perexchange type of electronic coupling) since the large value of $V$ itself enables a long-range electron transfer to occur in this system.

## Conclusion

The experimental data show the marked effect that temperature imposes on the photoinduced electron transfer from excited tris(2,2'-bipyridine)ruthenium(II) to methylviologen. The unusual kinetics are quantitatively explained by a mechanism whereby the nuclear contribution to the electron transfer probability is dominated by coupling with high-frequency molecular modes. Similar experimental data have been reported earlier for electron transfer in bio systems, while the present data provide the opportunity to check the theoretical consideration in a quantitative fashion.

**Acknowledgment.** The authors thank the National Science Foundation for support of this work via Grant NO. CHE 85-02062.

**Registry No.** $MV^{2+}$, 4685-14-7; $Ru(bpy)_3^{2+}$, 15158-62-0.

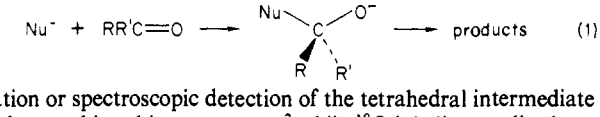
# Ab Initio and Monte Carlo Calculations for a Nucleophilic Addition Reaction in the Gas Phase and in Aqueous Solution

**Jeffry D. Madura and William L. Jorgensen***

*Contribution from the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907.* *Received November 18, 1985*

**Abstract:** Energy profiles for the nucleophilic addition of hydroxide ion to formaldehyde in the gas phase and in aqueous solution have been determined with quantum and statistical mechanical methods. Ab initio calculations at the 6-31+G* level were utilized to study the gas-phase potential energy surface. In the gas phase, the conversion of reactants to the tetrahedral intermediate is exothermic by 35 kcal/mol and proceeds via an ion-dipole complex. Monte Carlo simulations were then carried out for the reacting system solvated by 269 water molecules at 25 °C and 1 atm with solute-solvent potential functions derived from ab initio calculations in conjunction with the TIP4P model of water. Importance sampling methods were employed to obtain the potential of mean force in solution. A substantial activation barrier is introduced by hydration with the transition state occurring at a C–O separation of ca. 2 Å. The calculated height of the free energy barrier is 24–28 kcal/mol and the tetrahedral intermediate is 10–14 kcal/mol above the reactants. Finally, fixed solute simulations corresponding to the reactants, transition state, and product were carried out to probe the origin of the activation barrier. This established that it is not so much a change in the number of solute–water hydrogen bonds along the reaction path but rather in their strengths that is primarily responsible for the solvent-induced barrier. The charge-localized hydroxide ion participates in six strong hydrogen bonds with water molecules. Upon proceeding to the more charge-delocalized transition state and product, there are still 6–7 water–solute hydrogen bonds; however, they are ca. 40% weaker than for hydroxide ion.

Nucleophilic addition to a carbonyl group is a fundamental process in organic chemistry and biochemistry. Common examples include ester and amide hydrolysis, aldol condensation, and numerous reactions of organometallic reagents and ylides.[1,2] Extensive experimental work has led to general acceptance of the "tetrahedral mechanism" for these processes;[2,3] the key feature is formation of a tetrahedral intermediate that collapses to products (eq 1).



(1)

Isolation or spectroscopic detection of the tetrahedral intermediate have been achieved in some cases,[2] while $^{18}O$ labeling studies have provided additional compelling evidence for its existence.[4] Furthermore, Guthrie's analyses of thermodynamic and kinetic data have provided estimates of the free energy profiles for ester and amide hydrolyses.[5] The formation of the tetrahedral intermediate is normally the rate-determining step.

Still greater knowledge of the mechanism and structure of the intermediate and transition states are important goals. Valuable information has been obtained from recent gas-phase experiments by several groups.[6] A striking observation is the general tendency of esters to undergo competitive proton transfers, eliminations, and $S_N2$ reactions in favor of $B_{AC}2$ processes (eq 1) in the absence of solvent. However, reversal of this pattern has been found upon addition of only one solvent molecule.[6d] Asubiojo and Brauman have also found the $B_{AC}2$ mechanism to be operative for gas-phase additions to acyl halides.[6e] Their analysis of the kinetics led to the provocative proposal that the reactions in the gas phase feature a double-well potential surface with the tetrahedral species as a transition state and ion–dipole complexes as minima.

Additional insight on the reaction surface has been provided by Burgi et al.[7a,b] Using crystallographic data for various nucleophilic groups interacting with different carbonyl centers, they were able to suggest likely pathways for approach of the nu-

(1) For reviews see: (a) Bender, M. L. *Chem. Rev* **1960**, *60*, 53. (b) Johnson, S. L. *Adv. Phys. Org. Chem.* **1957**, *5*, 237. (c) Bruice, T. C.; Benkovic, S. J. *Bioinorganic Mechanisms*; Benjamin: New York, 1966; Vol. I, Chapter 1. (d) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, 1969. (e) Blow, D. *Acc. Chem. Res.* **1976**, *9*, 145.

(2) March, J. *Advanced Organic Chemistry*; Wiley: New York, 1985; Chapters 10 and 16.

(3) (a) Bender, M. L. *J. Am. Chem. Soc.* **1951**, *73*, 1626. (b) Barnett, R. E. *Acc. Chem. Res.* **1973**, *6*, 41. (c) Jencks, W. P. *Acc. Chem. Res.* **1980**, *13*, 161.

(4) Bender, M. L.; Heck, H. d'A, *J. Am. Chem. Soc.* **1967**, *89*, 1211.

(5) (a) Guthrie, J. P. *J. Am. Chem. Soc.* **1973**, *95*, 6999. (b) Guthrie, J. P. *J. Am. Chem. Soc.* **1974**, *96*, 3608.

(6) (a) Faigle, J. F. G.; Isolani, P. C.; Riveros, J. M. *J. Am. Chem. Soc.* **1976**, *98*, 2049. (b) Comisarow, M. *Can. J. Chem.* **1977**, *55*, 171. (c) Takashima, K.; Riveros, J. M. *J. Am. Chem. Soc.* **1978**, *100*, 6128. (d) Fukuda, E. K.; McIver, R. T. *Ibid.* **1979**, *101*, 2498. (e) Asubiojo, O. I.; Brauman, J. I. *Ibid.* **1979**, *101*, 3715. (f) Bohme, D. K.; Mackay, G. I.; Tanner, S. D. *Ibid.* **1980**, *102*, 407. (g) Bartmess, J. E.; Hays, R. L.; Caldwell, G. *Ibid.* **1981**, *103*, 1338. (h) McDonald, R. N.; Chowdhury, A. K. *Ibid.* **1982**, *104*, 901. (i) DePuy, C. H.; Della, E. W.; Filley, J.; Grabowski, J. J.; Bierbaum, V. M. *Ibid.* **1983**, *105*, 2481. (j) Johlman, C. L.; Wilkins, C. L. *Ibid.* **1985**, *107*, 327.

(7) (a) Burgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am Chem. Soc.* **1973**, *95*, 5065. (b) Burgi, H. B.; Dunitz, J. D. *Acc. Chem. Res.* **1983**, *16*, 153. (c) Burgi, H. B.; Lehn, J. M.; Wipff, G. *J. Am. Chem. Soc.* **1974**, *96*, 1956. (d) Burgi, H. B.; Dunitz, J. D.; Lehn, J. M.; Wipff, G. *Tetrahedron* **1974**, *30*, 1563. (e) Williams, I. H.; Maggiora, G. M.; Schowen, R. L. *J. Am. Chem. Soc.* **1980**, *102*, 7831. (f) Maggiora, G. M.; Williams, I. H. *THEOCHEM* **1982**, *88*, 23. (g) Scheiner, S.; Lipscomb, W. N.; Kleier, D. A. *J. Am. Chem. Soc.* **1976**, *98*, 4770. (h) Alagona, G.; Scrocco, E.; Tomasi, J. *J. Am. Chem. Soc.* **1975**, *97*, 6976. (i) Dewar, M. J. S.; Storch, D. M. *J. Chem. Soc., Chem. Commun.* **1985**, 94.

0002-7863/86/1508-2517$01.50/0 

cleophile. They then proceeded to perform SCF-LCGO-MO calculations on a simple model reaction, hydride addition to formaldehyde. The calculations yielded an energy surface with no barrier to the tetrahedral intermediate, which was 48 kcal/mol below the reactants.[7c,d] The approach vector indicated by the crystal data was reproduced in the calculations. Other quantum mechanical studies of addition reactions by several groups have produced similar reaction surfaces with no barrier to the tetrahedral intermediate.[7e-i] Thus, these theoretical findings seem to be in contradiction with the reaction profile proposed by Asubiojo and Brauman.[6e]

In order to provide further insight on the potential surfaces for addition reactions, results of an extensive theoretical study are reported here for the prototype reaction of hydroxide ion and formaldehyde both in the gas phase and in aqueous solution. The computational approach features ab initio quantum mechanical calculations to characterize the gas-phase reaction surface, followed by Monte Carlo statistical mechanics simulations to determine a free energy profile (potential of mean force) for the reaction in water. The same methodology has been applied previously with good success to the $S_N2$ reaction of chloride ion with methyl chloride in the gas phase, in aqueous solution, and in DMF.[8] Besides characterization of the reaction profiles, the calculations provide extensive details on structural and energetic changes along the reaction paths. As discussed below, these changes are particularly significant for the addition reaction since the substantial activation barrier in aqueous solution is essentially entirely solvent-induced.

The basic procedure in this approach involves three steps.[8] First, the lowest energy reaction path is determined for the gas phase by using ab initio calculations. Then, intermolecular potential functions are obtained to describe the interactions between the substrate and a solvent molecule. These vary along the reaction coordinate and are based on numerous ab initio calculations for complexes of the substrate with a solvent molecule. Finally, with analytical descriptions of the gas-phase reaction surface and the substrate–solvent and solvent–solvent potential functions, Monte Carlo simulations can be carried out to determine the free energy profile for the reaction in solution. This is a substantial undertaking since importance sampling methods are required to explore all values of the reaction coordinate. It is important to emphasize that the results of the statistical mechanics calculations correspond to the effects of solvating the gas-phase reaction. To detect a change in mechanism in solution would require the construction of multidimensional potentials of mean force. This is current beyond the range of most available computational resources.

Related computations by Weiner et al. on the addition of hydroxide ion to formamide should be noted.[9] Quantum mechanics was employed to obtain the reaction profile in the gas phase, and subsequently energy minimization was used for the substrate surrounded by water to estimate the solvent effect on the energetics.[9] The latter procedure does not take into account thermal, configurational averaging as in the present statistical mechanical calculations of the potential of mean force.

## Reaction Surface in the Gas Phase

**Calculations.** The gas-phase potential surface was determined with ab initio theory at 6-31+G* level.[10c] This basis set includes d orbitals and diffuse s and p functions on non-hydrogen atoms. The calculations were carried out with the GAUSSIAN 82[10] series of programs converted to run on the GOULD CONCEPT 32/8750 computer in our laboratory. Optimizations were performed by using the Berny algorithm with analytical gradients.[11] Frequency calculations used first and second analytical derivatives.[12]

Several alternative basis sets were evaluated before deciding on 6-31+G* (Table I). The last column in Table I gives the relative energy between the reactants and the tetrahedral complex. The interaction energy is found to be particularly sensitive to the inclusion of diffuse functions in the basis set. This is consistent with earlier work that has pointed out the general necessity of diffuse functions for computing reasonable energetic results for small, charge-localized anions such as $OH^-$.[10c] Gao et al. have

**Table I.** Basis Set, Zero-Point, and Correlation Dependence of $OH^-$ + $H_2C{=}O$ Energies[a]

| method, basis set[b] | $OH^-$ | $H_2C{=}O$ | complex | $-\Delta E$[c] |
|---|---|---|---|---|
| 3-21G | -74.86863 | -113.22182 | -188.20379 | 71.11 |
| 3-21+G | -74.99575 | -113.24467 | -188.30353 | 39.60 |
| ZPE/3-21+G//3-21+G | -74.98748 | -113.21597 | -188.25848 | 34.53 |
| 4-31++G//4-31G[d] | -75.29063 | -113.69924 | -189.04301 | 33.34 |
| 6-31G* | -75.32660 | -113.86633 | -189.27436 | 51.09 |
| 6-31+G* | -75.37642 | -113.87116 | -189.30360 | 35.15 |
| 6-311G* | -75.35328 | -113.89576 | -189.32754 | 49.25 |
| MP2/6-311G*//6-311G* | -75.55566 | -114.21801 | -189.86463 | 57.07 |
| MP3/MP2/6-311G*//6-311G* | -75.55711 | -114.22258 | -189.86834 | 55.62 |

[a] Total energies are in hartrees. [b] The notation x//y refers to a calculation with basis set x on geometries optimized with basis set y. MP2 and MP3 mean correlation energy corrections have been included with second- and third-order Moller–Plesset theory, and ZPE indicates inclusion of the zero-point vibrational energy. For more details, see ref 10. [c] Energy differences are in kcal/mol. [d] Unpublished results of Dr. R. Squires.

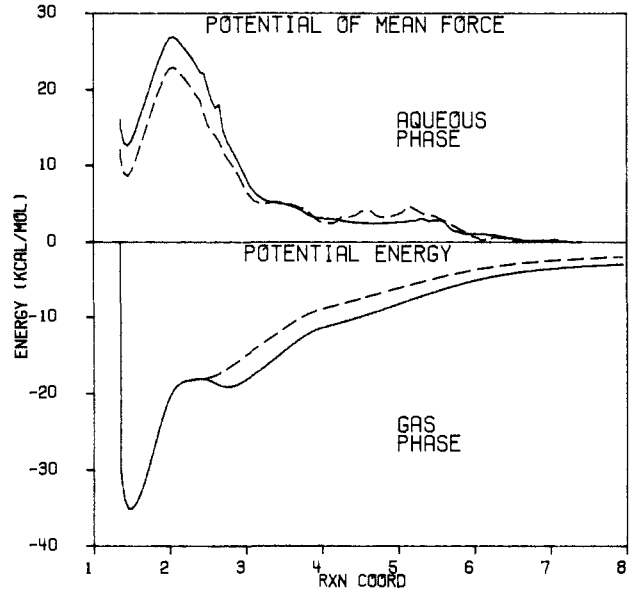



**Figure 1.** Energy profiles in the gas phase and in aqueous solution for the hydroxide + formaldehyde reaction. Solid lines represent the collinear trajectory, while the dashed lines are for the prependicular approach. The reaction coordinate is the distance in angstroms between hydroxyl oxygen and carbonyl carbon.

(8) (a) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 3049. (b) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154. (c) Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974.

(9) Weiner, S. J.; Singh, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1985**, *107*, 2219.

(10) (a) Binkley, J. S.; Whiteside, R. A.; Raghavachari, K.; Seeger, R.; DeFrees, D. J.; Schlegel, H. B.; Frisch, M. J.; Pople, J. A.; Kahn, L. R. *GAUSSIAN 82 Release H*; Carnegie-Mellon University: Pittsburgh, 1982. (b) 3-21G: Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939. Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797. (c) 3-21+G and 6-31+G*: Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294. (d) 6-31G*: Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213. (e) For a general discussion of basis sets and methodology, see: Clark, T. *A Handbook of Computational Chemistry*; Wiley: New York, 1985; Chapter 5.

(11) (a) Schlegel, H. B. *J. Chem. Phys.* **1982**, *77*, 3676. (b) Schlegel, H. B. *J. Comput. Chem.* **1982**, *3*, 214.

(12) (a) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; DeFrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A.; Hout, R. F.; Hehre, W. J. *Int. J. Quantum Chem.* **1980**, S15, 269. (b) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* **1979**, S13, 225. (c) Murrell, J. N.; Laidler, K. J *Trans. Faraday Soc.* **1968**, *64*, 371.

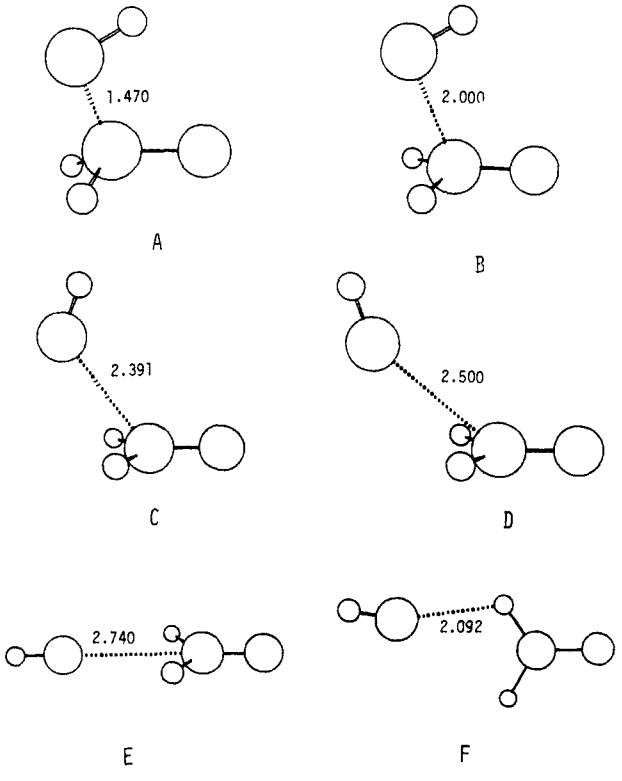



**Figure 2.** Selected optimized 6-31+G* structures along the reaction path.

also found that 6-31+G* is significantly better than 6-31G* for reproducing experimental anion–water interaction energies.[13] The effects of including zero-point and correlation energy corrections were also studied for the reaction. As shown in Table I, the corrections are both ca. 5–6 kcal/mol and since they have opposite sign, they nearly cancel. This result is affected by basis set choice and certainly may not be the case at all points on the reaction surface. Unfortunately, the magnitude of these calculations is such that we could not study the reaction surface including zero-point and correlation corrections with the 6-31+G* basis set. All things considered, our best estimate of the energy release on forming the tetrahedral complex is roughly the 6-31+G* value of 35 kcal/mol. In view of this and the above analysis, the 6-31+G* basis set was adopted for studying the entire surface for the gas phase. Unfortunately, no experimental guidance appears to be available for the complexation energy. However, the 6-31+G* results for anion–water complexes including $OH^-\cdots H_2O$ (vide infra) are very good.[13]

After the 6-31+G* basis set was decided upon, the optimal reaction path was then determined in $C_s$ symmetry (bottom half of Figure 1, solid line). The reaction coordinate for Figure 1 is the distance between the hydroxyl oxygen and the carbonyl carbon. According to this definition, the separated reactants are at the far right and the tetrahedral intermediate is the minimum on the left. At each value of the reaction coordinate complete optimization was carried out for all other geometrical variables within $C_s$ symmetry. Key points along the pathway are illustrated in Figure 2, structures A–E. The approach at large separation is collinear up to the apparent minimum (E) at $R_{CO}$ = 2.74 Å. The hydroxide then rises out of the plane to the transition state (C) at 2.39 Å. From there it is downhill to the tetrahedral complex (A) with $R_{CO}$ = 1.47 Å.

Frequency calculations provided further characterization of the stationary points. Due to the high demands on both computer time and disk space, these calculations could not be carried out above the 3-21+G level. However, the 3-21+G and 6-31+G* reaction profiles and geometries are very similar. The tetrahedral complex was shown to be a true minimum at the 3-21+G level. Study of the rotational potential about the C–OH bond further established the syn conformer (A) to be the global minimum for $HOCH_2O^-$. The transition state (C) was also verified as having only one imaginary frequency. However, the apparent minimum (E) turns out to be a saddle point between two equivalent structures of $C_1$ symmetry. The latter correspond to F in Figure 2, which may be called a hydrogen bonded structure. The energy difference between F and the ion–dipole geometry (E) turns out to be almost negligible with both basis sets, ca. 0.1 kcal/mol. Consequently, for computational convenience, the fully $C_s$ pathway has been employed in the Monte Carlo calculations, though it deviates very slightly from the minimum energy reaction path (MERP) for the gas phase.

(13) Gao, J.; Garner, D.; Jorgensen, W. L., submitted for publication.

**Table II.** 6-31+G* Energy as a Function of the Reaction Coordinate[a]

| $R_{CO}$ | energy, hartreess | rel energy, kcal/mol |
|---|---|---|
| 1.00 | −189.02398 | 140.28 |
| 1.30 | −189.28741 | −24.99 |
| 1.40 | −189.30151 | −33.84 |
| 1.47 | −189.30360 | −35.15 |
| 2.00 | −189.28009 | −20.40 |
| 2.28 | −189.27659 | −18.20 |
| 2.39 | −189.27635 | −18.05 |
| 2.50 | −189.27663 (−189.27613) | −18.23 (−17.91) |
| 2.62 | −189.27752 (−189.27543) | −18.79 (−17.48) |
| 2.74 | −189.27809 (−189.27436) | −19.14 (−16.80) |
| 2.85 | −189.27778 (−189.27320) | −18.95 (−16.08) |
| 3.00 | −189.27654 (−189.27146) | −18.17 (−14.98) |
| 3.40 | −189.27190 (−189.26697) | −15.26 (−12.17) |
| 4.00 | −189.26569 (−189.26169) | −11.36 ( −8.85) |
| 6.00 | −189.25576 (−189.25348) | −5.13 ( −3.70) |
| 8.00 | −189.25223 (−189.25067) | −2.92 ( −1.94) |
| 10.00 | −189.25059 (−189.24957) | −1.89 ( −1.25) |
| 12.00 | −189.24968 | −1.32 |
| ∞ | −189.24758 | 0.0 |

[a] Numbers in parentheses correspond to a second trajectory in which the OCO angle was fixed at 126.9°.

A second trajectory has also been used in the Monte Carlo calculations for comparison. It corresponds to a more perpendicular attack consistent with Burgi's crystallographic approach vector[7a,b] and with conventional geometrical ideas for addition reactions. In this case the OCO angle was fixed at its value of 126.9° at the transition state (C) for all larger values of $R_{CO}$. The resulting energy surface in $C_s$ symmetry is shown by the dashed curve in Figure 1 and no longer contains the ion-dipole minimum.

Numerical details on the variation of the energy and geometry along the reaction paths are given in Tables II and III. As the reactants approach each other from infinity very little change in the reactants' structures occurs until they are 2.74 Å apart; however, the interaction energy is −19.14 kcal/mol relative to the reactants. It is at this point that the geometric parameters start to undergo noticeable change. The change in angle of approach from linear toward tetrahedral is accomplished between 2.74 and 2.39 Å, which is at the transition state. From $R_{CO}$ = 2.39 Å to the minimum at 1.47 Å, the principal changes involve reorientation of the hydroxyl hydrogen, pyramidilization at carbon, and lengthening of the C=O bond. The O–H bond length remains constant throughout the reaction at 0.95 Å. Charges as calculated from Mulliken population analysis begin changing at 4.0 Å. Both oxygens show smooth charge change along the reaction path. The hydroxyl oxygen starts with a charge of −1.33*e* and ends with −0.83*e* at the tetrahedral intermediate, while the carbonyl oxygen goes from −0.39*e* to −0.95*e*. The hydroxyl hydrogen starts with 0.33*e* and ends with 0.48*e*. Note that the charge delocalization between the oxygens at the tetrahedral intermediate is substantial in contrast to conventional notions based on Lewis structures.

**Discussion.** Overall, the lowest energy reaction path found here is similar to the one obtained by Burgi et al. for hydride attack.[7c] However, in their study and in the previous calculations with smaller basis sets for $OH^-$ adding to formaldehyde and formamide, no intermediates were found along the reaction paths.[7e,h] The well depth that we find of only 1.1 kcal/mol could easily be swamped

**Table III.** Geometric Variations at the 6-31+G* Level for the Reaction $H_2C{=}O + OH^-$ [a]

| $R_{CO}$ | C=O | CH | H—C=O | XHC=O[b] | O—C=O | H—O—C | O—H |
|---|---|---|---|---|---|---|---|
| 1.00 | 1.374 | 1.185 | 107.4 | 128.1 | 114.0 | 110.8 | 0.950 |
| 1.30 | 1.330 | 1.121 | 112.5 | 121.7 | 112.5 | 104.0 | 0.950 |
| 1.40 | 1.315 | 1.113 | 114.0 | 119.5 | 112.3 | 102.2 | 0.949 |
| 1.47 | 1.305 | 1.109 | 115.0 | 117.9 | 112.2 | 101.1 | 0.949 |
| 2.00 | 1.229 | 1.087 | 120.5 | 103.4 | 114.2 | 100.4 | 0.950 |
| 2.28 | 1.210 | 1.082 | 121.6 | 96.3 | 120.6 | 111.1 | 0.950 |
| 2.39 | 1.206 | 1.081 | 121.9 | 93.9 | 126.9 | 124.5 | 0.950 |
| 2.50 | 1.204 | 1.081 | 122.3 | 91.3 | 139.9 | 150.2 | 0.950 |
| 2.62 | 1.204 | 1.082 | 123.0 | 90.0 | 167.7 | 173.1 | 0.950 |
| 2.74 | 1.203 | 1.084 | 123.2 | 90.0 | 180.0 | 180.0 | 0.950 |
| 2.85 | 1.202 | 1.085 | 123.1 | 90.0 | 180.0 | 180.0 | 0.950 |
| 3.00 | 1.200 | 1.086 | 123.1 | 90.0 | 180.0 | 180.0 | 0.950 |
| 3.40 | 1.197 | 1.087 | 122.8 | 90.0 | 180.0 | 180.0 | 0.950 |
| 4.00 | 1.194 | 1.088 | 122.4 | 90.0 | 180.0 | 180.0 | 0.950 |
| 6.00 | 1.190 | 1.089 | 122.0 | 90.0 | 180.0 | 180.0 | 0.950 |
| 8.00 | 1.188 | 1.090 | 121.9 | 90.0 | 180.0 | 180.0 | 0.950 |
| 10.00 | 1.188 | 1.090 | 121.9 | 90.0 | 180.0 | 180.0 | 0.950 |
| 12.00 | 1.187 | 1.090 | 121.9 | 90.0 | 180.0 | 180.0 | 0.950 |
| ∞ | 1.186 | 1.091 | 121.8 | 90.0 | | | 0.953 |

[a] All bond lengths in angstroms and bond angles in degrees. [b] XHC=O is the dihedral angle for the formaldehyde hydrogens with X being perpendicular to the formaldehyde plane.

by the 30–40 kcal/mol overestimates of the exothermicity with smaller basis sets. Of course, the viability of our shallow minimum upon basis set extension and zero-point and correlation energy considerations is uncertain. However, confidence can be expressed for the two-stage nature of the reaction profiles for the gas phase in Figure 1, i.e., predominant ion–dipole attraction at long range with covalent bond formation dominating below $R_{CO}$ = 2.5 Å. It may be noted that computations by Wu et al. for methoxide addition to formaldehyde are consistent with the present results.[14] They find an intervening barrier of ca. 4 kcal/mol and an energy difference between the reactants and tetrahedral intermediate of ca. 46 kcal/mol at the 3-21G level.

The rate data of Asubiojo and Brauman for substitution reactions of acyl halides indicate the presence of a barrier between reactants and product. Though they suggest that the corresponding transition state is the tetrahedral structure, for the present system the tetrahedral structure is clearly a minimum. The intervening complex is a hydrogen bonded structure that has approximately the same energy as the ion–dipole complex with $C_s$ symmetry. The occurrence of such intermediates along the gas-phase reaction path seems particularly reasonable for systems like $Cl^- + CH_3COCl$ studied by Brauman, in which the acyl halide has a large dipole moment and formation of the tetrahedral species is less exothermic than in the present case. However, a triple-well surface seems possible with the tetrahedral species still as an energy minimum.

## Potential Functions

**Analytic Description of the Gas-Phase Surface.** The statistical mechanics calculations require information on the gas-phase reaction to be in the form of continuous functions along the reaction coordinate.[8] A convenient and versatile method to convert the data calculated for discrete points to a continuous function is through the use of cubic spline interpolation.[15] Briefly, the procedure involves calculating the coefficients for a cubic polynomial for an interval between two points, at the same time smoothly splicing the intervals together, giving as the final result a continuous function. Full details of the method are given in reference 15a. The FORTRAN subroutines provided in that reference were used directly for the splining in our program. The gas-phase curves in Figure 1 are the result of spline interpolation of the data in Table II. The geometrical variations along the reaction path were also described by splining the data in Table III.

**Intermolecular Potential Functions.** Before the simulation of the reaction in aqueous media is undertaken, a complete description of the intermolecular interactions is needed. The solvent–solvent and solute–solvent interactions are assumed to be pairwise additive and are described with Coulomb and Lennard-Jones terms acting between sites on the ions and molecules (eq 2). Standard combining rules are used; i.e., $A_{ij} = (A_{ii}A_{jj})^{1/2}$

$$\epsilon_{mn} = \sum_{i}^{\text{on } m} \sum_{j}^{\text{on } n} (q_i q_j e^2/r_{ij} + A_{ij}/r_{ij}^{12} - C_{ij}/r_{ij}^{6}) \qquad (2)$$

and $C_{ij} = (C_{ii}C_{jj})^{1/2}$. The $A$ and $C$ parameters may also be expressed in terms of Lennard-Jones $\sigma$'s and $\epsilon$'s as $A_{ii} = 4\epsilon_i\sigma_i^{12}$ and $C_{ii} = 4\epsilon_i\sigma_i^{6}$. The TIP4P model was used to describe the water–water interactions.[16] The interaction sites for water are located at the three nuclei and at a fourth point ($M$) on the bisector of the HOH angle 0.15 Å from the oxygen atom toward the hydrogens. The water monomers are held fixed in the experimental geometry for the gas phase with $r_{OH}$ = 0.9572 Å and an HOH angle of 104.52°. The nonzero Lennard-Jones terms and charges are $-2q_H = q_M = 1.04$, $A_{OO} = 6 \times 10^5$ kcal·Å$^{12}$/mol, and $C_{OO}$ = 610 kcal·Å$^6$/mol.

The solute–solvent interactions on the other hand must vary along the reaction coordinate. Determination of the solute parameters was based primarily on 6-31+G* calculations for the reacting system plus a water molecule at key points along the reaction. The optimizations consisted of holding the solutes at their 6-31+G* geometries, holding water at the TIP4P geometry, and optimizing the intermolecular variables. The key points treated along the reaction path were for $R_{CO}$ = 1.00, 1.47, 2.00, 2.50, 4.00 Å and the separated species. The sites of hydration for each point considered are summarized in Figures 3 and 4. Table IV compares the ab initio geometries and energies with those from the potential functions (eq 2) for the structures in Figure 4.

The ab initio interaction energy in Figure 3 for monohydrated hydroxide ion is in excellent agreement with the experimental value of Payzant et al., −25 kcal/mol.[17a] In a separate Monte Carlo calculation, we found that the potential function parameters selected for hydroxide ion give a heat of solution of −100 ± 5 kcal/mol.[18] This compares favorably with the experimental value of −115 ± 10.[17b] The interaction energies for formaldehyde–water

(14) Wu, Yun-Dong; Houk, K. N., unpublished work.

(15) (a) Gerald, C. *Applied Numerical Analysis*; Addison-Wesley: Reading, MA, 1980; p 474–488, 508–509. (b) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637.

(16) (a) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926. (b) Jorgensen, W. L.; Madura, J. D. *Mol. Phys.* **1985**, *56*, 1381.

(17) (a) Payzant, J. D.; Yamdagni, R.; Kerbarle, P. *Can. J. Chem.* **1971**, *49*, 3309. (b) Arshadi, M.; Kebarle, P. *J. Phys. Chem.* **1970**, *74*, 1483.

(18) The Monte Carlo simulation was carried out for hydroxide ion with 216 TIP4P water molecules. The calculation used standard procedures[19a] and covered 3 × 10$^6$ configurations.

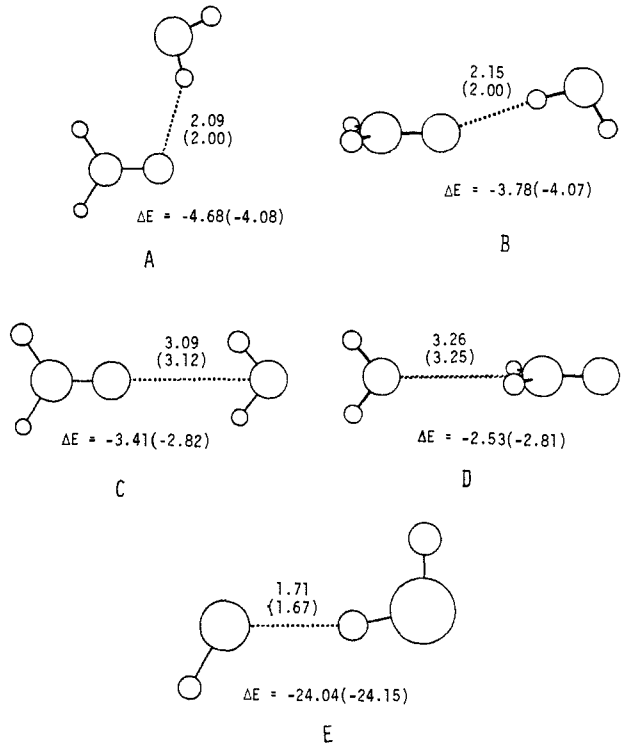



**Figure 3.** Comparison of geometries and interaction energies from the ab initio calculations and potential functions (in parentheses) for monohydrated reactants. Energies in kcal/mol and distances in angstroms.

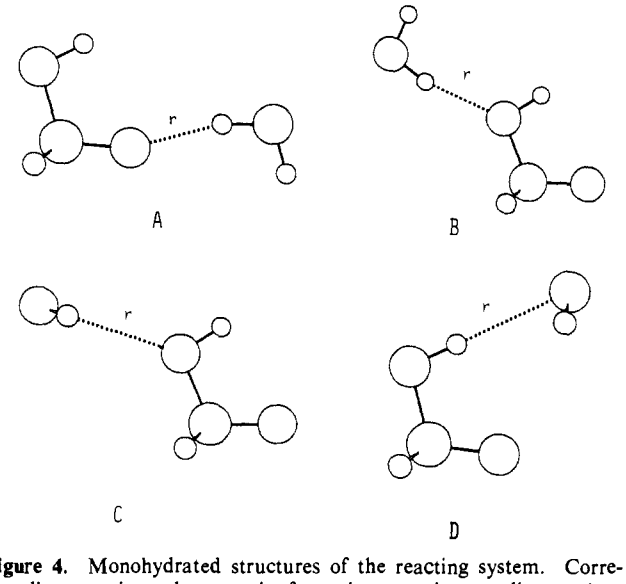



**Figure 4.** Monohydrated structures of the reacting system. Corresponding energies and geometries for various reaction coordinate values are in Tables IV and VI.

also seem reasonable and are similar to other higher level, ab initio results for water–carbonyl interactions.[19] The results in Table IV show a weakening of the hydrogen bond strength for the hydroxide oxygen as it goes from reactants to product. The ab initio results also reveal a change in the optimum monohydrated complex. This change occurs between 2.00 and 1.47 Å; the interaction is stronger for the hydroxyl oxygen at 2.00 Å while the carbonyl oxygen forms the stonger hydrogen bond at 1.47 Å. Naturally, this reflects the charge transfer from the hydroxyl oxygen to the carbonyl oxygen. The results with the potential functions parallel these trends. The development of the solute parameters will now be outlined.

(19) (a) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 1489. (b) Swaminathan, S.; Whitehead, R. J.; Guth, E.; Beveridge, D. L. *J. Am. Chem. Soc.* **1977**, *99*, 7817.

**Table IV.** Summary of Geometries and Interaction Energies from the ab Initio Calculations and the Potential Functions (eq 2)[a]

| | ab initio | | pot. functions | |
|---|---|---|---|---|
| $R_{CO}$ | $-\Delta E$ | $r^b$ | $-\Delta E$ | $r^b$ |
| 1.00 | | | | |
| A | 19.92 | 1.76 | 20.05 | 1.69 |
| B | 7.89 | 2.10 | 8.55 | 1.94 |
| 1.47 | | | | |
| A | 18.02 | 1.78 | 17.09 | 1.73 |
| B | 12.59 | 1.90 | 12.86 | 1.73 |
| C | 11.19 | 2.90 | 10.67 | 2.83 |
| D | 10.40 | 2.31 | 12.23 | 1.91 |
| 2.00 | | | | |
| A | 13.11 | 1.90 | 12.74 | 1.82 |
| B | 18.36 | 1.77 | 17.66 | 1.66 |
| C | 14.99 | 2.81 | 13.33 | 2.78 |
| 2.50 | | | | |
| A | 9.85 | 2.01 | 9.80 | 1.90 |
| B | 20.16 | 1.73 | 19.93 | 1.64 |
| C | 15.70 | 2.83 | 14.45 | 2.78 |
| 4.00 | | | | |
| A | 7.20 | 2.08 | 6.10 | 2.03 |
| B | 23.89 | 1.67 | 23.81 | 1.65 |
| C | 18.19 | 2.73 | 16.70 | 2.78 |

[a] Energies in kcal/mol, distances in angstroms. Structures a–d refer to the geometries in Figure 3. [b] $r$ for structures A and B is the O···H distance, while $r$ for structures C and D is the O···O distance.

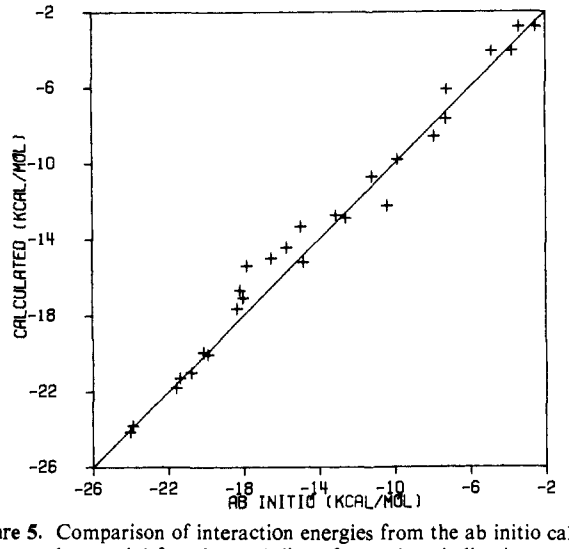



**Figure 5.** Comparison of interaction energies from the ab initio calculations and potential functions. A line of unit slope indicating perfect agreement is shown.

With the ab initio data from Figure 3 and Table IV the solute parameters were obtained by fitting to eq 2. The formaldehyde and hydroxide parameters were taken to be constant from infinity to $R_{CO}$ = 4.00 Å. This assumption is supported by the fact up to $R_{CO}$ = 4.00 Å very little change occurs in the geometries of the solutes. It was also supported by comparisons of interaction energies calculated by ab initio theory vs. the potential functions (eq 2) for formaldehyde with water in the absence of $OH^-$ and when $OH^-$ is located 4.0 Å from the formaldehyde. Consistent with the TIP4P model,[16] the Lennard-Jones terms for hydroxyl hydrogen were assigned to be zero throughout the reaction. However, the carbonyl hydrogens were assigned $\sigma$ = 1.3 Å and $\epsilon$ = 0.03 kcal/mol; these values also remained constant throughout the reaction. The charge on the carbon atom was determined through charge conservation such that the overall charge on the solute was −1. Initial guesses for the rest of the solute parameters were taken from earlier work on liquid hydrocarbons and carboxylate ions.[20a,b] Table V lists the final solute parameters as

(20) (a) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638. (b) Jorgensen, W. L.; Gao, J. *J. Phys. Chem.*, in press. (c) Chandrasekhar, J.; Spellmeyer, D. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 903.

**Table V.** Potential Function Parameters and Biasing Function as a Function of the Reaction Coordinate[a]

| | C[b] | | =O[c] | | $H_C$[d] | —O | | | $H_O$[e] | biasing |
|---|---|---|---|---|---|---|---|---|---|---|
| $R_{CO}$ | $\sigma$ | $\epsilon$ | $\epsilon$ | $q$ | $q$ | $\sigma$ | $\epsilon$ | $q$ | $q$ | function |
| 1.00 | 3.94 | 0.121 | 0.200 | −1.00 | −0.09 | 3.07 | 0.150 | −0.50 | 0.360 | −11.5 |
| 1.30 | 3.92 | 0.119 | 0.200 | −0.94 | −0.07 | 3.07 | 0.160 | −0.67 | 0.368 | −11.0 |
| 1.40 | 3.91 | 0.118 | 0.200 | −0.92 | −0.065 | 3.07 | 0.170 | −0.72 | 0.366 | −10.5 |
| 1.47 | 3.905 | 0.117 | 0.200 | −0.90 | −0.06 | 3.07 | 0.175 | −0.76 | 0.364 | −10.0 |
| 2.00 | 3.86 | 0.1155 | 0.210 | −0.70 | −0.01 | 3.10 | 0.195 | −0.99 | 0.350 | −30.0 |
| 2.28 | 3.85 | 0.114 | 0.220 | −0.61 | 0.03 | 3.12 | 0.200 | −1.10 | 0.340 | −28.0 |
| 2.39 | 3.84 | 0.113 | 0.225 | −0.59 | 0.04 | 3.13 | 0.205 | −1.13 | 0.336 | −26.0 |
| 2.50 | 3.83 | 0.112 | 0.230 | −0.57 | 0.05 | 3.14 | 0.210 | −1.17 | 0.333 | −24.0 |
| 2.62 | 3.82 | 0.111 | 0.235 | −0.55 | 0.06 | 3.15 | 0.215 | −1.19 | 0.330 | −22.0 |
| 2.74 | 3.81 | 0.110 | 0.242 | −0.52 | 0.07 | 3.16 | 0.220 | −1.22 | 0.326 | −20.0 |
| 2.85 | 3.80 | 0.109 | 0.244 | −0.50 | 0.08 | 3.17 | 0.225 | −1.24 | 0.323 | −18.0 |
| 3.00 | 3.79 | 0.108 | 0.246 | −0.48 | 0.09 | 3.18 | 0.230 | −1.27 | 0.320 | −14.0 |
| 3.40 | 3.77 | 0.107 | 0.248 | −0.42 | 0.095 | 3.19 | 0.240 | −1.29 | 0.310 | −10.0 |
| 4.00 | 3.75 | 0.105 | 0.250 | −0.38 | 0.10 | 3.20 | 0.250 | −1.30 | 0.300 | −6.0 |
| 6.00 | 3.75 | 0.105 | 0.250 | −0.38 | 0.10 | 3.20 | 0.250 | −1.30 | 0.300 | −2.0 |
| 8.00 | 3.75 | 0.105 | 0.250 | −0.38 | 0.10 | 3.20 | 0.250 | −1.30 | 0.300 | −1.0 |
| 10.00 | 3.75 | 0.105 | 0.250 | −0.38 | 0.10 | 3.20 | 0.250 | −1.30 | 0.300 | 0.0 |
| 12.00 | 3.75 | 0.105 | 0.250 | −0.38 | 0.10 | 3.20 | 0.250 | −1.30 | 0.300 | 0.0 |

[a] Units for $\sigma$ are in angstroms, $\epsilon$ in kcal/mol, $q$ in electrons, and the biasing function in kcal/mol. [b] Charge on C chosen to give a net charge of −1 for the reacting system. [c] $\sigma$ for =O is 3.20. [d] $\sigma$ and $\epsilon$ for $H_C$ are 1.30 and 0.03. [e] $\sigma$ and $\epsilon$ for $H_O$ are both zero.

**Table VI.** Summary of Geometries and Interaction Energies from the ab Initio Calculations and the Potential Functions (eq 2) for the Second Trajectory[a]

| | ab initio | | pot. functions | |
|---|---|---|---|---|
| $R_{CO}$ | $-\Delta E$ | $r^b$ | $-\Delta E$ | $r^b$ |
| 2.50 | | | | |
| B | 16.5 | 2.79 | 15.0 | 2.77 |
| C | 20.8 | 1.73 | 21.0 | 1.63 |
| 2.74 | | | | |
| B | 14.8 | 2.76 | 15.2 | 2.77 |
| C | 21.4 | 1.72 | 21.3 | 1.63 |
| 3.00 | | | | |
| B | 17.8 | 2.71 | 15.4 | 2.77 |
| C | 21.6 | 1.72 | 21.8 | 1.63 |

[a] Energies in kcal/mol, distances in angstroms. Structures B and C refer to the geometries in Figure 4. [b] $r$ for structure B is the O···H distance, while it is the O···O distance for C.

a function of the reaction coordinate. The parameters at $R_{CO}$ = 1.00, 1.47, 2.00, 2.50, and 4.00 Å are the points obtained from the fitting. The other values were obtained by cubic spline interpolation. This table was subsequently used as the basis for further interpolation in the Monte Carlo program.

The potential function variation with reaction coordinate for the second trajectory was taken to be the same as that used for the first trajectory. This was validated by running several water–solute ab initio optimizations at $R_{CO}$ = 2.50, 2.74 and 3.00 Å and comparing the results with those from the potential functions. Table VI summarizes these results and reveals only small discrepancies. Finally a plot of all the ab initio interaction energies vs. the values predicted by the potential functions is shown in Figure 5. The points fall along a line of unit slope with a mean error of 0.33 kcal/mol and a correlation coefficient of 0.98.

A key observation from these results is that the hydrogen bonding to the hydroxide ion becomes weaker as it proceeds along the reaction path toward the product. Though the interaction energies for the carbonyl oxygen become more favorable simultaneously, the interactions are stronger for the hydroxide ion. Thus, it is likely that the reactants are better solvated than the product or intermediate points on the reaction path.

## Monte Carlo Simulations

In order to determine the reaction profile in aqueous solution, the same computational approach was used as in the previous work on the $S_N2$ reaction.[8] The potential of mean force, $w(r_c)$, is calculated as a function of the reaction coordinate, $r_c$. Furthermore, $w(r_c)$ is given by $-k_BT \ln g(r_c)$, where $g(r_c)$ is the probability of occurrence of each value of $r_c$. Precise computation of $g(r_c)$ is not possible through a single long simulation but requires the use of importance sampling.[21] In this procedure, probability distributions are obtained for different overlapping sections of the reaction coordinate and spliced together to obtain a single distribution covering the entire range. This is accomplished by constraining the solute to remain in limited regions by using fictitious forces known as umbrella potentials. The regions or "windows" are chosen such that all values of interest for the reaction coordinate are adequately sampled. The Monte Carlo walk is generated in the usual manner,[22] but the probability of selecting a configuration is made proportional to $e^{-\beta(u+u')}$ rather than to the usual Boltzmann factor $e^{-\beta u}$. Here, $u$ is the true configurational energy, while $u'$ is the umbrella potential added to keep the solute within the chosen window, and $\beta = 1/k_BT$. All calculated properties from such simulations can be normalized to remove the effects of the non-Boltzmann sampling.[21] The true ensemble average of a property $Q$ is given by 3, where $\langle\ \rangle_w$ refers

$$\langle Q \rangle = \langle Qe^{\beta u'} \rangle_w / \langle e^{\beta u'} \rangle_w \quad (3)$$

to an average in the biased system. In particular, the frequency of occurrence of a value of the reaction coordinate in window $i$, $P_i(r_c)$, leads to the function $g_i(r_c)$ on normalization through eq 3. The normalized distributions, $g_i(r_c)$, for adjacent windows are matched at points of maximum overlap of the corresponding $P_i(r_c)$ functions. By splicing the results from all the windows, a single distribution $g(r_c)$ is obtained, from which the potential of mean force, $w(r_c)$, is calculated.

**Details of the Simulations.** Monte Carlo simulations in the NPT ensemble were performed on a orthorhombic box containing 269 TIP4P water molecules and the solutes (hydroxide and formaldehyde). Average dimensions of the cell were 16.50 × 19.80 × 24.75 Å. The shape was chosen to accommodate both trajectories by keeping at least 15 Å between images of solute atoms. Metropolis sampling and periodic boundary conditions were employed. The importance sampling involved nine windows for the first trajectory corresponding to the lowest energy $C_s$ pathway in the gas phase. The two solutes were constrained by a harmonic force of the form $^1/_2k(x - x_0)^2$, where $x_0$ is the center of the window and $k$ is the force constant. The windows used for the first trajectory were centered at $R_c$ = 1.47, 1.90, 2.30, 2.50, 2.60, 3.00, 4.00, 5.00, and 6.50 Å and the corresponding force constants were $k$ = 30.0, 60.0, 15.0, 5.0, 15.0, 3.0, 5.0, and 1.0 kcal/mol·Å$^2$.

As for the $S_N2$ reaction,[8] an additional biasing function was needed to permit adequate sampling of the high-energy regions

(21) Patey, G. N.; Valleau, J. P. *J. Chem. Phys.* **1975**, *63*, 2334. Pangali, C. S.; Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, *71*, 2975.
(22) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

along the reaction coordinate. The perfect biasing function would be the mirror image of the potential of mean force. Of course, this was unknown though some rough guidance on the possible barrier height and position could be obtained from the studies by Guthrie[5] and Weiner et al.[9] The final form was arrived at largely by trial and error and is summarized in Table V. Again, a continuous function was obtained by cubic splining.

Acceptance of new configurations was based on the Metropolis test modified for preferential sampling in which the probability of attempting to move a solvent molecule was made proportional to $1/(r^2 + c)$, where $r$ is the shorter of the hydroxide oxygen–water or carbonyl carbon–water distance.[23] The constant $c$ was set to 150 Å$^2$, which caused the water molecules nearest each solute to be moved 3–4 times as often as distant waters.

New configurations were generated in the usual manner. For solvent moves, the randomly chosen monomer was randomly translated in all three Cartesian directions and randomly rotated about a randomly chosen axis. Solute moves were attempted every 60 configurations. A solute move consisted of a random variation of $R_{CO}$ and a random translation along the long axis of the periodic cell. The new geometry, potential function parameters, and gas-phase energy were calculated with the cubic spline interpolation algorithm. Volume moves were attempted every 1250 configurations by scaling the coordinates. For each new configuration, the energy was computed from the sum over all the interaction energies (eq 2) out to a cutoff of 7.5 Å based upon the $O_w$–$O_w$, C–$O_w$ and $O_H$–$O_w$ distances. This includes interactions with ca. 3 shells of water around the solutes and is expected to be adequate for studying the differential effects along the reaction path that are of principal interest here. Due to the volume fluctuations in NPT simulations, increasing the cutoff significantly would require a larger system size. The ranges of the moves were chosen to yield an acceptance rate of about 40% for new configurations. The run for each window involved equilibration for ca. $5 \times 10^5$ to $10^6$ configurations, followed by averaging over an additional $(2–4) \times 10^6$ configurations. The first three windows covering $R_{CO}$ = 1.35–2.40 Å ran smoothly. The next two windows were more troublesome. This was due to the fact that the OCO angle changes rapidly over a small range of $R_{CO}$ near 2.5 Å. Thus, concomitant reorganization of the solvent is significant in this region. The difficulties were overcome with close spacing of the windows in this region and longer runs.

The problem is avoided in the second trajectory since the OCO angle is fixed beyond $R_{CO}$ = 2.39 Å. The centers of the windows for this second trajectory were at $R_{CO}$ = 1.47, 1.90, 2.30, 3.00, 3.60, 4.00, 5.40, 5.50, 5.70, and 6.50 Å and the corresponding force constants were 30.0, 60.0, 15.0, 1.0, 5.0, 10.0, 20.0, 5.0, 10.0, and 1.0 kcal/mol·Å$^2$. The latter windows were spliced with the first three windows of the first calculation. Finally, three simulations in which $R_{CO}$ was fixed at 1.45, 2.05, and 7.40 Å corresponding to the product, aqueous transition state, and reactants were performed. The setup was the same as in the simulations with importance sampling. Preferential sampling was used and averaging occurred over $(2–3) \times 10^6$ configurations. These simulations were needed to permit characterization of the changes in solvation at these key points. The runs with importance sampling do not yield a large enough number of configurations at precisely these points for this purpose.

## Results of Simulations

**Gas Phase and Aqueous Energy Profiles.** From the beginning one of the main goals was to observe the change in the energy profiles on going from the gas phase to aqueous solution. The results are shown in Figure 1, where the upper curves represent the aqueous phase potential of mean force and the lower curves are the gas-phase potential energy. The solid lines correspond to the initially collinear trajectory (trajectory 1), while the dashed lines are for the more orthogonal approach (trajectory 2). It must be emphasized that the top curves represent free energy while the bottom curves represent potential energy.

With this in mind, the differences can be analyzed. In aqueous solution, the free energy profile rises only gradually from the reactants to $R_{CO}$ = 3 Å. In this region the loss of solvation energy is being balanced by the increasing ion–dipole attraction for the reactants. However, between 2 and 3Å the gas-phase energy for the reactants is relatively constant. The desolvation is no longer offset and there is a rapid rise in the free energy profile for aqueous solution to the tranition state at $R_{CO}$ = 2.05 Å (roughly B in Figure 2). At this point the gas-phase energy begins its rapid descent with the onset of covalent bond formation. The descent is mirrored in the profile for the aqueous phase and indicates little change in the hydration energy for the transition state or product. Though the term "desolvation" has been used here, it is important to note, as shown below, that this in fact corresponds to a weakening of hydrogen bonds rather than to a reduction in their number. It should also be realized that these computations have provided an a priori determination of a transition state in solution. It is clear that in contrast to the $S_N2$ reaction,[8] the barrier for the addition reaction in solution is nearly entirely solvent-induced. The scope of this observation for addition reactions with other substrates, particularly acyl halides, needs further study.[7i]

Qualitatively, the results for the two trajectories are the same. In both cases, the reaction profiles in water show a single barrier with no statistically significant minima other than the product. There are no ion–dipole or solvent-separated intermediates. Quantitatively, the transition-state theory free energy of activation for the first trajectory is 28 kcal/mol and the product is 14 kcal/mol above the reactants. The corresponding values for the second trajectory are 24 and 10 kcal/mol. The discrepancy of 4 kcal/mol between the results of the two trajectories may reflect a difference in the free energies for the two geometries of the reactants, or, more likely, it is a reasonable estimate of the absolute uncertainty in the calculations. Some of this error comes from a degree of arbitrariness in the splicing procedure.

The reaction profiles calculated here are consistent with those reported by Guthrie[5] and by Weiner et al.[9] for alkaline hydrolyses of amides. For their systems, the activation energies are ca. 22 kcal/mol and the tetrahedral intermediate is 9–18 kcal/mol above the reactants. Activation energies for ester hydrolysis are typically somewhat lower at 15–20 kcal/mol.[5,24] The location of the transition state in the present case and in that of Weiner et al. is also approxmiately the same at $R_{CO} \approx 2.0$ Å.[9] Thus, the transition state may be classified as very late. The shift in its position from the gas phase ($R_{CO}$ = 2.39 Å) to the aqueous phase ($R_{CO}$ = 2.05 Å) is in accordance with Hammond's postulate.[25]

The potential sources of error in the present study should be made clear. The intermolecular potential functions and the assumption of pairwise additivity for the intermolecular interactions are certainly not perfect. The latter approximation is of greatest concern for the hydroxide ion and probably leads to some artificial lowering of the reactant end of the aqueous potential of mean force; i.e., the computed free energy of activation is likely to be somewhat too high. Also, there has been no treatment of the vibrational energy changes for the reactants, though the principal concern would be with the differential effects between the gas phase and aqueous solution. Since we could not compute the vibrational frequencies, a free energy curve could not be reported for the gas phase in Figure 1. Typically, ion–molecule association reactions have $\Delta S = -20 \pm 5$ cal/mol·deg in the gas phase[26] and would yield a $-T\Delta S$ contribution of ca. 6 kcal/mol at 25 °C. At the 3-21+G level, we find $\Delta H$ at 25 °C including the zero-point corrections to be 3.5 kcal/mol less exothermic than the SCF energy difference. This would likely be cancelled by the correlation correction. Thus, the well depth for the tetrahedral intermediate is anticipated to be ca. 6 kcal/mol less on the free energy scale than the SCF results in Figure 1.[27]

(23) (a) Owicki, J. C. *ACS Symp. Ser.* **1978**, *No. 86*, 159. (b) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304.

(24) Bender, M. L.; Ginger, R. D.; Unik, J. P. *J. Am. Chem. Soc.* **1958**, *80*, 1044.

(25) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334.

(26) Dougherty, R. C. *Org. Mass Spectrom.* **1974**, *8*, 85. Sharma, D. K. S.; Kebarle, P. *J. Am. Chem. Soc.* **1982**, *104*, 19.

**Table VII.** Summary of Thermodynamic Results from the Fixed-Solute Simulations Representing the Reactants, Transition State, and Product[a]

| | react. | t.s. | product |
|---|---|---|---|
| $E_{sx}$ | −221 ± 3 | −151 ± 1 | −154 ± 1 |
| $E_{ss}$ | −2646 ± 10 | −2648 ± 9 | −2636 ± 8 |
| $E_{ss}^{*}$ [b] | −2720 ± 8 | −2720 ± 8 | −2720 ± 8 |
| $\Delta E_{ss}$ | 74 ± 13 | 72 ± 12 | 84 ± 11 |
| $E_{xx}$ | −3.2 | −19.5 | −35.0 |
| $E_{tot}$ | −2870 ± 7 | −2818 ± 4 | −2825 ± 7 |
| $V$ | 8132 ± 25 | 8081 ± 19 | 8157 ± 23 |
| $V^{*}$ [b] | 8092 ± 24 | 8092 ± 24 | 8092 ± 24 |
| $\Delta V_{sol}$ | 40 ± 35 | −11 ± 31 | 65 ± 33 |

[a] Energies are in kcal/mol and volumes are in Å$^3$. See eq 4 and accompanying text for definition of computed quantities. [b] Computed from values in ref 16b.

**Thermodynamics from the Fixed-Point Simulations.** While the use of importance sampling is essential for the determination of the potential of mean force, the procedure is not well suited for obtaining solute–solvent structure and energy distributions. Thus, the fixed solute simulations were run for the reactants, aqueous transition state, and product. The total energy of a solution with two solutes is defined in eq 4 as the sum of the solute–solvent

$$E_{tot} = E_{sx} + E_{ss} + E_{xx} \quad (4)$$

energy ($E_{sx}$), the solvent–solvent energy ($E_{ss}$), and the solute–solute energy ($E_{xx}$). Table VII summarizes the results of the three simulations. The reported statistical uncertainities ($\pm 1\sigma$) were obtained from separate averages over blocks of $2 \times 10^5$ configurations. The $E_{sx}$ values for the three fixed solutes are remarkably similar to those reported by Weiner et al. for the corresponding points for the hydroxide ion plus formamide reaction using the TIP3P potential for water.[9] This quantity converges readily and has a relatively small statistical uncertainty. The similarity between the $E_{sx}$ for the transition state and product emphasizes that they are solvated relatively equally. Therefore, it is the $E_{xx}$ term that lowers the free energy profile for the product relative to the transition state. The stronger hydration of the reactants is clearly reflected in their lower $E_{sx}$ of −221 kcal/mol.

The statistical uncertainties for the total energies and volumes are substantially greater and undoubtedly underestimated in Table VII. This results from the dominant contributions of the water–water interactions to these quantities. For example, $\Delta E_{ss}$ measures the disruption of the solvent by the solutes and is calculated by substracting $E_{ss}$ from $E_{ss}^{*}$, which is the energy for the pure solvent.[16b] Experience with numerous simulations of ionic solutions has shown that as $E_{sx}$ becomes more favorable $\Delta E_{ss}$ increases and the ratio of the two is normally less than 2.5:1 for anions.[20b,c] On this basis the results for the product and transition state in Table VII appear reasonable; however, the solvent disruption seems underestimated for the reactants. This translates into the reactants having a ca. 45 ± 10 kcal/mol lower total energy than the product from the fixed-point simulations. The potential of mean force calculations indicate a much smaller difference, though entropy may disfavor the more strongly hydrated reactants. The substantial uncertainty in the direct computation of heats of solution makes fixed point simulations unsuitable for constructing reaction profiles with acceptable precision.

The computed volumes of solution have little analytical value except to point out that the density of the systems remained close to the value for pure water at 25 °C and 1 atm (8100 Å$^3$/269 = 30.1 Å$^3$/water molecule = 0.993 g cm$^{-3}$).

**Energy Distributions.** Important details on the solute–solvent bonding are provided in Figure 6. These distributions are for the individual solute–water interaction energies. As expected, the reactant curve (short dashes) shows a low-energy band representing the strong hydrogen bonding between water molecules and the hydroxide ion. Integration of this band to an energy to −11.4 kcal/mol yields 6.0 water molecules presumably solvating the hydroxide ion. The transition-state curve (long dashes) shows two bands at low energy. The lowest can be assigned to water molecules solvating the hydroxide fragment (see charges in Table V) while the next band is for the carbonyl oxygen that is becoming more negatively charged. Integration of the lower band to −13.0 kcal/mol yields 3.4 water molecules, and integration of the second band from −13.0 to −8.5 kcal/mol reveals 3.6 water molecules. The product also shows two energy bands. Integration of the first band to −12.5 reveals 3.6 water molecules, and the second band integrated from −12.5 to −8.5 yields 2.8. On the basis of the charge shifts, the assignment of the bands is likely reversed on going from the transition state to the product. The important observation here is that the total number of strong solute–water hydrogen bonds is relatively constant at 6–7 along the reaction path. However, the hydrogen bonds for the hydroxide ion are significantly stronger. The interactions between the hydroxide ion and water molecules in the second hydration shell are also expected to be more favorable. These interactions appear in the shoulder near −8 kcal/mol in Figure 6. Thus, the barrier for the reaction can be attributed primarily to reduction in hydrogen bond strengths rather than in numbers of hydrogen bonds. The same observation was made for the solvent-induced component of the barrier for the $S_N2$ reaction.[8]

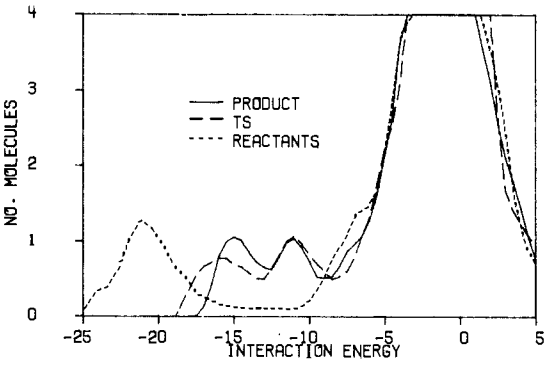



**Figure 6.** Computed solute–solvent energy pair distributions for the reactants, transition state, and product. Units for the ordinate are molecules per kcal/mol.

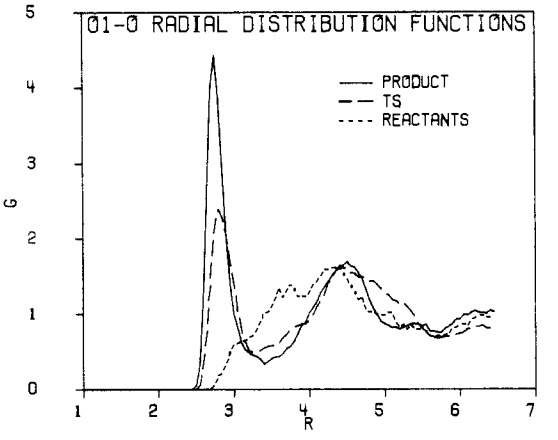



**Figure 7.** Computed O1–$O_w$ rdfs for the reactants, transition state, and product. Distances are in angstroms throughout. O1 is the carbonyl oxygen.

**Radial Distributions Functions.** The solute–solvent structure can be further characterized through radial distribution functions (rdfs), proximity analysis, and stereoplots, as presented in the following sections.

An rdf, $g_{xy}(r)$, gives the probability of finding an atom of type y a distance $r$ from an atom of type x. Integration of the first peak in an rdf to the first minimum also gives an indication of the coordination number for y atoms about x. The solute–solvent radial distribution functions are shown in Figures 7–12. Labeling in the figures uses the following convention. The first atom label is for the solute atom followed by a "1" or "2" to distinguish atoms from the formaldehyde and hydroxide fragments, respectively.

(27) For further discussion of free energy profiles in the gas phase, see: Garrett, B. C.; Truhlar, D. G. *J. Phys. Chem.* **1979**, *83*, 1052.

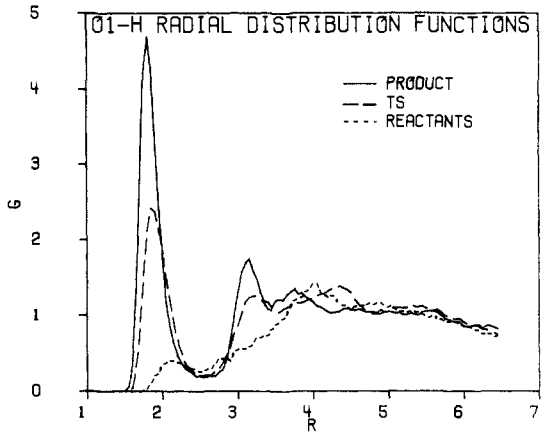



**Figure 8.** Computed O1–$H_w$ rdfs for the reactants, transition state, and product.

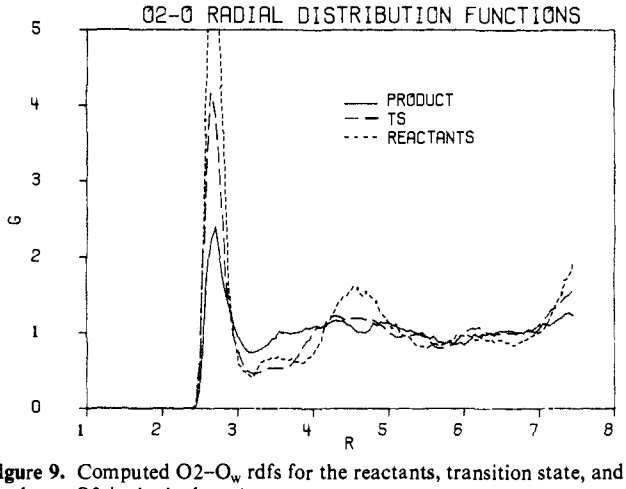



**Figure 9.** Computed O2–$O_w$ rdfs for the reactants, transition state, and product. O2 is the hydroxyl oxygen.

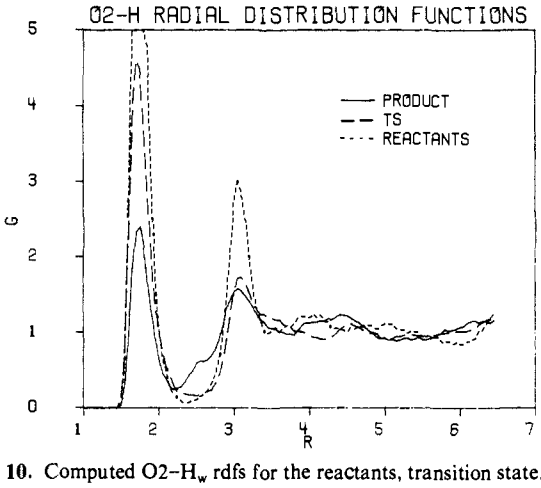



**Figure 10.** Computed O2–$H_w$ rdfs for the reactants, transition state, and product.

The second atom label is for the oxygen or hydrogen of water.

Key interactions occur around the hydroxide and carbonyl oxygens. The corresponding rdfs are found in Figures 7–10. The first observation is the sharp peaks near 2.8 Å for the O1–O and O2–O rdfs. These peaks reflect the first hydration shells around the oxygens and may be integrated to estimate hydration numbers. For the carbonyl oxygen, integration of the O1–O rdfs to 3.4 Å yields 1.7, 3.6, and 4.7 waters for the reactants, transition state, and product. The substantial increase in solvent structure about this atom in going from reactants to product is clearly evident in Figure 7. There is a concomitant decrease in solvent structure around the hydroxyl oxygen as shown in Figure 9; integration of the O2–O rdfs to 3.2 Å yields 6.3, 4.3, and 3.2 water molecules

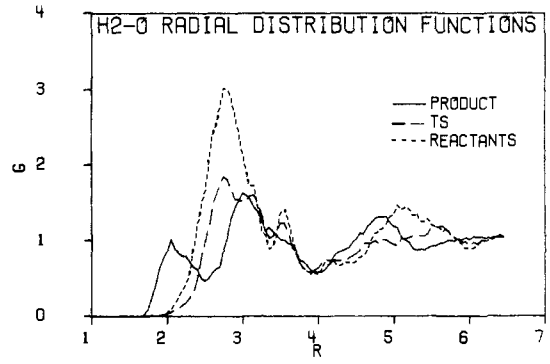



**Figure 11.** Computed H2–$O_w$ rdfs for the reactants, transition state, and product. H2 is the hydroxyl hydrogen.

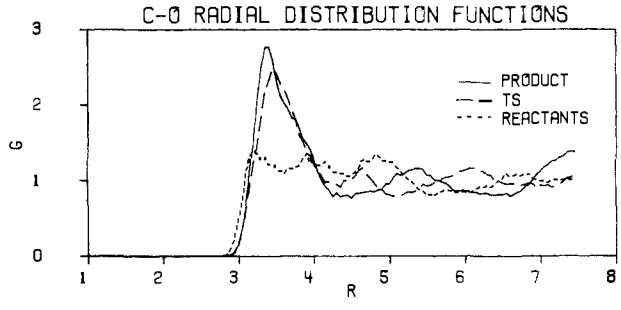



**Figure 12.** Computed C–$O_w$ rdfs for the reactants, transition state, and product.

for the reactants, transition state, and product. It should be noted that the sum of these coordination numbers remains nearly constant at 8 along the reaction path. This is consistent with the pattern from the energy pair distributions and from the proximity analyses presented below.

The first peaks in the O1–H and O2–H rdfs (Figures 8 and 10) also reflect the hydrogen bonding and provide information on the orientation of the water molecules. The peaks at 1.8 and 3.0 Å correspond to the near and far hydrogens of waters hydrogen bonding to oxygens of the solutes in a "linear" geometry. The first peaks in the O2–H rdfs integrate to 6.0, 4.1, and 2.3 for the hydroxide ion, transition state, and product. The corresponding data for the first peak of the O1–H rdfs are 0.8, 3.2, and 4.4. These values provide reasonable estimates of the number of hydrogen bonds for the oxygen atoms of the solute and are consistent with the integrals of the low-energy bands in the energy pair distributions. Hydrogen bonding may also be anticipated at the hydroxyl hydrogen of the solute. This should be reflected in the first peaks for the H2–O rdfs in Figure 11. For the product, the first peak centered at 2.0 Å contains 1.0 water oxygen, presumably of a hydrogen bond accepting water molecule. However, for the transition state and hydroxide ion, the first peak is higher and shifted to larger separation. In these cases, the substantial negative charge on the hydroxide oxygen (Table V) prevents orientation of water molecules with oxygens toward the hydroxide hydrogen. Thus, the broad peaks centered near 2.6 Å are reflecting the first shell water molecules that are primarily oriented for hydrogen bonding at the hydroxide oxygen.

The C–O rdfs (Figure 12) give some indication of the change in solvation at the methylene group. The structureless rdf for the reactant is characteristic of a comparatively hydrophobic solute. Such groups are surrounded by a solvent cage, which can be observed in the stereoplots. The strong peaks at 3.3 Å for the transition state and product rdfs arise from the waters that are interacting with the two oxygen atoms.

**Proximity Analysis.** To further clarify the origin of the activation barrier upon solvation, proximity analyses of the results for the three-fixed-point calculations were performed.[28] The procedure was developed by Beveridge and co-workers to help in

(28) Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102*,4287.

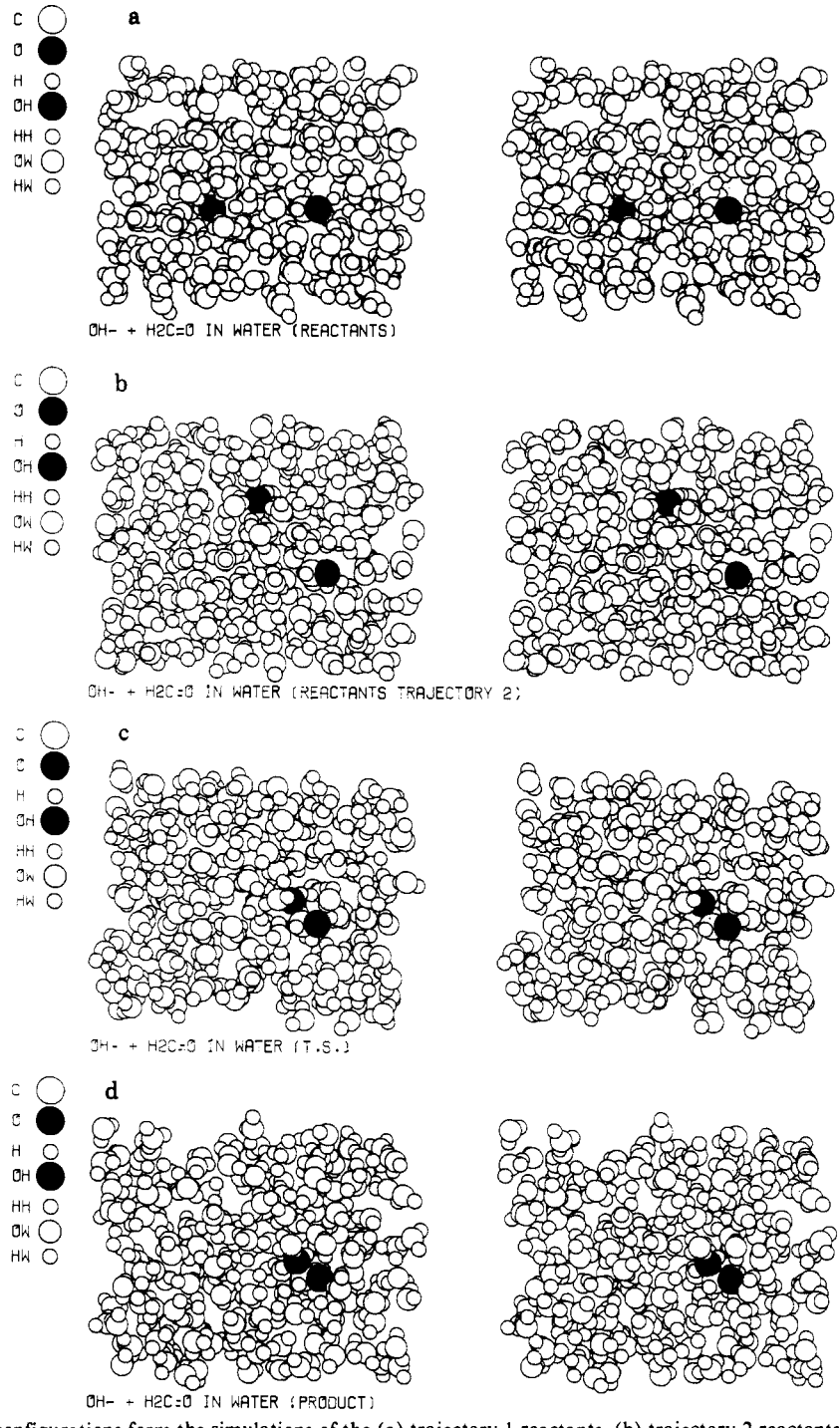



**Figure 13.** Stereoplots of configurations form the simulations of the (a) trajectory 1 reactants, (b) trajectory 2 reactants, (c) transition state, and (d) the product. Water molecules more than 3 Å in front of the solute have been removed for clarity.

the analysis of coordination numbers for nonspherical solutes. The principal interest is in the first solvation shell. This may be defined by using cutoff distances for each type of atom in the solute. Then if a solvent molecule is nearer that atom than any other atom in the solute and it is within the cutoff range, it can be counted as in the primary coordination shell for the atom. Once the solvent molecules are assigned in this manner, numerous distributions for the coordination numbers and solute–solvent energies can be defined.

The cutoffs are the same as those used previously for amides[19] except for $H_C$–$O_w$. The $H_C$–$O_w$ cutoff was set at 3.0 Å, which coincides with the location of the first minimum in $H_1$–O rdf for formaldehyde. $H_C$–$O_w$ rdfs for neutral solutes typically have little structure so this cutoff is rather arbitrary.

The results of the proximity analyses for primary coordination numbers based on the three fixed point simulations are presented in Table VIII. The average solute–primary water interaction energies were also computed and are listed as well. The results are from analyses of 200–300 configurations saved at regular intervals during each Monte Carlo run.

From the results in Table VIII, it is apparent that the primary coordination numbers for the formaldehyde carbon and hydrogens are fairly constant throughout the reaction. The solvation of the carbonyl and hydroxyl oxygens shows the expected variations. The 5.5 primary waters for the hydroxide ion are consistent with the coordination number from the O2–O rdf. Also, the hydration numbers for the alcoholic oxygen and hydrogen in the product are very similar to those for methanol,[29] i.e., 2 for oxygen and 1 for the hydrogen. The total number of water molecules in the primary hydration shells and the primary energy contributions

(29) Jorgensen, W. L.; Madura, J. D. *J. Am. Chem. Soc.* **1983**, *105*, 1407.

**Table VIII.** Summary of Proximity Analyses[a]

| atom | 1° coord no. | | | solute-water energy[b] | | |
|---|---|---|---|---|---|---|
| | reactants | t.s. | product | reactants | t.s. | product |
| C | 0.13 | 0 | 0 | -0.5 | 0 | 0 |
| H | 1.95 | 1.14 | 1.27 | -4.0 | -4.8 | -6.9 |
| H | 1.23 | 1.61 | 1.55 | -2.0 | -7.9 | -3.9 |
| O= | 0.80 | 2.59 | 3.01 | -2.5 | -10.4 | -14.2 |
| O— | 4.71 | 2.03 | 1.73 | -21.0 | -15.8 | -10.4 |
| $H_O$ | 0.79 | 0.50 | 1.04 | -11.8 | -7.8 | -10.6 |
| total | 9.79 | 7.87 | 8.60 | | | |
| 1° energy contrib | | | | -122.7 | -81.1 | -86.8 |
| $E_{sx}$ | | | | -221 | -151 | -154 |
| % contrib of 1° energy | | | | 56% | 54% | 56% |

[a] Results from the three fixed-solute simulations. Cutoffs for each atom are C, 5.35 Å; $H_C$ 3.00 Å; O=, 3.20 Å; O—, 3.20 Å; $H_O$ 2.55 Å. [b] Average solute-water interaction energy (kcal/mol) for water molecules in the 1° coordination shell of the given atom.

reveal the origin of the activation barrier in solution. Specifically, the 5.5 strong interactions for $OH^-$ become five or six weaker ones for the transition state and product. This reflects hydrogen bond weakening upon charge delocalization as noted above for the monohydrated complexes (Figures 3 and 4).

The quantitative impact of this effect is indicated by the contributions to the total solute-water interaction energies, $E_{sx}$, from the waters in the primary hydration shells. As listed in Table VIII, these contributions are -123, -81, and -87 kcal/mol for the reactants, transition state, and product. The difference of 42 kcal/mol between the reactants and transition state is undoubtedly the principal component of the hydration-induced barrier to the addition reaction.

**Stereoplots.** In closing, four stereoplots representing the last Monte Carlo configuration for the product, the transition state, and the reactants in both trajectories are shown in Figure 13. For clarity, water molecules more than 3 Å in front of the solutes have been removed. These plots illustrate many of the structural ideas discussed above.

For the reactants, the stereoplots clearly show the hydroxide ion participating in about 6 hydrogen bonds with one hydrogen of each water molecule pointing toward the hydroxide oxygen. The formaldehyde sits more in a solvent cavity with about 1 hydrogen bonded water molecule on the carbonyl oxygen. At the transition state, both oxygens of the solute appear to be participating in 3 hydrogen bonds. The density of water molecules around the transformed carbonyl oxygen increases in continuing to the product, though the total number of solute-water hydrogen bonds appears to remain constant at 6-7. In all of the plots except perhaps for the product, it does not appear that the hydroxide hydrogen in participating as a hydrogen bond donor. This seems reasonable, though not obvious a priori, for a relatively small anion on the electrostatic grounds discussed above. It should be remembered that these stereoplots show only one of millions of configurations that were sampled in the Monte Carlo calculations.

## Conclusions

Through the use of quantum and statistical mechanical methods, a comprehensive examination of a nucleophilic addition in the gas phase and in aqueous solution has been carried out. The reaction of $OH^- + H_2C{=}O$ in the gas phase was found to have a shallow ion-dipole minimum along the lowest energy path to the tetrahedral intermediate. Upon hydration, a ca. 25 kcal/mol free energy barrier is introduced; the transition state was located at a CO separation of 2.05 Å. These results and the overall free energy change in water are consistent with prior experimental studies of hydrolysis reactions and with the theoretical results of Weiner et al. for the addition of hydroxide to formamide in water.[5,9,24] Further confirmation of these findings is possible through application of integral equation methods now that the necessary potential functions have been reported here.[30] There are a number of potential sources of error in the present study that should be kept in mind, including the choice and form of the intermolecular potential functions and the lack of intramolecular vibrations. Also, only two trajectories for the reaction in water could be studied since even this required the equivalent of ca. 300 days of VAX 11/780 time. Nevertheless, the present work and the earlier studies of an $S_N2$ reaction[8] represent a significant advance in the theoretical treament of organic reactions in solution. Both studies have yielded not only reasonable energetic results but also detailed insight into the change in solvation along the reaction paths. In particular, a general observation appears to be emerging for these polar reactions: it is not so much the change in number of hydrogen bonds along the reaction paths but rather in their strengths that is primarily responsible for the solvent-induced activation barriers.

**Acknowledgment.** Gratitude is expressed to the National Science Foundation for support of this work.

(30) Chiles, R. A.; Rossky, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 6867.

# Chemisorption Patterns as Predicted by Band Calculations, by Frontier Crystal Orbitals, and by Cluster Calculations: Hydrogen Atoms on Graphite

**John P. LaFemina and John P. Lowe***

*Contribution from the Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802.* *Received September 13, 1985*

**Abstract:** Extended-Huckel band calculations for five patterns of ordered overlayers of hydrogen atoms on unreconstructed graphite (11% coverage) show energy differences as large as ~15 kcal $mol^{-1}$. These energy differences are primarily due to differences in interactions between the H atoms and graphite and not direct interactions between H atoms. The results of the band calculations are, to some degree, in accord with predictions based on the frontier crystal orbitals of graphite, though there are significant differences. Embedded cluster calculations on these systems yield the same frontier orbitals as band calculations, hence the same qualitative predictions, *only when cluster size meets a certain modular requirement*. This prevents smooth convergence with increasing cluster size. As a result, embedded-cluster frontier-orbital predictions of relative pattern stabilities get worse for these systems when one goes from an 18-carbon cluster to a 32-carbon cluster.

When atoms or molecules chemisorb on clean, crystalline surfaces, they sometimes form ordered overlayers, even at coverages of a small fraction of a monolayer.[1-3] At such low coverages, distances between adspecies can be too large for direct

0002-7863/86/1508-2527$01.50/0 