, to examine also these transfer modes in the adiabatic approximation wherein all nuclei relax to their lowest energy positions at each stage of transfer. Previous studies using both the STO-3G basis set[11] and Gaussian lobe functions[10] found that C–O bond distance relaxation during the synchronous proton transfer in formic acid dimer resulted in a reduction of the energy barrier by 23%. Using STO-3G geometry optimizations, we find by 4-31G that structure D is 23.3 kcal/mol higher in energy than A (and C). Optimization therefore reduces the energy barrier by about 50% but preserves the double-minimum nature of the well.

It was expected that structure B might also be stabilized by relaxation of the rigidity constraint. A *full* STO-3G optimi-

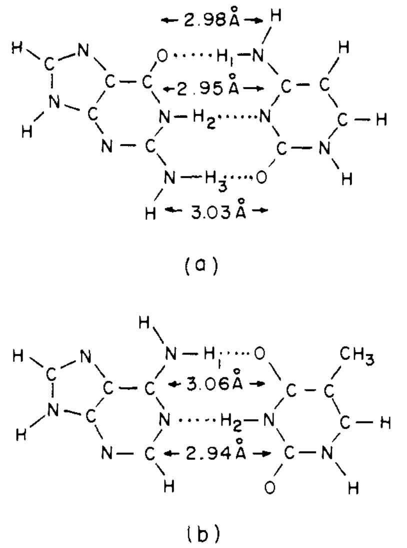


Figure 5. (a) G-C and (b) A-T base pairs. Hydrogen bonds are shown as dotted lines.

Table I. PRDDO Energies of DNA Bases

base	E au	base	E au
G	−537.9531	A	−463.3965
C	−391.5234	T	−450.2329
G-C	−929.5432	A-T	−913.6738

zation of the geometry of B was therefore carried out (including optimization of the intermolecular hydrogen bond distance). The 4-31G energy of the resulting configuration was, however, only 2 kcal/mol more stable than the unoptimized geometry of B. Allowing the nuclei to relax is thus seen to reduce significantly the energy barrier for double proton transfers, whereas it has little effect on the energetics of a single transfer.

DNA Base Pairs

Calculations were performed on the DNA base pairs guanine–cytosine (G–C) and adenine–thymine (A–T) using the PRDDO method.[22] The geometries of the pairs, depicted in Figure 5, were taken from Spencer's X-ray diffraction data[25] of the structure of DNA.[26] The PRDDO energies[28] of the base pairs as well as the isolated bases are presented in Table I.

Previous ab initio calculations[27–31] have found the highest occupied and lowest unoccupied molecular orbitals (MOs) of each base to be of π symmetry. Our PRDDO calculations agree with this result; in addition, we find that the highest occupied and lowest vacant MOs of both the G–C and A–T base pairs are of π symmetry as well.[32] Clementi et al.[9] have reached a similar conclusion for the highest occupied MO of the G–C pair. First ionization potentials are thus predicted to be of π symmetry and excitations of lowest energy are of $\pi \rightarrow \pi^*$ type for all bases and pairs studied here.

Hydrogen bonds in the base pairs are of two types, NH···N and NH···O. The position of a fully transferred proton in either bond type was defined such that $r(XH) = 1.0$ Å, X = O, N. A proton transferred to N bisected the associated CNC angle whereas $\angle COH = 110°$ for a proton transferred to O. Since the hydrogen bond lengths, $r(NX)$, are different each proton required a different total displacement (disp) in order to complete its transfer. These quantities are listed in Table II. We have defined the fractional proton transfer, δ, as the fraction of the total displacement moved by a proton at any point along its straight line path. For example, $\delta = 0$ for those positions depicted in Figure 5 and $\delta = 1$ for a fully transferred proton. All nuclei other than the protons being transferred are held stationary.

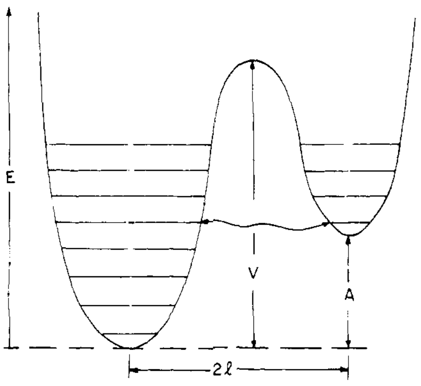


Figure 6. Model potential used for tunneling rate calculations. The barrier is formulated using an Eckart function and the wells are of approximately quadratic form. Wavy arrow denotes proton tunneling.

Table II. Hydrogen Bond Lengths, Total Displacements, and Energies for Single Proton Transfer

	$r(NX)$, Å[a]	disp, Å	ΔE, kcal/mol
		G-C	
H_1	2.98	0.99	55.5
H_2	2.95	0.95	54.4
H_3	3.03	1.06	79.4
		A-T	
H_1	3.06	1.12	81.3
H_2	2.94	0.94	67.6

[a] Taken from ref 25.

The PRDDO energy[28] is a monotonically increasing function of δ for each single transfer (a single-well potential), in agreement with the calculations of Clementi et al.[9] The total energy required for each transfer is presented in Table II as ΔE. A correlation between ΔE, the total displacement, and the hydrogen bond length for each transfer may be noted. The potentials obtained for multiple proton transfers in the base pairs are also given in ref 28. Multiple transfers were carried out in a coupled manner such that values of δ for each proton being transferred were equal at every point in the transfer. The double transfers are characterized by double-well potentials with barrier height V and upper well height A (see Figure 6). The PRDDO values of V and A are given in Table III.

Our results on formic acid dimer suggest that a single- or double-well potential obtained in the PRDDO approximation is likely to be confirmed by extended basis set calculations. The steepness of double-well potentials, their barrier heights, and the energy difference between the two wells are overestimated by PRDDO but the qualitative features nonetheless remain. It is therefore our expectation that, when extended basis set potentials become computationally practical for base pairs, single proton transfers will lead to potentials of single-well character but less steep than those shown in Figures 2 and 3 of ref 28 by a factor of perhaps ½. Double proton transfers are likely to be characterized by double-well potentials but with energy barriers considerably smaller than those reported here. We also anticipate increased stability of the right-hand wells relative to the left (see Figure 4). Similar effects are expected if the heavy atoms of the bases are allowed to relax their positions as the proton transfers occur. Our calculations on the formic acid dimer suggest that the barriers to double proton transfers may also be facilitated by uncoupling the motions of the two protons or by zero-point vibrational deformations.[13] All of these effects tend to *reduce* the PRDDO energy barrier.

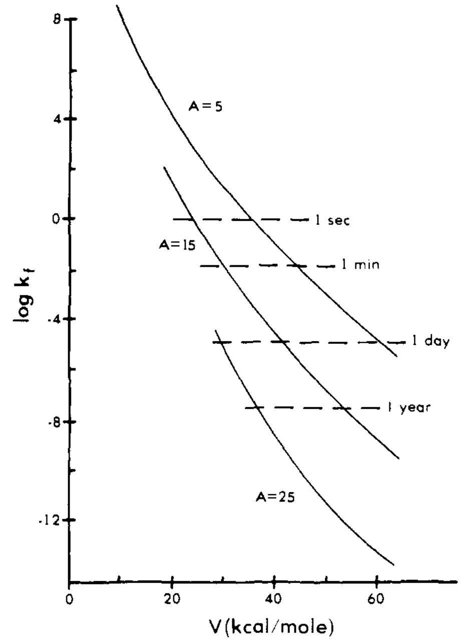


Figure 7. Functional dependence of the forward tunneling rate constant on barrier height V for various values of well height difference A (kcal/mol). Also shown are approximate times required for attainment of equilibrium.

Table III. Energetics of Double Proton Transfers

	V, kcal/mol	A, kcal/mol
	G-C[a]	
H_1, H_2	61.6	26.3
H_1, H_3	94.3	76.8
	A-T	
H_1, H_2	82.8	57.2

[a] The transfer of (H_2, H_3) was not investigated since this would lead to a very unstable pair of doubly charged bases.

The proton transfer potentials in hydrogen-bonded systems have been shown to be sensitive to the distance between the two monomers.[11] We therefore decided to investigate the effect on the potentials of varying the interbase distance. In particular, we were interested in the double transfer of protons H_1 and H_2 of the G-C pair since this mode is computed to be the most facile. It was found that, when the hydrogen bond distance, r_{NN}, is decreased from 2.95 to 2.75 Å (see Figure 5), the height of the barrier to (H_1H_2) transfer is reduced from 62 to 28 kcal/mol. In addition, A, the difference in energy between the left- and right-hand wells, is decreased much less, from 26 to 23 kcal/mol. When r_{NN} in further reduced to 2.55 Å, the double proton transfer is characterized by a single-well potential with a minimum located at $\delta \sim 1/4$. Our calculations therefore indicate that a double-well potential is characteristic of the (H_1H_2) double proton transfer for r_{NN} greater than about 2.6-2.7 Å.

Rein and Harris[7] have estimated the equilibrium proportions of the tautomers formed as a result of the transfer of protons H_1 and/or H_2 in the G-C pair. They have also computed the times required for each of these protons to tunnel to the complementary base. Their results were based on semiempirical molecular orbital calculations on (primarily) the π-electron system of the base pair which yielded double-well potentials for single proton transfers, contrary to the all-electron calculations reported here and by Clementi et al.[9] Since our *double* transfers do, by contrast, yield double-well potentials, we have computed values for the equilibrium and rate constants associated with these modes. We treat the quantum mechanical tunneling of a single quasi-particle (of mass equal to that of two protons) from the left well to the right well (see Figure 6). Löwdin[18] has estimated that the probability per second for tunneling from one side of the barrier to the other is the product of the transmission coefficient (g) and the classical oscillatory frequency (ν) of the particle in the original well. We obtain the frequency ν by assuming a quadratic form of the potential near the bottom of each well. Thus

$$\nu = \frac{1}{2\pi}\sqrt{\alpha/\mu}$$

where μ is equal to the mass of our quasi-particle and the force constant α is obtained from our computed potential.

The next step involves the identification of our computed potential with a model function containing an Eckart barrier for which the transmission coefficient may be calculated exactly.[33] Each model function is characterized by a barrier height V, barrier width $2l$, and right well height A as shown in Figure 6. For appreciable values of A ($\gtrsim 3$ kcal/mol) tunneling to the right well occurs only from vibrationally excited states of the left well for which $E > A$. It is therefore necessary to include in our calculations the probability of finding the particle initially in a suitable excited state of the left well. The rate of tunneling, k, is hence expressed as

$$k = \nu g b$$

where b is the appropriate Boltzmann constant.[7c]

We focus our attention on the most facile mode of proton transfer: the double transfer of H_1 and H_2 in the G-C pair. This mode leads to the "rare" tautomeric forms of purines and pyrimidines which have been observed in aqueous solution.[34] Using the calculated (PRDDO) values of V and A equal to 62 and 26 kcal/mol, respectively, we compute a forward (left → right) rate constant, k_f, of 10^{-14} s^{-1} at 37 °C. (The reverse (left ← right) rate $k_r \cong 10^4$ s^{-1} is many orders of magnitude faster.) This extremely small value of k_f implies that several million years are required for the system to attain equilibrium. In addition, the ratio K of rare to normal tautomeric forms of the base pair present at equilibrium is only 6×10^{-19}.

The potential giving rise to these values was computed using a minimum basis set of atomic orbitals. As shown above for the formic acid dimer, use of an extended basis set might significantly reduce both the energy barrier V and the right well height A. Guided by the data shown in Figure 4 we estimate that V and A would be about equal to 45 and 15 kcal/mol, respectively, if an extended basis (e.g., 4-31G) were used. These parameters lead to values of $k_f \sim 10^{-6}$ s^{-1} and $K \sim 10^{-10}$. The equilibrium ratio, within the lower limit of the estimated "spontaneous" frequency of changes per base pair per generation,[35] is nevertheless several orders of magnitude smaller than the relative proportions of rare to normal tautomers of purines and pyrimidines normally present in aqueous solution.[34]

The extreme sensitivity of the calculated tunneling rate constant to the parameters V and A is illustrated in Figure 7. The equilibrium constant K is also very sensitive to the transfer energetics, obeying approximately the relation $K = \exp(-A/RT)$. The wide range of tunneling data obtained with different transfer potentials indicates that very accurate quantum mechanical calculations are required in order to obtain reliable kinetic and equilibrium data from first principles.

The potentials derived in this paper were calculated for the base pairs in vacuo. We have thus neglected the effect of the chemical environment of the bases. Moreover, we have assumed rigid structures for the bases during the proton transfers. Relaxation of these restrictions, although computationally expensive, might produce important changes in our calculated potentials and hence the contribution of proton tunneling to mutation rates.

Acknowledgments. We thank Professor I. Shavitt for bringing the Eckart function to our attention and Professor T. A. Halgren for providing us a copy of the PRDDO program. We acknowledge generous grants of computer time from the Instruction and Research Computer Center of The Ohio State University and the Southern Illinois University Academic Computing Center. S.S. was a Weizmann Postdoctoral Fellow during the course of much of this work. Dr. S. Abdulnur kindly brought ref 16 to our attention.

References and Notes

(1) (a) Southern Illinois University; (b) The Ohio State University.
(2) P. A. Kollman in "Applications of Electronic Structure Theory", H. F. Schaefer, Ed., Plenum Press, New York, 1977, pp 109–152, and references cited therein.
(3) J. D. Watson and F. H. C. Crick, *Nature (London)*, **171**, 737, 964 (1953).
(4) B. Pullman and A. Pullman, *Biochim. Biophys. Acta*, **36**, 343 (1959).
(5) H. Berthod and A. Pullman, *Biopolymers*, **2**, 483 (1964); *J. Chim. Phys. Phys.-Chim. Biol.*, **62**, 942 (1965).
(6) R. Rein and J. Ladik, *J. Chem. Phys.*, **40**, 2446 (1964).
(7) (a) R. Rein and F. E. Harris, *J. Chem. Phys.*, **41**, 3393 (1964); (b) *ibid.*, **42**, 2177 (1965); (c) *ibid.*, **43**, 4415 (1965); (d) *ibid.*, **45**, 1797 (1966); (e) *Science*, **146**, 649 (1964).
(8) S. Lunell and G. Sperber, *J. Chem. Phys.*, **46**, 2119 (1967).
(9) E. Clementi, J. Mehl, and W. von Niessen, *J. Chem. Phys.*, **54**, 508 (1971).
(10) E. Ady and J. Brickman, *Chem. Phys. Lett.*, **11**, 302 (1971).
(11) J. E. Del Bene and W. L. Kochenour, *J. Am. Chem. Soc.*, **98**, 2041 (1976).
(12) K. Morokuma, S. Iwata, and J. A. Lathan in "The World of Quantum Chemistry", R. Daudel and B. Pullman, Eds., Reidel, Holland, 1974, pp 277–316.
(13) W. A. Sokalski, H. Romanowski, and A. Jaworski, *Adv. Mol. Relaxation Interaction Processes*, **11**, 29 (1977).
(14) H. Morita and S. Nagakura, *Theor. Chim. Acta*, **27**, 325 (1972).
(15) P. Schuster, *Int. J. Quantum Chem.*, **3**, 851 (1969).
(16) J. Marañon, O. M. Sorarrain, H. Grinberg, S. Lamdan, and C. H. Gaozza, *J. Theor. Biol.*, **74**, 11 (1978). See also, S. Abdulnur, "Quantum Science, Methods, and Structure", Plenum Press, New York, 1976, pp 567–575.
(17) A. C. Blizzard and D. P. Santry, *J. Theor. Biol.*, **25**, 461 (1969).
(18) P. O. Löwdin, *Adv. Quantum Chem.*, **2**, 213 (1965).
(19) G. Karle and L. O. Brockway, *J. Am. Chem. Soc.*, **66**, 574 (1944).
(20) R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724 (1971); W. J. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton, and J. A. Pople, GAUSSIAN 70, Program No. 236, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Ind., 1974.
(21) S. Scheiner, D. A. Kleier, and W. N. Lipscomb, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 2606 (1975).
(22) T. A. Halgren and W. N. Lipscomb, *J. Chem. Phys.*, **58**, 1569 (1973). Standard orbital exponents were used for all atoms except H, for which 1.2 was used.
(23) E. Clementi, *J. Chem. Phys.*, **46**, 4737 (1967).
(24) S. Huzinaga, *J. Chem. Phys.*, **42**, 1293 (1962).
(25) M. Spencer, *Acta Crystallogr.*, **12**, 59, 66 (1959).
(26) The atomic coordinates of the G–C base pair were taken from ref 27. The coordinates given for thymine in ref 27 do not agree with the internal geometry of this base as provided in ref 25. The coordinates for the A–T pair used here were therefore calculated directly from Spencer's data.
(27) E. Clementi, J. M. André, M. Cl. André, D. Klint, and D. Hahn, *Acta Phys Acad. Sci. Hung.*, **27**, 493 (1969).
(28) A preliminary account has appeared: S. Scheiner and C. W. Kern, *Chem. Phys. Lett.*, **57**, 331 (1978).
(29) R. E. Christoffersen and R. P. Angeli in "New World of Quantum Chemistry", B. Pullman and R. Parr, Eds., Reidel, Boston, 1976, pp 189–210.
(30) B. Mely and A. Pullman, *Theor. Chim. Acta*, **13**, 278 (1969).
(31) L. C. Snyder, R. G. Shulman, and D. B. Neumann, *J. Chem. Phys.*, **53**, 256 (1970).
(32) Because the hydrogens of the methyl group of thymine lie out of the plane of the remainder of the molecule, the MOs of T (and A–T) cannot be strictly identified as pure σ or π MOs. However, such a classification is possible for most of the relevant MOs as they are of predominantly σ or π type.
(33) C. Eckart, *Phys. Rev.*, **35**, 1303 (1930).
(34) D. M. Brown and M. J. E. Hewlins, *J. Chem. Soc. C*, 2050 (1968); R. V. Wolfenden, *J. Mol. Biol.*, **40**, 307 (1969).
(35) E. Freese, *J. Theor. Biol.*, **3**, 82 (1962); J. W. Drake, "The Molecular Basis of Mutation", Holden-Day, San Francisco, 1970, pp 59–62.

Effect of Electron Correlation on Theoretical Equilibrium Geometries

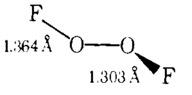
Douglas J. DeFrees,[1a] **Beverly A. Levi,**[1a] **Steven K. Pollack,**[1a] **Warren J. Hehre,***[1a] **J. Stephen Binkley,**[1b] **and John A. Pople**[1b]

Contribution from the Department of Chemistry, University of California, Irvine, California 92717, and the Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 *Received January 24, 1979*

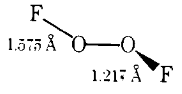
Abstract: Theoretical HF/6-31G* (Hartree-Fock, 6-31G* basis set) and MP2/6-31G* (second-order Møller-Plesset, 6-31G* basis set) equilibrium geometries have been obtained for a set of one- and two-heavy-atom molecules containing only first-row elements and for which experimental structural data are available. Geometrical parameters from the MP2/6-31G* calculations, which treat effects of electron correlation explicitly, are generally in better agreement with the experimental data than are those derived from the corresponding single-determinant wave functions. This is particularly true for bonds involving highly electronegative atoms (e.g., O and F), which generally are too short at the Hartree-Fock limit. Here the correlation corrections result in significant bond lengthening; the MP2/6-31G* values for single bond lengths are in good accord with experimental values. However, MP2/6-31G* lengths of multiple linkages are generally overestimated. The 6-31G* polarization basis set appears to be among the simplest representations suitable for use in molecular orbital calculations which take partial account of electron correlation.

Introduction

Hartree–Fock theory has proven to be remarkably successful in the calculation of the equilibrium structures of small molecules.[2a] Basis sets as simple as STO-3G generally yield bond lengths to within a few hundredths of an ångstrom of the experimental values, and bond angles to within a few degrees. The level of agreement of theoretical equilibrium geometries with experiment increases with the basis set size. Nevertheless, with very large basis sets the theoretical geometries for some very simple molecules are seriously in error. For example, the best single determinant calculations yet performed[2b] on the fluorine molecule yield a bond length 0.1 Å shorter than the experimental value. Worse yet, the near Hartree–Fock equilibrium bond lengths for the fluorine hydroperoxide molecule are very different from the experimental values. Partial consideration



Hartree–Fock



experimental

0002-7863/79/1501-4085$01.00/0