Hydration of the Carbonyl Group. A Theoretical Study of the Cooperative Mechanism

Saul Wolfe,*,‡ Chan-Kyung Kim,† Kiyull Yang,§ Noham Weinberg,‖ and Zheng Shi

Contribution from the Department of Chemistry, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Received June 7, 1994. Revised Manuscript Received February 22, 1995[⊗]

Abstract: The thermochemical parameters, vibrational frequencies, solvent isotope effects, and proton inventories for the neutral hydration of formaldehyde by water and by clusters containing two, three, and four water molecules have been calculated at 298 K for the gas phase, and also for water solvent, using ab initio molecular orbital theory at the MP2/6-31G* level and the self-consistent reaction field method. All of the stationary points required for an examination of a cyclic (cooperative) mechanism, first proposed by Eigen, have been found. Basis set superposition error has been taken into account, and this has allowed the calculation of the free energy changes associated with the different ways in which CH_2O and $(H_2O)_n$ (n = 1, 2, 3, 4, 8) can reach transition states containing different numbers of water molecules. In the gas phase, a major reaction channel involves the formation of a complex containing three water molecules, which then proceeds to the product. In water solvent, when concentrations and entropic effects associated with the loss of translational and rotational motion are taken into account, 99.9% of the reaction proceeds via this complex, and the experimental pseudo-first-order rate constant for the hydration of formaldehyde in water is reproduced. These findings are consistent with the results of R. P. Bell and co-workers, who concluded that uncatalyzed hydration of a carbonyl group proceeds via a cyclic transition state containing two extra water molecules. Although the process is disfavored entropically, the entropy loss is almost exactly balanced by the gain in enthalpy resulting from more favorable O···H···O hydrogen bonding in an 8-membered ring containing three water molecules than in a 6-membered ring containing two water molecules, as suggested by Gandour. A similar, favorable, hydrogen bonding geometry is present in the cyclic water tetramer. The enthalpy change and solvent isotope effect calculated for the conversion of the complex of formaldehyde with three water molecules to methanediol solvated by two water molecules are in good agreement with the experimental results in water solvent. The different active hydrons of the water molecules of this complex make different normal contributions to the solvent isotope effect, with proton transfer to the adjacent water molecule from the water molecule that forms the C−O bond significantly more advanced than proton transfer to the carbonyl oxygen. Nevertheless, the process is characterized by a non-linear (dome-shaped) proton inventory.

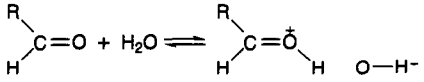
Introduction

Addition to the carbonyl group is a reaction of some importance in organic chemistry[1] and biochemistry.[2] The simplest real example of this reaction is the uncatalyzed addition of water to formaldehyde in water solvent, to form methanediol (eq 1, R = H).[3] At 298 K, the free energy change for this

$$RCH{=}O + H_2O \rightleftharpoons RCH(OH)_2 \qquad (1)$$

reaction is −4.6 kcal/mol,[3a] corresponding to an equilibrium constant of 2000, and the free energy of activation for the addition, estimated from the rate constant,[3b] is 16 kcal/mol. For comparison, the corresponding data for the uncatalyzed addition of water to acetaldehyde (eq 1, R = Me) at 298 K are ΔG = −0.1 kcal/mol, K = 1.2,[4a] and $\Delta G^{\ddagger}$ = 20.7 kcal/mol.[4b] In the case of acetaldehyde, the solvent isotope effect has also been measured; at 293 K, k_{H_2O}/k_{D_2O} is 3.6.[4c]

Since the hydration of a carbonyl group is known to be catalyzed by general acids and bases,[5] the uncatalyzed reactions may, in principle, proceed by proton transfer to the carbonyl oxygen from the general acid, water (eq 2), followed by bond



(2)

formation between the two ions. However, this mechanism presumes that the hydroxide ion of eq 2 can move through the solvent, from the plane of the protonated aldehyde into the perpendicular plane, to reach the carbonyl carbon, prior to the capture of the carbocation by a second water molecule (eq 3). In the latter case, a third molecule of water would then be needed

‡ This work was begun during the tenure of a Canada Council Killam Research Fellowship.

† Present address: Department of Chemistry, Inha University, Inchon 402-751, Republic of Korea.

§ On leave from the Department of Chemical Education, Gyeongsang National University, Chinju 660-701, Republic of Korea.

‖ Present address: Department of Chemistry, University College of the Fraser Valley, Abbotsford, B.C.

⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1995.

(1) Bell, R. P. *Adv. Phys. Org. Chem.* **1966**, *4*, 1−29.

(2) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover Publications, Inc.: New York, 1987.

(3) (a) Zavitsas, A. A.; Coffiner, M.; Wiseman, T.; Zavitsas, L. R. *J. Phys. Chem.* **1970**, *74*, 2746−2750; (b) Calculated by Williams et al. (Williams, I. H.; Spangler, D.; Femec, D. A.; Maggiora, G. M.; Schowen, R. L. *J. Am. Chem. Soc.* **1983**, *105*, 31−40) using a pseudo-first-order rate constant, of 9.2 s^{-1}. This was deduced from the kinetic data of Bell and Evans,[3c] for dehydration of methanediol at 298 K, the $[CH_2O]/[CH_2(OH)_2]$ equilibrium constant, determined by Valenta at 293 K,[3d] and application of the van't Hoff equation to the data of ref 3a. (c) Bell, R. P.; Evans, P. G. *Proc. R. Soc. London, Ser. A.* **1966**, *291*, 297−323; (d) Valenta, P. *Collect. Czech. Chem. Commun.* **1960**, *25*, 853−861.

(4) (a) Bell, R. P.; Rand, M. H.; Wynne-Jones, K. M. A. *Trans. Faraday Soc.* **1956**, *52*, 1093−1102. Critchlow, J. E. *J. Chem. Soc., Faraday Trans. 1* **1972**, *68*, 1774−1992. (b) Based on the rate constant of 4.3 × 10^{-3} s^{-1} given by Critchlow.[4a] (c) Pocker, Y. *Proc. Chem. Soc.* **1960**, 17−18.

(5) Funderburk, L. H.; Aldwin, L.; Jencks, W. P. *J. Am. Chem. Soc.* **1978**, *100*, 5444−5459. Sørensen, P. E.; Jencks, W. P. *J. Am. Chem. Soc.* **1987**, *109*, 4675−4690.

0002-7863/95/1517-4240$09.00/0

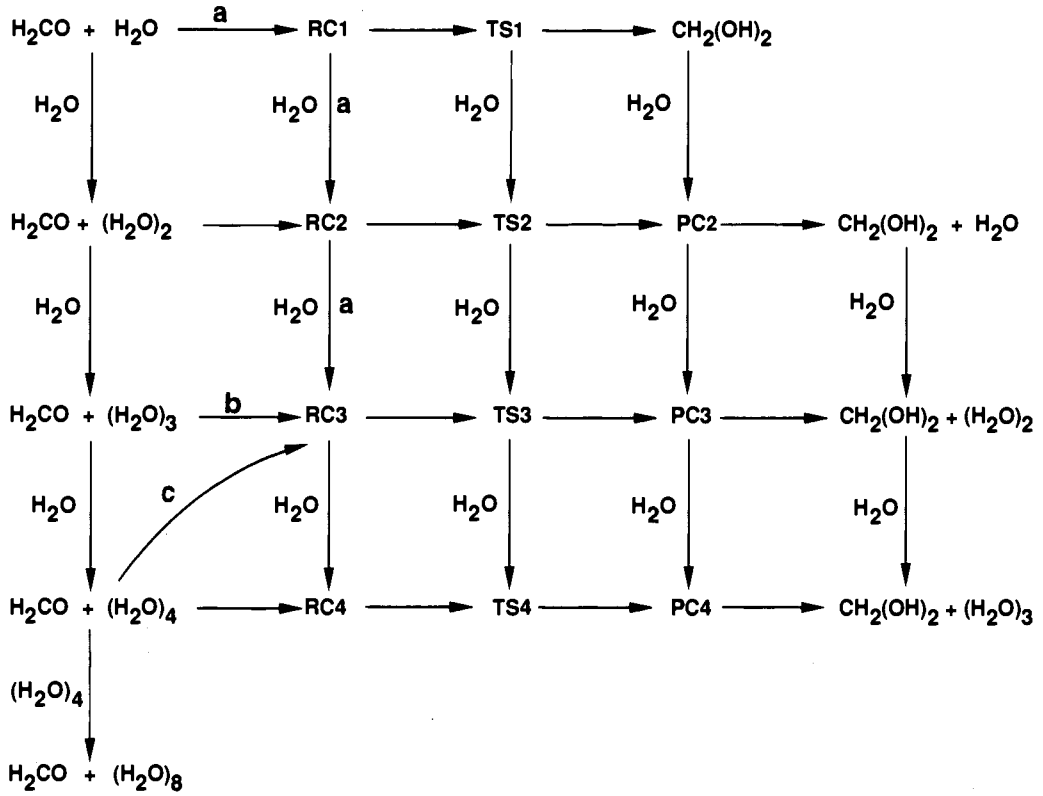
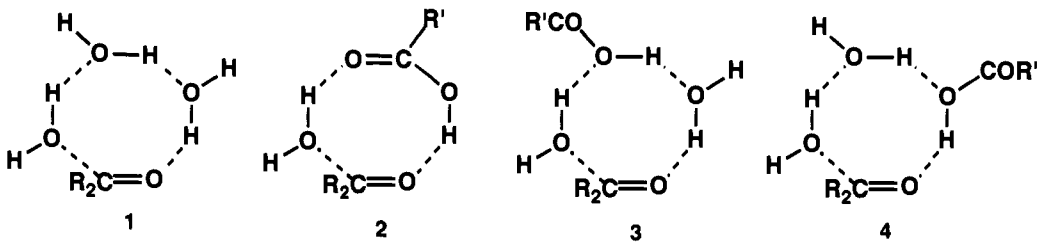


Figure 1. A summary of the processes that describe the hydration of formaldehyde by $(H_2O)_n$ (horizontal arrows) and the stepwise hydration of each of these species (vertical arrows).

Chart 1



(eq 4) to complete the formation of a 1,1-diol.

$$H_2O + RCH{=}\overset{+}{O}H + OH^- \rightleftharpoons H_2\overset{+}{O}{-}CHR{-}OH + OH^- \quad (3)$$

$$H_2O + H_2\overset{+}{O}{-}CHR{-}OH + OH^- \rightleftharpoons H_3O^+ + HO{-}CHR{-}OH + OH^- \quad (4)$$

This reasoning leads to the idea[1,6] that up to two additional water molecules could play an intimate role in the neutral hydration of a carbonyl group and be more than spectators.

In his opening remarks at the 1965 Faraday Society Discussion on Proton Transfer,[6] Eigen commented that stepwise proton transfer mechanisms, e.g., eqs 2–4, require the solvation and desolvation of charged intermediates as these appear and disappear, and should be disfavored with respect to a cyclic, cooperative mechanism.

The hydration of a carbonyl group by a cooperative mechanism involving three water molecules would imply the formation of a cyclic eight-membered reaction complex, e.g., **1**, and three protons moving, more or less synchronously, in the transition state. This proposal was examined experimentally by several groups,[7] especially Bell and his co-workers,[8] who carried out a series of studies on the hydration of dichloroacetone in dioxane. They found the kinetics to be approximately third order in water; with $x_w = 0.1744$, an activation energy of 4.8 kcal/mol and a solvent isotope effect of 4.0 were observed at 298 K. Catalysis by carboxylic acids was also observed and, in the presence of acetic acid or benzoic acid, the kinetics were approximately first order in water and first order in carboxylic acid. For $x_w = 0.1744$, the benzoic acid-catalyzed hydration proceeded with a solvent isotope effect of 2.5 at 298 K, lower than that of the uncatalyzed reaction, and an activation energy of 8.6 kcal/mol, *higher than that of the uncatalyzed reaction.*

It was concluded that the uncatalyzed reaction proceeds via a cyclic transition state containing two extra water molecules. Although proton movement was thought to be partly synchronous with the making or breaking of carbon–oxygen bonds, the experimental data did not reveal whether more than one proton moves simultaneously. For the catalyzed reactions, it was argued initially that the two extra water molecules are

(6) Eigen, M. *Discuss. Faraday Soc.* **1965**, *39*, 7–15.

(7) Dahn, H.; Aubort, J.-D. *Helv. Chim. Acta* **1968**, *51*, 1348–1352. Haldna, U. L.; Erreline, L. E.-J.; Kuura, H. J. *Org. React. (USSR)* **1968**, *5*, 202; Sørensen, P. E. *Acta Chem. Scand., Ser. A* **1976**, *30*, 673–679.

(8) Bell, R. P.; Millington, J. P.; Pink, J. M. *Proc. R. Soc. London, Ser. A* **1968**, *303*, 1–16. Bell, R. P.; Critchlow, J. E. *Proc. R. Soc. London, Ser. A* **1971**, *325*, 35–55. Bell, R. P.; Sörensen, P. E. *J. Chem. Soc., Perkin Trans. 2*, **1972**, 1740–1743.

Table 1. Electronic Energies (E),[a] Zero-Point Energies (Z),[b] and Entropies (S),[c] in the Gas Phase and in Water Solvent, Calculated at Different Levels of Theory for the Network of Reactions Shown in Figure 1

		3-21G[d]		MP2=FC/ 6-31G*//3-21G		MP2=FULL/ 6-31G*[d]
molecule		gas phase	water	gas phase	water	gas phase
H_2O	E	−75.58596	−75.59046	−76.19648	−76.20035	−76.19924
	Z	13.66696	13.67364			13.48158
	S	45.087	45.104			45.133
$(H_2O)_2$	E	−151.18940	−151.20184	−152.40387	−152.41349	−152.41027
	Z	30.31128	30.54866			29.29792
	S	67.179	65.441			69.731
$(H_2O)_3$	E	−226.81434	−226.81870	−228.62421	−228.62909	−228.63575
	Z	49.13660	49.27555			46.95218
	S	73.868	72.732			77.147
$(H_2O)_4$	E	−302.44126	−302.44126	−304.84270	−304.84269	−304.86049
	Z	66.26626	65.79388			63.35010
	S	84.772	86.034			90.338
$(H_2O)_8$	E	−604.94305	−604.94305	−609.72907	−609.72910	−609.76446
	Z	138.07542	137.67542			*e*
	S	114.356	114.688			*e*
CH_2O	E	−113.22182	−113.22558	−114.16682	−114.16966	−114.17496
	Z	18.18158	18.20715			17.13421
	S	52.139	52.146			52.257
RC1	E	−188.82234	−188.83030	−190.37376	−190.37978	−190.38534
	Z	34.75921	34.20013			33.04057
	S	69.785	74.073			71.946
TS1	E	−188.75680	−188.76396	−190.30334	−190.30797	−190.31800
	Z	34.17139	34.28136			32.29882
	S	61.261	61.177			61.629
$CH_2(OH)_2$	E	−188.84694	−188.84695	−190.38613	−190.38616	−190.39761
	Z	38.00068	37.83922			36.68645
	S	62.109	62.392			62.039
RC2	E	−264.43402	−264.43431	−266.58534	−266.58557	−266.60399
	Z	51.72648	51.47621			49.35306
	S	83.545	84.549			86.488
TS2	E	−264.41306	−264.41606	−266.54571	−266.54637	−266.56452
	Z	51.61249	51.22940			48.89158
	S	69.868	71.441			69.363
PC2	E	−264.46168	−264.46344	−266.60140	−266.60257	−266.61715
	Z	55.36682	55.19062			53.10373
	S	74.031	74.336			75.434
RC3	E	−340.05542	−340.05549	−342.80096	−342.80095	−342.82301
	Z	69.56745	69.18236			65.69607
	S	92.590	94.130			100.238
TS3	E	−340.04587	−340.04881	−342.76933	−342.77176	−342.79069
	Z	68.11057	68.15650			63.73529
	S	80.526	81.967			80.286
PC3	E	−340.08425	−340.08623	−342.81870	−342.81931	−342.83978
	Z	72.39751	71.97457			69.29523
	S	86.016	87.857			89.722
RC4	E	−415.67082	−415.67295	−419.01122	−419.01470	−419.03998
	Z	86.13865	85.67703			81.65572
	S	107.857	108.372			117.815
TS4	E	−415.66461	−415.66849	−418.97954	−418.99069	−419.00423
	Z	84.09555	85.52256			78.51227
	S	94.217	97.937			94.037
PC4	E	−415.70115	−415.70312	−419.03074	−419.03155	−419.05745
	Z	88.82877	88.40678			85.08862
	S	101.186	102.042			107.218

[a] In atomic units (1 au = 627.5 kcal/mol). [b] In kcal/mol. [c] Units are cal mol^{-1} K^{-1}. [d] Optimized structure. [e] We were unable to calculate the frequencies of water octamer in this case because the file size exceeded the limit for our Unix system. The free energy change shown in Figure 12 for the dimerization of water tetramer to water octamer was obtained by extrapolation of the linear plot (r = 0.999) of ΔG(MP2=FULL/6-31G*//MP2=FULL/6-31G*) versus ΔG(MP2=FULL/6-31G*//3-21G) for the reactions $(H_2O)_{n-1} + H_2O \rightarrow (H_2O)_n$.

replaced, in the eight-membered ring, by a carboxyl group (i.e., **2**). However, this was later revised to **3**, because the hydrate of the carboxylic acid[9] should be the species which enters into the reaction. Structures **2** and **3** depict two of the ways (**4** is a third) in which a carbonyl group, water, and a carboxyl group might be organized to achieve acid-catalyzed hydration in solution.

The crystal structures of class A β-lactamase enzymes[10a] contain geometrical motifs[10b] that are relevant to the observations and conclusions just discussed. These enzymes, produced by bacteria as a defense against the action of β-lactam antibiotics,[11] have an active site serine residue whose hydroxyl group is acylated by the β-lactam ring; the acyl-enzyme is then

(9) Grunwald, E.; Jumper, C. F.; Meiboom, S. *J. Am. Chem. Soc.* **1963**, *85*, 522−528. Grunwald, E.; Meiboom, S. *J. Am. Chem. Soc.* **1963**, *85*, 2047−2050. Luz., Z.; Meiboom, S. *J. Am. Chem. Soc.* **1963**, *85*, 3923.

(10) (a) β-Lactamases are divided into four classes (A, B, C, and D) on the basis of their structures and specificities. The class A are the most commonly encountered. See, e.g.: Ambler, R. P.; Coulson, A. F.; Frère, J.-M.; Ghuysen, J.-M.; Forsman, M.; Joris, B.; Levesque, R.; Tiraby, G.; Waley, S. G. *Biochem. J.* **1991**, *276*, 269−270. Joris, B.; Ledent, P.; Dideberg, O.; Fonzé, E.; Lamotte-Brasseur, J.; Kelly, J. A.; Ghuysen, J.-M.; Frère, J.-M. *Antimicrob. Agents Chemother.* **1991**, *35*, 2294−2301. (b) For the crystal structure of a β-lactamase from *Staphylococcus aureus*, see: Herzberg, O. *J. Mol. Biol.* **1991**, *217*, 701−719. For the β-lactamase from *Bacillus licheniformis*, see: Knox, J. R.; Moews, P. C. *J. Mol. Biol.* **1991**, *220*, 435−455. For the crystal structure of the RTEM-1 enzyme from *Escherichia coli*, see: Jelsch, C.; Lenfant, F.; Masson, J. M.; Samama, J. P. *FEBS Lett.* **1992**, *299*, 135−142. Strydnadka, N. C. J.; Adachi, H.; Jensen, S. E.; Johns, K.; Slelecki, A.; Betzel, C.; Sutoh, K.; James, M. N. G. *Nature* **1992**, *359*, 700−705.

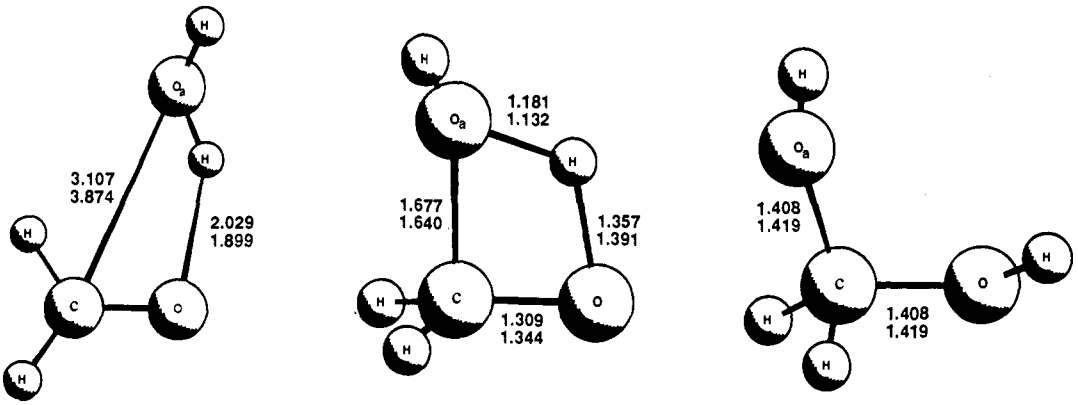


Figure 2. The structures of $(RC)_1$ (left), $(TS)_1$ (center), and $(PC)_1$ (right). Bond lengths refer to the gas phase (upper) and to water (lower). The imaginary frequencies of the transition states are 1753i (gas phase) and 1624i (water).

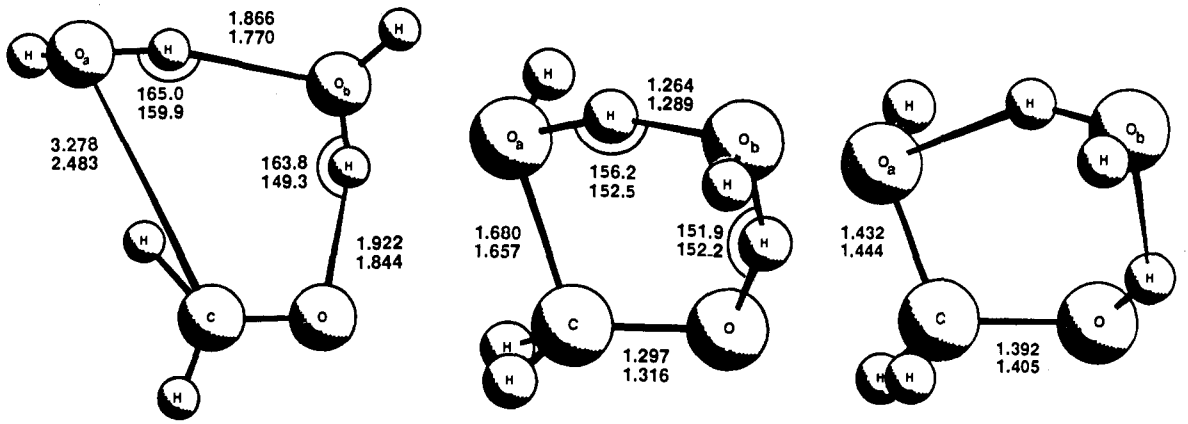


Figure 3. The structures of $(RC)_2$ (left), $(TS)_2$ (center), and $(PC)_2$ (right). Bond lengths and bond angles refer to the gas phase (upper) and to water (lower). The imaginary frequencies of the transition states are 1450i (gas phase) and 961i (water).

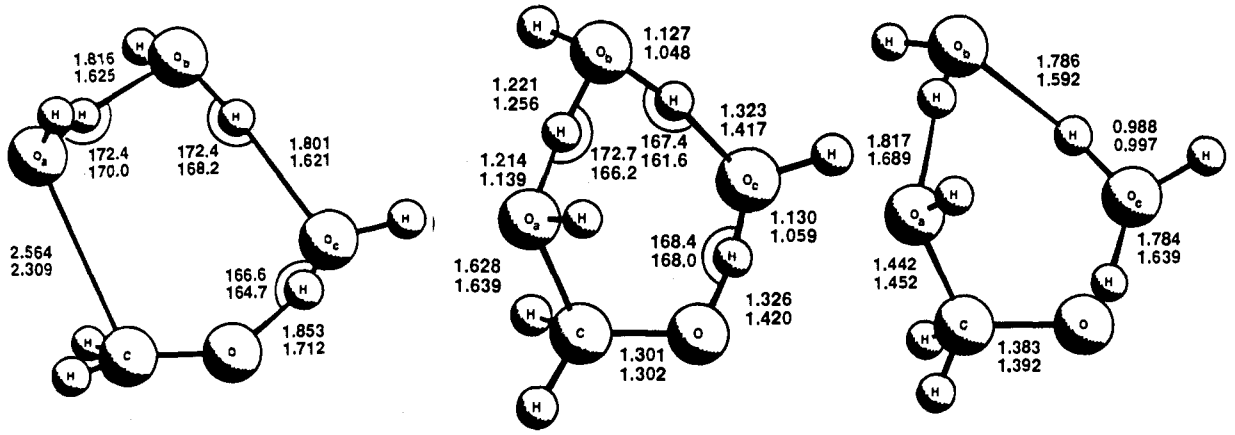


Figure 4. The structures of $(RC)_3$ (left), $(TS)_3$ (center), and $(PC)_3$ (right). Bond lengths and bond angles refer to the gas phase (upper) and to water (lower). The imaginary frequencies of the transition states are 1290i (gas phase) and 457i (water).

hydrolyzed. In the crystal structures there is a conserved geometrical relationship, reminiscent of **2**, between the serine hydroxyl group, *one* water molecule, and the carboxyl group of a glutamic acid residue that has a significant catalytic effect upon the deacylation reaction. The replacement of this residue by alanine does not affect the acylation step, but the rate of deacylation is decreased one-million-fold.[12] These observations suggest that carboxylic acid catalysis of the addition of water to a carbonyl group could be much more pronounced than in the experiments of Bell and co-workers, if the carbonyl group, the carboxyl group, and one water molecule were constrained to a cyclic structure such as **2**.

The foregoing considerations raise several questions: (i) Is a cooperative three-water molecule hydration mechanism uniquely compatible with the available kinetic, thermodynamic, and solvent isotope effect data? (ii) What is the synchroneity of multiple proton transfer? (iii) What are the structures of the reaction complexes and transition states in the cooperative mechanism? (iv) How is carboxylic acid catalysis of this mechanism manifested? (v) In particular, if the carboxylic acid-catalyzed hydration of a carbonyl group proceeds via complex **3** in a solvent, how would the enforcement of structure **2** by an enzyme affect the kinetics?

(11) Abraham, E. P.; Chain, E. *Nature* **1940**, *146*, 837. Kelly, J.; Dideberg, O.; Charlier, P.; Wéry, J.; Libert, M.; Moews, P.; Knox, J.; Duez, C.; Fraipont, C.; Joris, B.; Dusart, D.; Frère, J.-M.; Ghuysen, J.-M. *Science* **1986**, *231*, 1429–1431. Abraham, E. P. *BioEssays* **1990**, *12*, 601–606. Abraham, E. P. *J. Chemother.* **1991**, *3*, 67–74.

(12) Escobar, W. A.; Tan, A. K.; Fink, A. L. *Biochemistry* **1991**, *30*, 10783–10787.

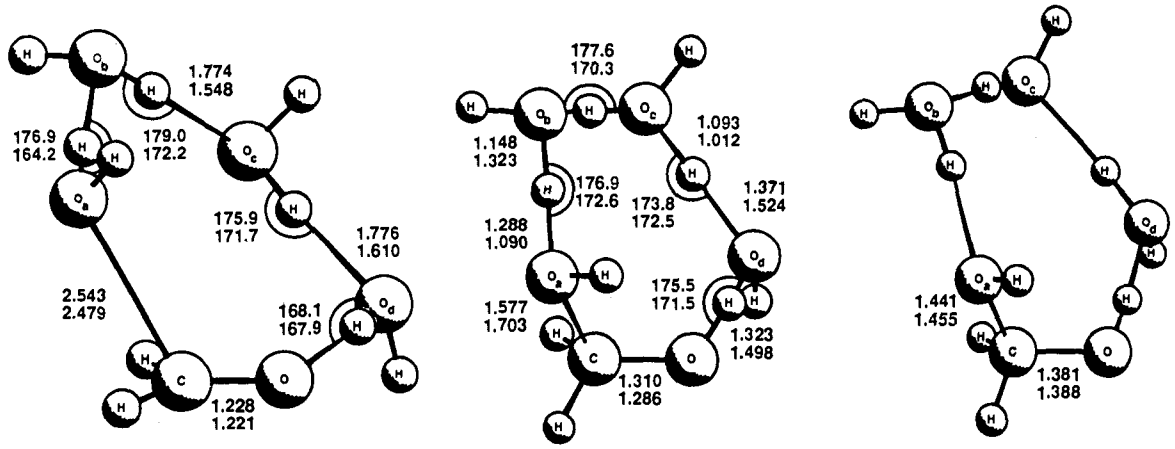


Figure 5. The structures of $(RC)_4$ (left), $(TS)_4$ (center), and $(PC)_4$ (right). Bond lengths and bond angles refer to the gas phase (upper) and to water (lower). The imaginary frequencies of the transition states are 1115i (gas phase) and 493i (water).

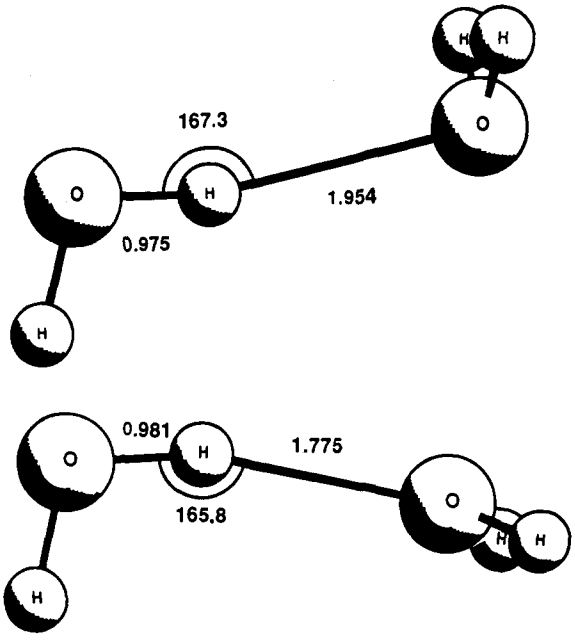


Figure 6. Calculated structure of water dimer in the gas phase (upper) and in water (lower). Geometrical parameters refer to optimization at MP2=FULL/6-31G* (upper) and 3-21G(SCRF) (lower).

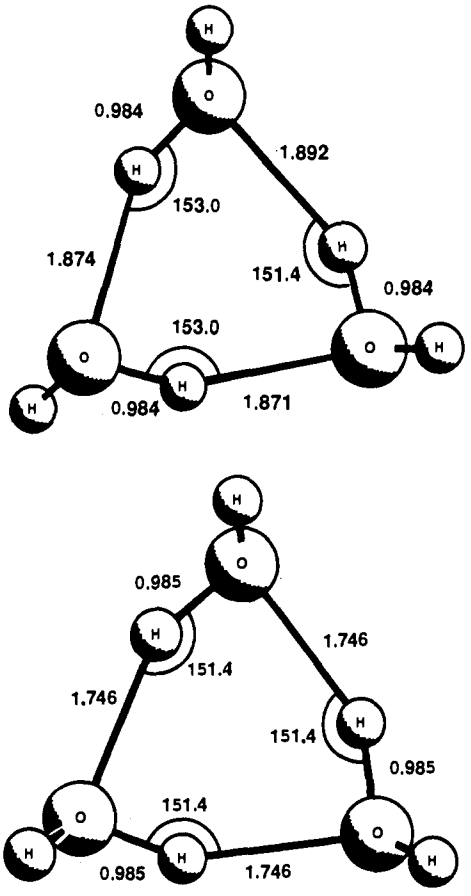


Figure 7. Calculated structure of water trimer in the gas phase (Figure 7a, upper) and in water (Figure 7b, lower). Geometrical parameters refer to optimization at MP2=FULL/6-31G* (upper) and 3-21G(SCRF) (lower).

We have used ab initio molecular orbital theory to examine these questions. The present paper, which is the first of a series, is concerned with the neutral hydration of formaldehyde. We have also examined the nature of carboxylic acid catalysis, rate–equilibrium relationships in acid catalysis, and structure–reactivity relationships in the hydration of formaldehyde, acetaldehyde, and substituted acetaldehydes. These studies will be reported in subsequent papers.

There have been many previous calculations on the energetics of hydration of formaldehyde by one or two water molecules.[3b,13] As stated above, our objectives were to find a *three*-water molecule reaction, to compute the thermochemical quantities, solvent isotope effects, and proton inventories[14] associated with this mechanism and, at the highest level of theory available to us, to compare the results to those calculated for the one- and two-water molecule mechanisms. For processes termed "gas phase", this level of theory was MP2=FULL/6-31G*,[15] i.e., a second-order Möller–Plesset treatment using all of the orbitals generated by the 6-31G* basis set, for the optimization of all structures and the calculations of the frequencies used to obtain the thermochemical quantities and isotope effects.

While the work was in progress, new studies were published on the constitution of water[16] and, in particular, the characteriza-

(13) Madura, J. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 2517–2527. Ventura, O. N.; Coltiño, E. L.; Irving, K.; Iglesias, A. *J. Mol. Struct. (Theochem)* **1990**, *210*, 427–440. Peeters, D.; Leroy, G. *Can. J. Chem.* **1991**, *69*, 1376–1387. Tachibana, A.; Fueno, H.; Yamato, M.; Yamabe, T. *Int. J. Quantum Chem.* **1991**, *40*, 435–456. Ventura, O. N.; Apostolova, E. S. *J. Org. Chem. USSR* **1991**, *27*, 573–579. Coltiño, E. L.; Lledós, A.; Bertran, J. *J. Comput. Chem.* **1992**, *13*, 1037–1046. For related work, see: Kozmutza, C.; Ozoróczy, Zs.; Kapuy, E. *J. Mol. Struct. (Theochem)* **1990**, *204*, 101–109.

(14) Venkatasubban, K. S.; Schowen, R. L. *CRC Crit. Rev. Biochem.* **1981**, *17*, 1–44.

(15) Hehre, W. J.; Radom, L.; Schleyer, P. v R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.

(16) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Clarendon Press: Oxford, 1969; Franks, F., Ed. *Water, A Comprehensive Treatise*; Plenum Press: New York; Vols. 1–7.

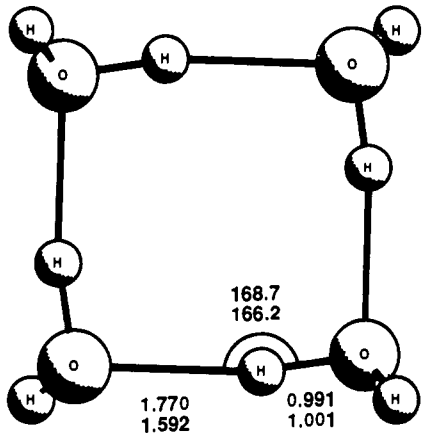


Figure 8. Calculated structure of water tetramer. Geometrical parameters refer to optimization at MP2=FULL/6-31G* (upper) and at 3-21G(SCRF) (lower).

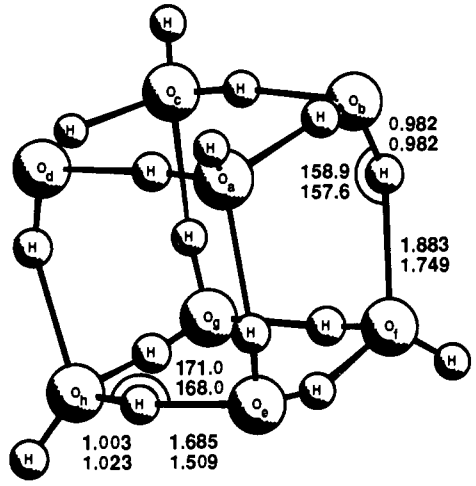


Figure 9. Calculated structure of water octamer. Geometrical parameters refer to optimization at MP2=FULL/6-31G* (upper) and at 3-21G(SCRF) (lower).

tion[17] and computation[18] of different water clusters. These led us to extend our work to the search for a four-water molecule mechanism and to the roles played by water dimer,[18a] cyclic water trimer,[18b] cyclic water tetramer,[18c] and cubic water octamer.[18d]

Figure 1 summarizes the network of reactions that was eventually calculated at MP2=FULL/6-31G*. A left-to-right movement on this figure describes, successively, the organization of formaldehyde and a water cluster $(H_2O)_n$ into a reactant complex $(RC)_n$, which is transformed, over a barrier $(TS)_n$, to a product complex $(PC)_n$, and then to methanediol and $(H_2O)_{n-1}$. Vertical movement on Figure 1 refers to the stepwise hydration or dehydration of each species. Where E is an energy (enthalpy, H, or free energy, G), for each value of n the quantity $[E(CH_2O) + E(H_2O)_n - E(RC)_n]$ is the energy required to assemble the reactant complex, $[E(PC)_n - E(RC)_n]$ is ΔE, the energy change for the reaction, and $[E(TS)_n - E(RC)_n]$ is $\Delta E^{\ddagger}$, the reaction barrier.

To illustrate the use of Figure 1, the reactant complex for $n = 3$ might be formed by the successive attachment of three water molecules to formaldehyde (Path a), the insertion of formaldehyde into the cyclic water trimer (Path b), the displacement of one water molecule from the cyclic water tetramer (Path c), or the displacement of water from a corner of the cubic octamer and dissociation into $(RC)_3$, water, and water tetramer. The energy changes for all of these processes were calculated.

These energies and structures pertain to the gas phase and, perhaps, to the mainly hydrophobic environment at the active site of an enzyme. However, as discussed earlier, the experimental data to which the computational results have to be compared refer mainly to water solvent or to mixed solvents.

There are two conceptually different ways to treat solvation. The more rigorous has been thought to be one which treats solvent molecules explicitly. The molecule or reaction is placed in a box of several hundred solvent molecules, and molecular dynamics or Monte Carlo statistical mechanics simulations are carried out on the resulting system. This approach has been very successful,[19] but it is computationally very demanding, and would not have been feasible for the treatment of the full network of reactions shown in Figure 1. A second approach to solvation is based on the classical reaction field models of Born,[20] Kirkwood,[21] and Onsager.[22] In the Onsager treatment the system is placed in a spherical cavity immersed in a continuous medium of dielectric constant ϵ. If the structure possesses a permanent dipole moment, a dipole moment is induced in the solvent, and the interaction between the two stabilizes the system. The resulting energy can be calculated to self-consistency, hence Self Consistent Reaction Field (SCRF).[23]

If SCRF theory could be implemented, the computational cost of this procedure would be negligible, but there has been

(17) Dimer: Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*, 498–510. Odutola, J. A.; Hu, T. A.; Prinslow, D.; O'Dell, S. E.; Dyke, T. R. *J. Chem. Phys.* **1988**, *88*, 5352–5361. Slanina, Z. *Ber. Bunsenenges. Phys. Chem.* **1993**, *97*, 558–561. Pugliano, N.; Saykally, R. J. *J. Chem. Phys.* **1992**, *96*, 1832–1839. Cyclic trimer: Pugliano, N.; Saykally, R. J. *Science* **1992**, *257*, 1937–1940. Saykally, R. J.; Blake, G. A. *Science* **1993**, *259*, 1570–1575. Cyclic tetramer and cubic octamer: Benson, S. W.; Siebert, E. D. *J. Am. Chem. Soc.* **1992**, *114*, 4269–4276. Bertagnolli, H. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 1577–1578.

(18) (a) Dimer: Slanina, Z. *Chem. Phys.* **1991**, *150*, 321–329. Bertran, J.; Ruiz-López, M. F.; Rinaldi, D.; Rivail, J. L. *Theor. Chim. Acta* **1992**, *84*, 181–194. Millot, C.; Stone, A. J. *Mol. Phys.* **1992**, *77*, 439–462. (b) Cyclic trimer: Mó, O.; Yáñez, M.; Elguero, J. *J. Chem. Phys.* **1992**, *97*, 6628–6638. Vegiri, A.; Farantos, S. C. *J. Chem. Phys.* **1993**, *98*, 4059–4075. Xantheas, S. S.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 8037–8040. Wales, D. J. *J. Am. Chem. Soc.* **1993**, *115*, 11180–11190. van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *Chem. Phys.* **1993**, *175*, 271–281. (c) Cyclic tetramer: Koehler, J. E. H.; Saenger, W.; Lesyng, B. *J. Comput. Chem.* **1987**, *8*, 1090–1098. Ferrari, A. M.; Garrone, E.; Ugliengo, P. *Chem. Phys. Lett.* **1993**, *212*, 644–648 and references cited therein. (d) Octamer: Coronglu, G.; Clementi, E. *Chem. Phys. Lett.* **1993**, *214*, 367–372. Kim, J.; Mhin, B. J.; Lee, S. J.; Kim, K. S. *Chem. Phys. Lett.* **1994**, *219*, 243–246 and references cited therein.

(19) Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304–5314. Dean, P. M. *Molecular Foundations of Drug-Receptor Interaction*; Cambridge University Press: Cambridge, UK, 1987; pp 211f. Jorgensen, W. L. *Chemtracts: Org. Chem.* **1991**, *4*, 91–119 and references cited.

(20) Born, M. *Z. Phys.* **1920**, *1*, 45–48. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

(21) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351–361.

(22) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.

(23) Tapia, O.; Goscinski, O. *Mol. Phys.* **1975**, *29*, 1653–1661. Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129. Pascual-Ahuir, J. L.; Silla, E.; Tomasi, J.; Bonaccorsi, R. *J. Comp. Chem.* **1987**, *8*, 778–787. Bianco, R.; Miertuš, S.; Persico, M.; Tomasi, J. *Chem. Phys.* **1992**, *168*, 281–292. Milne, G. W. A.; Nicklaus, M. C.; Hodoš ek, M. *J. Mol. Struct.* **1993**, *291*, 89–103. Olivares del Valle, F. J.; Aguilar, M. A.; Contador, J. C. *Chem. Phys.* **1993**, *170*, 161–165. Aguilar, M.; Bianco, R.; Miertus, S.; Persico, M.; Tomasi, J. *Chem. Phys.* **1993**, *174*, 397–407. Mingos, D. M. P.; Rohl, A. L.; Burgess, J. *J. Chem. Soc., Dalton Trans.* **1993**, 423–426. Honig, B.; Sharp, K.; Yang, A.-S. *J. Phys. Chem.* **1993**, *97*, 1101–1109. Kim, H. J.; Bianco, R.; Gertner, B. J.; Hynes, J. T. *J. Phys. Chem.* **1993**, *97*, 1723–1728. Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 446–454.

(24) (a) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Chem. Phys.* **1991**, *95*, 8991–8998. Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776–4782. (b) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, S.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; DeFrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92, Revision B*; Gaussian, Inc.: Pittsburgh, PA, 1992.

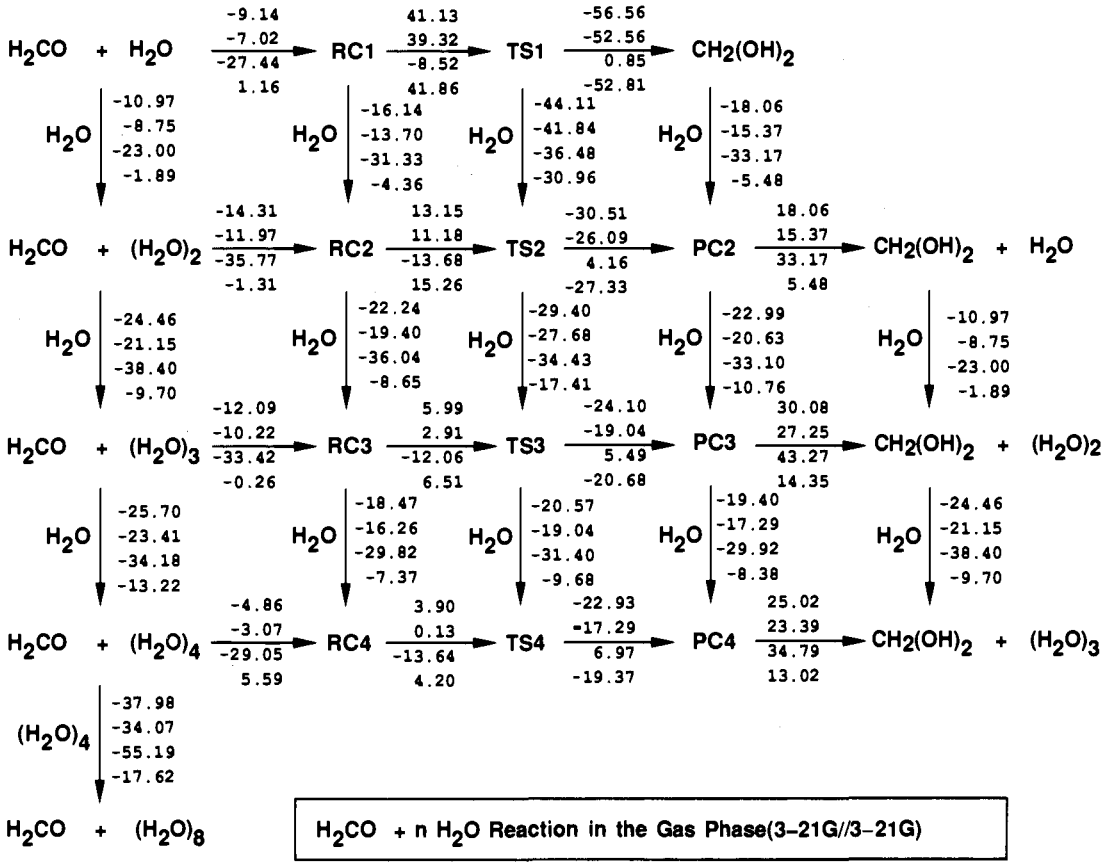


Figure 10. Details of Figure 1 with all structures and thermochemical parameters calculated using 3-21G. The numbers shown beside each arrow refer, from top to bottom, to ΔE, ΔH, ΔS, and ΔG. The temperature is 0 K for ΔE and 298 K for the other parameters.

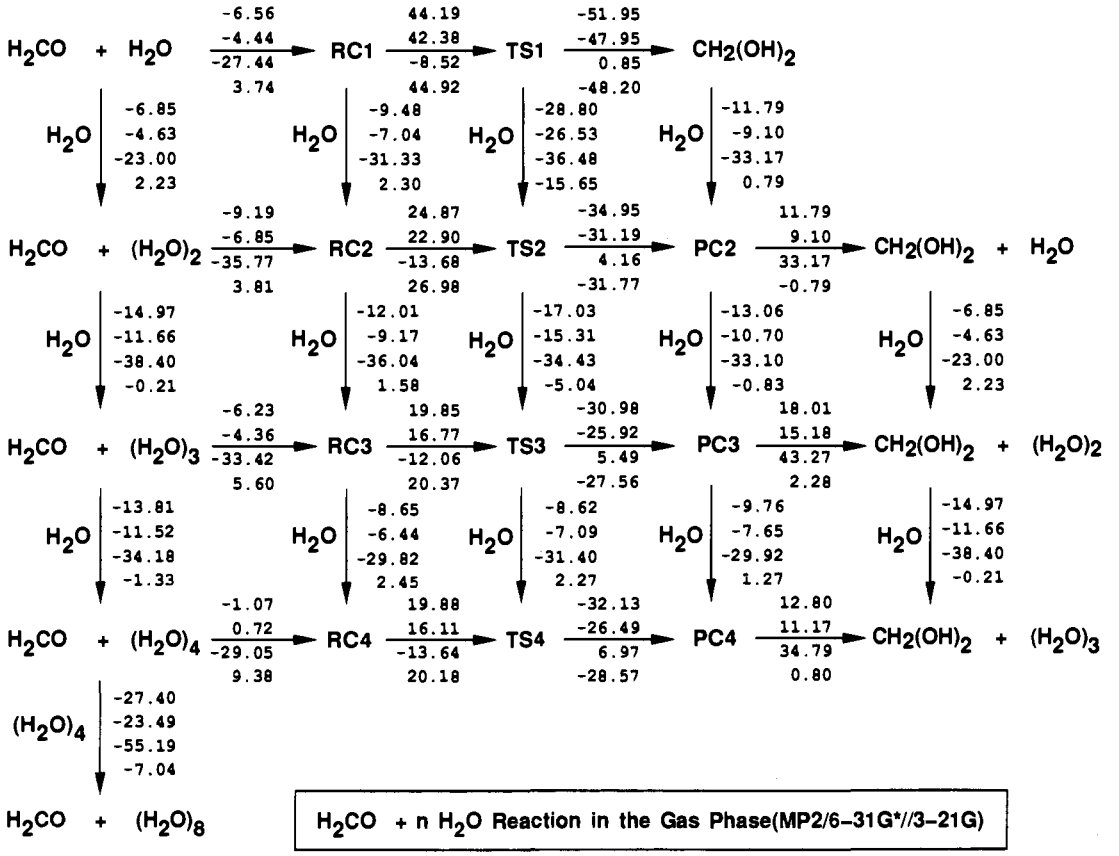


Figure 11. Details of Figure 1 with structures and vibrational frequencies optimized at 3-21G and energies calculated at MP2=FC/6-31G*. The ordering of the numbers is the same as in Figure 10.

uncertainty regarding how to specify the radius of the sphere and how this radius varies along a reaction coordinate. Solutions to these questions have recently been proposed by Wong, Wiberg, and Frisch[24a] and incorporated into the latest versions

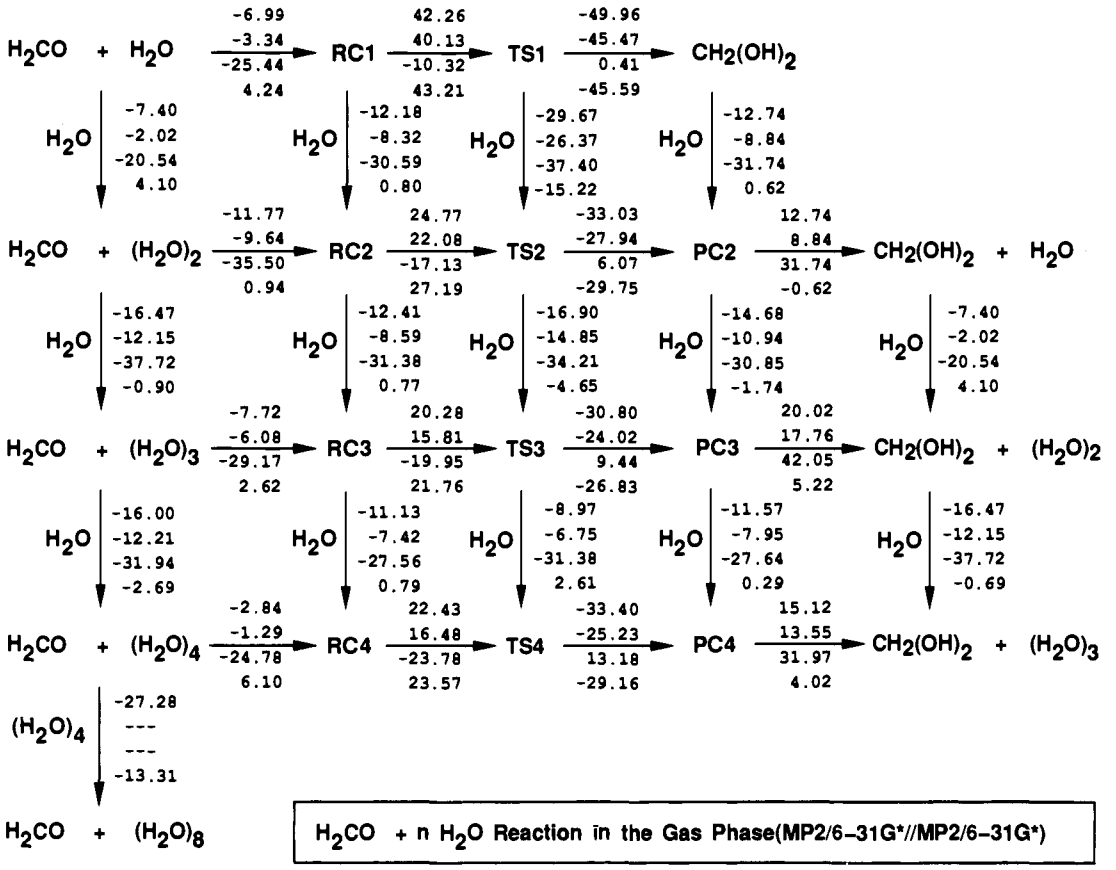


Figure 12. Details of Figure 1 with all structures and thermochemical parameters calculated at MP2=FULL/6-31G*. The ordering of the numbers is the same as in Figure 10.

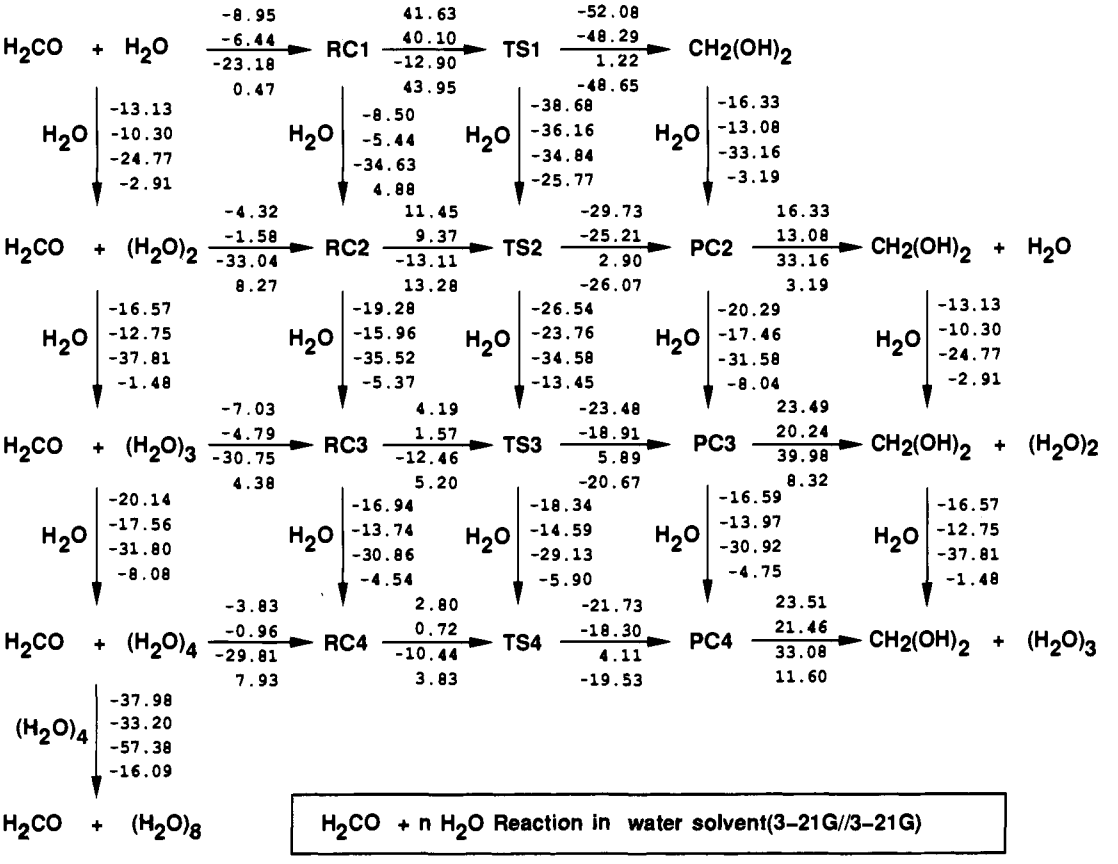


Figure 13. Details of Figure 1 with all structures and thermochemical parameters calculated at 3-21G(SCRF). The ordering of the numbers is the same as in Figure 10.

of GAUSSIAN.[24b] Numerous applications of the ab initio SCRF treatment have been published,[25] including side-by-side comparisons with the Jorgensen procedure.[26] The results are encouraging, and we proposed to use the SCRF procedure as

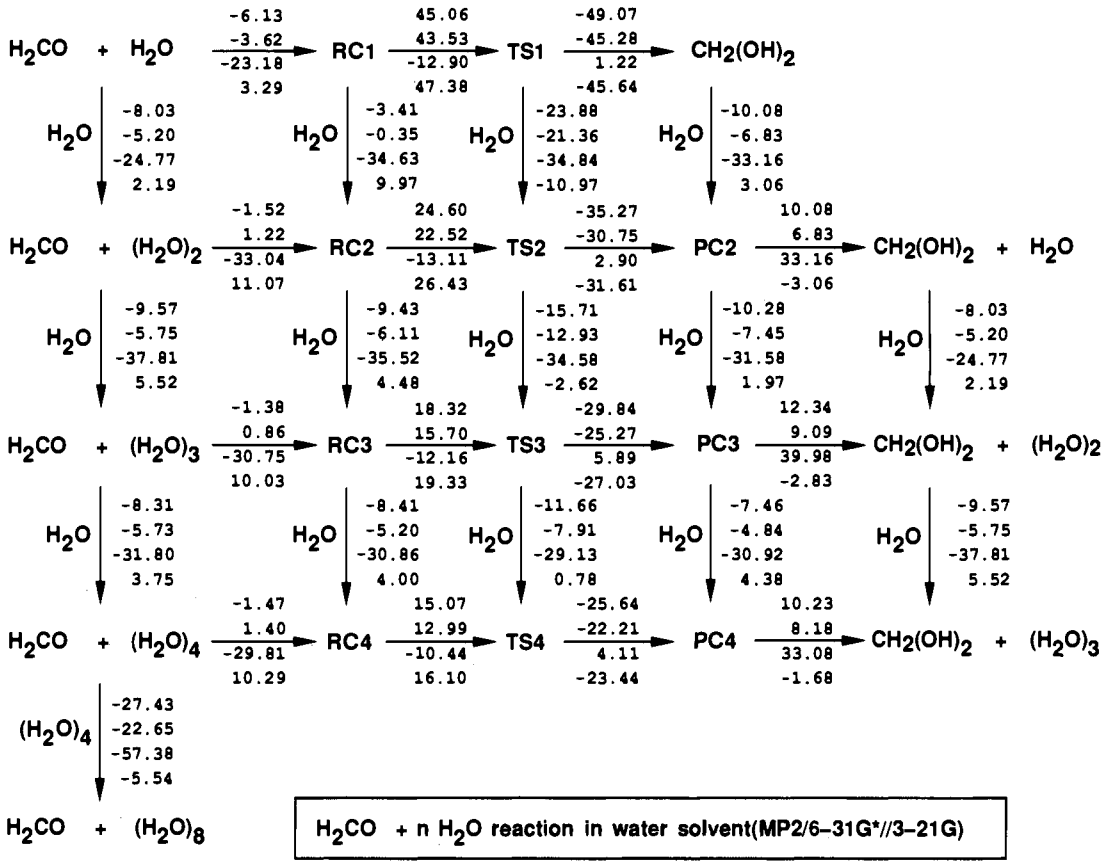


Figure 14. Details of Figure 1 with structures and vibrational frequencies calculated at 3-21G(SCRF) and energies calculated at MP2=FC/6-31G*(SCRF). The ordering of the numbers is the same as in Figure 10.

implemented in Gaussian 92[24b] to extend Figure 1 from the gas phase to water solvent.

Only single-point SCRF calculations are currently possible at the MP2 level. These were performed, with the 6-31G* basis set, on SCRF-optimized 3-21G[27] structures, with exclusion of the inner-shell orbitals, i.e., MP2=FC/6-31G*(SCRF)//3-21G-(SCRF) (FC = frozen core). The designation "water solvent" refers to energies computed in this way and geometries, frequencies, and isotope effects computed at 3-21G(SCRF). To check whether this strategy degrades the proposed comparisons between gas-phase and water solvent, we examined five variations of Figure 1: 3-21G//3-21G, MP2=FULL/6-31G*// MP2=FULL/6-31G*, MP2=FC/6-31G*//3-21G, 3-21G(SCRF)/ /3-21G(SCRF), and MP2=FC/6-31G*(SCRF)//3-21G(SCRF).

Methods

Transition structures $(TS)_n$ were located by the reaction coordinate method. Using $(RC)_n$ as a starting point, trial values of the forming O···C=O and C=O···H bond lengths were selected, all other geometrical parameters were optimized, and the magnitudes and signs of the gradients of the two reaction coordinates were used to improve the trial values. After several such calculations, a structure close to the saddle point was reached; at this stage, a final refinement was achieved using the OPT=CALCALL option, in which force constants are calculated in every iteration. All transition structures exhibited one imaginary frequency.

For SCRF calculations, the previously optimized gas-phase structures were used to estimate the radius of the cavity, and this value was employed in the optimization. A dielectric constant of 78.5 was used to describe water solvent.

Isotope effects were calculated with the Bigeleisen equation[28a] and refer to 298 K, unless stated otherwise. The frequencies employed as input to the Bigeleisen equation were obtained by scaling of the computed values.[28b,c,d] Standard scaling factors were used.[28e,f] Thermodynamic properties were calculated from the frequencies by standard procedures.[28g] In the gas phase, the standard state is 1 atm; in water solvent the standard state is 1 M for formaldehyde and 55.6 M for water (*cf.* eq 6).

In a mixture of HOH and DOD, the rate of the reaction between formaldehyde and one water molecule, which has one active hydron, one passive hydron, and two exchangeable sites, can be written[14]

$$V_n = k_{HH}c_{HOH} + k_{DH}c_{DOH} + k_{HD}c_{HOD} + k_{DD}c_{DOD} \qquad (5)$$

where c is the concentration of a species having isotopes i and j, k_{ij} is a rate constant associated with this species, and n is the atom fraction of deuterium. In a solvent containing no DOD, the rate can be expressed

(25) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Am. Chem. Soc.* **1992**, *114*, 1645–1652. Cieplak, A. S.; Wiberg, K. B. *J. Am. Chem. Soc.* **1992**, *114*, 9226–9227. Parchment, O. G.; Hillier, I. H.; Green, D. V. S.; Burton, N. A.; Morley, J. O.; Schaefer, H. F., III *J. Chem. Soc., Perkin Trans. 2* **1992**, 1681–1684. McDouall, J. J. W. *J. Org. Chem.* **1992**, *57*, 2861–2864. Wiberg, K. B.; Wong, M. W. *J. Am. Chem. Soc.* **1993**, *115*, 1078–1084. Tuñón, I.; Silla, E.; Tomasi, J. *J. Phys. Chem.* **1992**, *96*, 9043–9048. Parchment, O. G.; Green, D. V. S.; Taylor, P. J.; Hillier, I. H. *J. Am. Chem. Soc.* **1993**, *115*, 2352–2356. Wong, M. W.; Leung-Toung, R.; Wentrup, C. *J. Am. Chem. Soc.* **1993**, *115*, 2465–2472. Luque, F. J. *J. Phys. Chem.* **1993**, *97*, 9380–9384. Olivares del Valle, F. J.; Aguilar, M. A.; Tolosa, S. *J. Mol. Struct. (Theochem)* **1993**, *279*, 223–231. Young, P. E.; Hillier, I. H. *Chem. Phys. Lett.* **1993**, *215*, 405–408. Gould, I. R.; Hillier, I. H. *J. Chem. Soc., Perkin Trans 2* **1993**, 1771–1773. Tortonda, F. R.; Pascual-Ahuir, J.-L.; Silla, E.; Tuñón, I. *J. Phys. Chem.* **1993**, *97*, 11087–11091.

(26) Orozco, M.; Jorgensen, W. L.; Luque, F. J. *J. Comput. Chem.* **1993**, *14*, 1498–1503. Wiberg, K. B.; Marquez, M. *J. Am. Chem. Soc.* **1994**, *116*, 2197–2198.

(27) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

(28) (a) Bigeleisen, J. *J. Chem. Phys.* **1949**, *17*, 675–678. (b) Wolfe, S.; Hoz, S.; Kim, C.-K.; Yang, K. *J. Am. Chem. Soc.* **1990**, *112*, 4186–4191. (c) Wolfe, S.; Kim, C.-K. *J. Am. Chem. Soc.* **1991**, *113*, 8056–8061. (d) Boyd, R. J.; Kim, C.-K.; Shi, Z.; Weinberg, N.; Wolfe, S. *J. Am. Chem. Soc.* **1993**, *115*, 10147–10152. (e) Aljibury, A. L.; Snyder, R. G.; Strauss, H. L.; Raghavachari, K. *J. Chem. Phys.* **1986**, *84*, 6872–6878. (f) Pupyshev, V. I.; Panchenko, Y. N.; Bock, C. W.; Pongor, G. *J. Chem. Phys.* **1991**, *94*, 1247–1252. (g) Shaik, S. S.; Schlegel, H. B. Wolfe, S. *Theoretical Aspects of Physical Organic Chemistry. The S_N2 Mechanism*; Wiley-Interscience: New York, 1992; pp 80–84.

in terms of the total concentration of water,

$$V_0 = k_{HH}c^{\circ}_{HOH} \quad (6)$$

The equilibrium constant for the reaction between HOH and DOD (eq 7) is

$$HOH + DOD \rightleftharpoons 2HOD \quad (7)$$

4.00,[29] i.e.,

$$K = [HOD]^2/[HOH][DOD] = x_{HOD}^2/x_{HOH}x_{DOD} = 4 \quad (8)$$

The solvent molecules, N_w, are distributed as $N_w = N_{HOH} + N_{DOD} + N_{HOD}$. Since each of these contains two hydrons, we write

$$N_{H+D} = 2N_w \quad (9)$$

where N_{H+D} is the total number of hydrogen and deuterium atoms. The number of hydrogen atoms, N_H, is

$$N_H = 2N_{HOH} + N_{HOD} = 2x_{HOH}N_w + x_{HOD}N_w = (2x_{HOH} + x_{HOD})N_w \quad (10)$$

From eq 9 and eq 10,

$$N_H = 0.5(2x_{HOH} + x_{HOD})N_{H+D}$$

$$N_H/N_{H+D} = n_H = x_{HOH} + 0.5x_{HOD} \quad (11)$$

where n_H is the atom fraction of hydrogen. Similarly, for the atom fraction of deuterium,

$$N_D/N_{H+D} = n_D = x_{DOD} + 0.5x_{HOD} \quad (12)$$

Summarizing, with $n_D = n$ and $n_H = 1 - n$,

$$x_{HOD}^2 = 4x_{HOH}x_{DOD} \quad (13)$$

$$1 - n = x_{HOH} + 0.5x_{HOD} \quad (14)$$

$$n = x_{DOD} + 0.5x_{HOD} \quad (15)$$

From eq 14 minus eq 15,

$$1 - 2n = x_{HOH} - x_{DOD}$$

$$x_{HOH} = 1 - 2n + x_{DOD} \quad (16)$$

From eq 14 plus eq 15,

$$1 = x_{HOH} + x_{DOD} + x_{HOD} = 1 - 2n + x_{DOD} + x_{DOD} + x_{HOD}$$

$$x_{HOD} = 2n - 2x_{DOD} \quad (17)$$

From eqs 13, 16, and 17,

$$(2n - 2x_{DOD})^2 = 4x_{DOD}(1 - 2n + x_{DOD})$$

$$x_{DOD} = n^2 \quad (18)$$

$$x_{HOH} = 1 - 2n + n^2 = (1 - n)^2 \quad (19)$$

$$x_{HOD} = 2n(1 - n) \quad (20)$$

From the ratio (eq 5)/(eq 6), the relative velocity in a mixture of HOH and DOD is

$$V_n/V_0 = \frac{c_{HOH}}{c^{\circ}_{HOH}} + \frac{k_{DH}}{k_{HH}}\frac{c_{HOD}}{c^{\circ}_{HOH}} + \frac{k_{HD}}{k_{HH}}\frac{c_{DOH}}{c^{\circ}_{HOH}} + \frac{k_{DD}}{k_{HH}}\frac{c_{DOD}}{c^{\circ}_{HOH}} \quad (21)$$

Since $c_{HOD} = c_{DOH}$, eqs 18–21 lead to

$$V_n/V_0 = x_{HOH} + (k_{DH}/k_{HH} + k_{HD}/k_{HH})x_{HOD} + x_{DOD}k_{DD}/k_{HH}$$
$$= (1 - n)^2 + 2n(1 - n)(k_{DH}/k_{HH} + k_{HD}/k_{HH}) + n^2k_{DD}/k_{HH} \quad (22)$$

Taking symmetry into account, i.e., HOH = 2, DOD = 2, HOD = 1, leads to eq 23

$$V_n/V_0 = (1 - n)^2 + n(1 - n)(k_{DH}/k_{HH} + k_{HD}/k_{HH}) + n^2k_{DD}/k_{HH} \quad (23)$$

which relates the isotope effects to the solvent composition in the one-water molecule reaction.[14] The gas-phase isotope effects computed at MP2=FULL/6-31G*//MP2=FULL/6-31G* are $k_{HH}/k_{DD} = 4.18$ and $k_{HH}/k_{HD} = 4.36$ for the active hydron and $k_{HH}/k_{DH} = 0.94$ for the passive hydron. Insertion of the inverse of these values into eq 23 and variation of n from 0.0 to 1.0 gives the gas-phase proton inventory for this reaction.

Three isotope effects have to be computed to obtain this proton inventory. For the proton inventory of the two-water molecule reaction, 15 isotope effects are needed; 63 isotope effects are required to obtain the proton inventory of the three-water molecule reaction; and the inventory of the four-water molecule reaction requires the computation of 255 isotope effects. These were inserted into (the more complex) equations, derived in the same manner as eq 23, to obtain the results described later.

Results and Discussion

I. Structures and Energies. Table 1 collects the electronic energies, zero-point energies, and entropies of all structures calculated during this work. Figures 2–5 show, from left to right in each case, the reactant complex $(RC)_n$, transition structure $(TS)_n$, and product complex $(PC)_n$ for, respectively, the one-, two-, three-, and four-water molecule reactions. The bond lengths and bond angles shown on these structures refer to MP2=FULL/6-31G* optimization (upper; gas phase) and 3-21G(SCRF) optimization (lower; water). Figures 6–9 give the structures calculated in the gas phase and in water for water dimer (C_s symmetry),[18a] trimer,[18b] tetramer (S_4 symmetry),[18c] and octamer (D_2 symmetry).[18d] In the case of water trimer, the structure changes from C_1 symmetry in the gas phase (Figure 7a), to C_3 symmetry in water (Figure 7b). This result comprises support for the efficacy of the SCRF treatment since, in the more polar environment, the structure having the higher dipole moment is preferred.

Figures 10–14 give the thermochemical details of the reaction network at 3-21G//3-21G, MP2=FC/6-31G*//3-21G, MP2=FULL/6-31G*//MP2=FULL/6-31G*, 3-21G(SCRF)//3-

(29) Schowen, R. L. In *Isotope Effects on Enzyme-Catalyzed Reactions*; Cleland, W. W., O'Leary, M. H., Northrop, D. B., Eds.; University Park Press: Baltimore, 1972; p 64. A Referee has questioned our choice of 4.00 for this equllibrium constant, and not the experimental value of 3.78.[29a] This choice had been made for two reasons: a value of 4.00 greatly simplifies the derivations of the equations, which become increasingly more complicated as the number of water molecules increases from 1 to 2, 3, and 4; and, in dealing with the same point, Schowen (*op cit*) had previously commented that the difference between the two values "turns out to have only the slightest effect on the distribution of the three modifications of water (and) it has been demonstrated algebraically[29b] that violations of the rule of the geometric mean need arouse no concern in mechanistic applications of isotope effects...". In view of the Referee's concern, we calculated the mole fractions of HOH and HOD, and DOD in eq 8 for K = 4.00 and 3.78. There was no atom fraction of deuterium between 0 and 1 for which the different values of K led to a difference greater than ±0.01 in any of x(HOH), x(HOD), or x(DOD). We then derived the equation $V_n/V_0 = x(HOH) + x(HOD)[k(DH)/2k(HH) + k(HD)/2k(HH)] + x(DOD)\text{-}k(DD)/k(HH)$, which is a more general version of eq 23 (*vide infra*), and used the mole fractions x calculated at the different atom fractions n, with K = 4.00 and 3.78, to obtain the proton inventories. There was no difference. (a) Friedman, L.; Shiner, V. J. *J. Chem. Phys.* **1966**, *44*, 4639–4640. Van Hook, W. A. *J. Chem. Soc., Chem. Commun.* **1972**, 479–480. (b) Albery, W. J.; Davies, M. H. *Trans. Faraday Soc.* **1969**, *65*, 1059–1065.

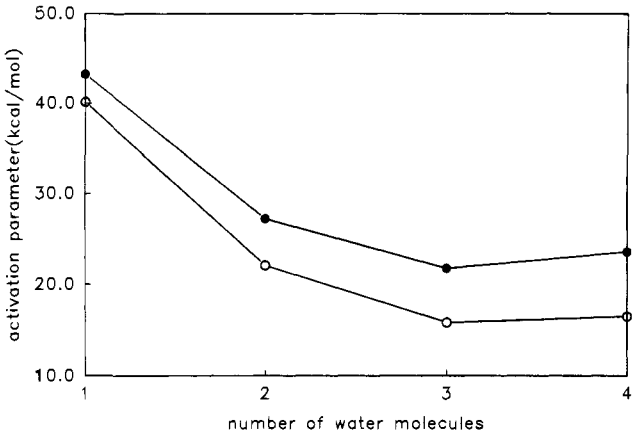


Figure 15. Variations in $\Delta H^{\ddagger}$ (lower plot) and $\Delta G^{\ddagger}$ (upper plot) for the hydration of formaldehyde by $(H_2O)_n$ as a function of n, with full optimization of all structures at MP2=FULL/6-31G*.

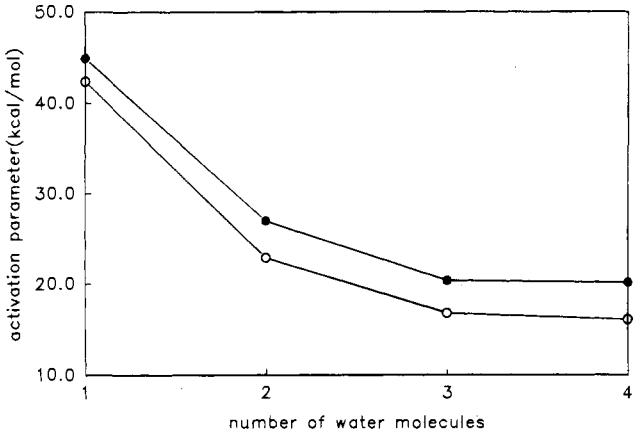


Figure 16. Variations in $\Delta H^{\ddagger}$ (lower plot) and $\Delta G^{\ddagger}$ (upper plot) for the hydration of formaldehyde by $(H_2O)_n$ as a function of n, with structures and vibrational frequencies calculated at 3-21G and energies calculated at MP2=FC/6-31G*.

21G(SCRF), and MP2=FC/6-31G*(SCRF)//3-21G(SCRF). In each case, the numbers attached to each arrow refer, from top to bottom, to ΔE, ΔH, ΔS, and ΔG, i.e., the changes in electronic energy, enthalpy, entropy, and free energy. The global consistency of the free energies shown in each of Figures 10−14 can be seen upon inspection of the many thermochemical cycles contained in these figures.

II. The Trends in the Activation Energies Are Compatible with a Cooperative Mechanism That Employs at Least Two Catalytic Water Molecules. Figure 15 shows the variation in $\Delta H^{\ddagger}$ and $\Delta G^{\ddagger}$, i.e., the energy difference between $(RC)_n$ and $(TS)_n$ in the gas phase as a function of n, with all structures optimized at MP2=FULL/6-31G*. Comparison of Figure 15 with Figure 16 suggests that one-point MP2=FC/6-31G* calculations on structures optimized at 3-21G afford reliable trends in these energies. Figure 17 is obtained when this strategy is applied to the SCRF results in water.

The decrease in $\Delta H^{\ddagger}$ associated with the cooperative mechanism and bifunctional catalysis by one water molecule, i.e., the change between $n = 1$ and 2, is 18.1 kcal/mol in the gas phase and 21.0 kcal/mol in water. A second water molecule (from $n = 2$ to 3) decreases $\Delta H^{\ddagger}$ by an additional 6.27 kcal/mol in the gas phase and 6.8 kcal/mol in water solvent. However, the third water molecule (from $n = 3$ to 4) has an inhibitory effect in the gas phase, and only a modest additional catalytic effect in water solvent.

Based on these trends in the *activation energies*, it can be concluded that, if sufficient water is available, the neutral hydration of formaldehyde can proceed via a cooperative

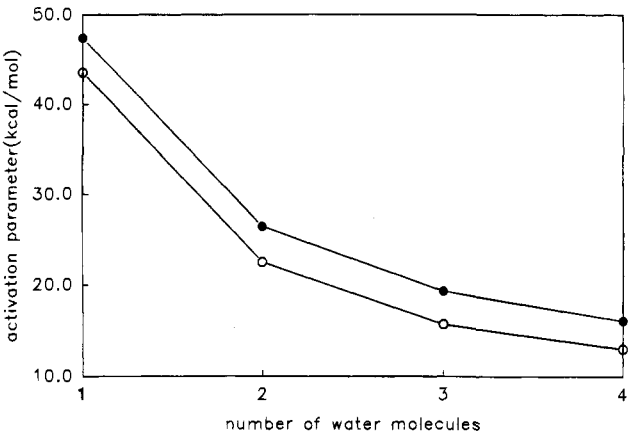


Figure 17. Variations in $\Delta H^{\ddagger}$ (lower plot) and $\Delta G^{\ddagger}$ (upper plot) for the hydration of formaldehyde by $(H_2O)_n$ as a function of n, with structures and vibrational frequencies calculated at 3-21G(SCRF) and energies calculated at MP2=FC/6-31G*(SCRF).

mechanism in the gas phase and also in water solvent, and that this mechanism involves active catalysis by two additional water molecules in the gas phase and either two or three additional water molecules in solution.

III. Enthalpies of Reaction. The heat of hydration (ΔH) of formaldehyde is −8.4 kcal/mol in water solvent.[3a] Table 2 shows the enthalpies of hydration calculated in the gas phase and in water solvent, from Figures 12 and 14. The first entries in the Table refer to the formation of methanediol from the separated reactants, but these values are unreliable because of basis set superposition error (BSSE).[30] This results from the use of different numbers of basis functions to describe the separated reactants and the product when a standard basis set is employed; the result is that the separated reactants are less well described than the product, and their energies are artificially too high. A treatment of the BSSE in these reactions, which allows the entire reaction coordinate to be discussed, will be given in Section VIA.

The BSSE problem is avoided in the remaining entries of Table 2 because $(RC)_n$ and $(TS)_n$ are calculated using the same number of basis functions. For the SCRF calculations in water, the largest heat of reaction is seen with $n = 3$. The ΔH calculated in this case (−9.57 kcal/mol) is slightly larger than the experimental value, and not very different from the values calculated for $n = 2$ and 4. Thus, the calculated heats of reaction, like the calculated trends in the activation energies, are consistent with the existence of a cooperative mechanism in water solvent, but neither of these thermochemical criteria permits a choice between active catalysis by two water molecules ($n = 3$) and active catalysis by three water molecules ($n = 4$). We must therefore proceed to an examination of the isotope effects.

IV. Solvent Isotope Effects. An Isotope Effect of 4.0 Is Consistent with Two Catalytic Water Molecules. Table 3 summarizes the solvent isotope effects (k_{HOH}/k_{DOD}) calculated

(30) (a) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553. (b) Ventura, O. N.; Coltiño, E. L.; Irving, K.; Iglesias, A.; LLedós, A. *J. Mol. Struct. (Theochem)* **1990**, *210*, 427−440. Cook, D. B.; Sordo, T. L.; Sordo, J. A. *J. Chem. Soc., Chem. Commun.* **1990**, 185−186. Åstrand, P.-O.; Wallqvist, A.; Karlström, G. *J. Phys. Chem.* **1991**, *95*, 6395−6396. Eggenberger, R.; Gerber, S.; Huber, H.; Searles, D. *Chem. Phys. Lett.* **1991**, *183*, 223−226. Valiron, P.; Vibók, A.; Mayer, I. *J. Comput. Chem.* **1993**, *14*, 401−409. Turi, L.; Dannenberg, J. J. *J. Phys. Chem.* **1993**, *97*, 2488−2490. Dobado, J. A.; Molina, J. M. *J. Phys. Chem.* **1993**, *97*, 7499−7504. Davidson, E. R.; Chakravorty, S. J. *Chem. Phys. Lett.* **1993**, *217*, 48−56. Mayer, I.; Vibók, A.; Valiron, P. *Chem. Phys. Lett.* **1994**, *224*, 166−174. Wind, P.; Heully, J.-L. *Chem. Phys. Lett.* **1994**, *230*, 35−40. Neuheuser, T.; Hess, B. A.; Reutel, C.; Weber, E. *J. Phys. Chem.* **1994**, *98*, 6459−6467. Kieninger, M.; Suhai, S.; Mayer, I. *Chem. Phys. Lett.* **1994**, *230*, 485−490.

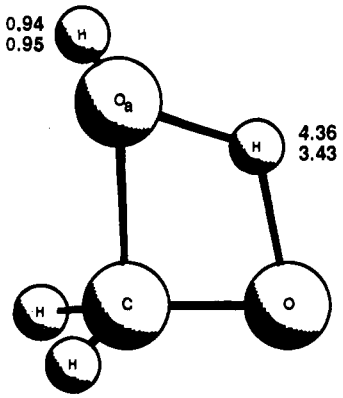


Figure 18. The contribution of each hydron to the overall solvent isotope effect for the reaction of formaldehyde with one water molecule in the gas phase (upper numbers) and in water (lower numbers).

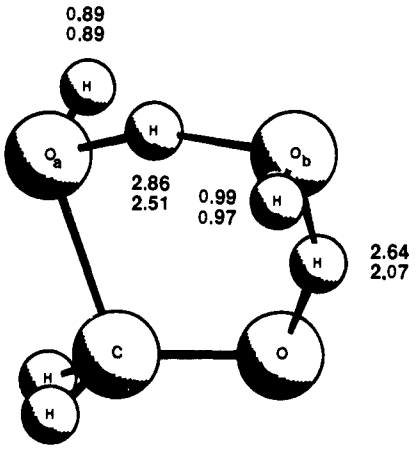


Figure 19. The contribution of each hydron to the overall solvent isotope effect for the reaction of formaldehyde with two water molecules in the gas phase (upper numbers) and in water (lower numbers).

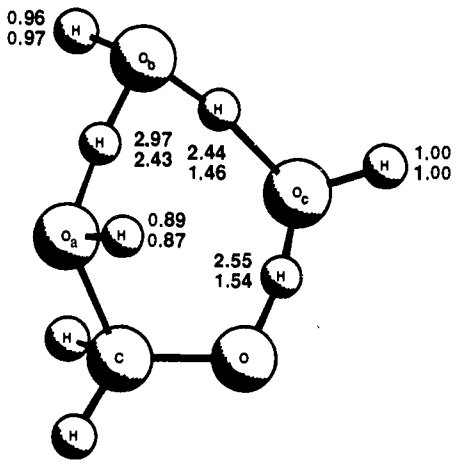


Figure 20. The contribution of each hydron to the overall solvent isotope effect for the reaction of formaldehyde with three water molecules in the gas phase (upper numbers) and in water (lower numbers).

at 298 K for the one-, two-, three-, and four-water molecule reactions. These were obtained using the Bigeleisen equation,[28a] which is a semiclassical treatment that does not take quantum mechanical tunneling into consideration. As we have discussed elsewhere,[31] hydron tunneling is unlikely to be significant in these reactions when $n > 1$; in any event, because of its high barrier, a cyclic one-water molecule mechanism is experimentally improbable.

The results of Pocker, in water,[4c] and Bell, in aqueous dioxane,[8] suggest that the experimental H/D solvent isotope

(31) Wolfe, S.; Kim, C.-K. *Isr. J. Chem.* **1993**, *33*, 295–305.

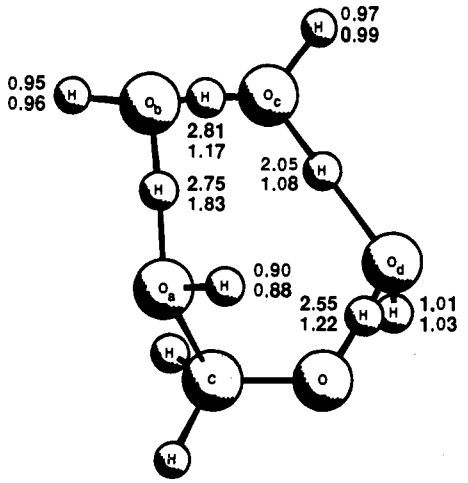


Figure 21. The contribution of each hydron to the overall solvent isotope effect for the reaction of formaldehyde with four water molecules in the gas phase (upper numbers) and in water solvent (lower numbers).

Table 2. Enthalpies of Hydration

	ΔH (kcal/mol)	
reaction	gas phase[a]	water[b]
$CH_2O + H_2O \rightarrow (PC)_1$	−8.68	−5.37
$(RC)_1 \rightarrow (PC)_1$	−5.34	−1.75
$(RC)_2 \rightarrow (PC)_2$	−5.86	−8.23
$(RC)_3 \rightarrow (PC)_3$	−8.21	−9.57
$(RC)_4 \rightarrow (PC)_4$	−8.75	−9.22

[a] MP2=FULL/6-31G*//MP2=FULL/6-31G* calculations. [b] MP2=FC/6-31G*//3-21G(SCRF) calculations.

Table 3. Solvent Isotope Effects for the Process $(RC)_n \rightarrow (TS)_n$

	k_{HOH}/k_{DOD}			k_{HOH}/k_{DOD}	
n	gas phase[a]	water[b]	n	gas phase[a]	water[b]
1	4.18	3.36	3	17.39	4.69[c]
2	7.21	4.65	4	36.13	2.45

[a] Calculated at MP2=FULL/6-31G*. [b] Calculated at 3-21G(SCRF). [c] For this case the gas-phase kinetic isotope effect is 6.49.

Table 4. Variation of the Dipole Moments Along the Reaction Coordinate for the Process $(RC)_n \rightarrow (TS)_n \rightarrow (PC)_n$

	dipole moment (debye)					
	gas phase[a]			water[b]		
n	$(RC)_n$	$(TS)_n$	$(PC)_n$	$(RC)_n$	$(TS)_n$	$(PC)_n$
1	0.893	1.281		2.543	1.929	
2	0.318	0.843	0.736	0.557	1.502	1.032
3	0.244	0.919	0.633	0.298	1.741	1.481
4	0.720	1.539	0.775	2.344	1.836	1.550

[a] Refers to MP2=FULL/6-31G* calculations. [b] Refers to 3-21G(SCRF) calculations.

Table 5. Basis Set Superposition Error in the Reactions $CH_2O + (H_2O)_n \rightarrow (RC)_n$ in the Gas Phase[a]

n	E^{BSSE} (kcal/mol)	n	E^{BSSE} (kcal/mol)
1	3.37	3	5.15
2	4.49	4	5.61

[a] MP2=FULL/6-31G* calculations.

effect for the uncatalyzed hydration of a carbonyl group is about 4.0. The gas-phase solvent isotope effects shown in Table 3 are all larger than 4.0, and they increase rapidly as the number of catalytic water molecules increases. The SCRF solvent isotope effects are in much better agreement with experiment. If agreement between experimental and theoretical solvent isotope effects is a proper criterion to distinguish between $n =$

Table 6. Basis Set Superposition Error for the Formation of $(H_2O)_n$ by Addition of H_2O to $(H_2O)_{n-1}$ in the Gas Phase[a]

n	E^{BSSE} (kcal/mol)
2	2.29
3	5.18 ± 0.06
4	5.45

[a] MP2=FULL/6-31G* calculations.

3 and 4, then the solvent isotope effects are best interpreted in terms of a cooperative mechanism having two catalytic water molecules.

Thus, each of the theoretical criteria we have employed supports Eigen's cooperative mechanism for the hydration of a carbonyl group[6] and Bell's conclusion that the neutral process proceeds with *two* catalytic water molecules. The agreement between the isotope effects observed experimentally and the isotope effects calculated for water solvent by the above mechanism is particularly striking and deserves comment as an additional illustration of how the SCRF treatment manifests itself. As seen in Table 4, this treatment leads to an increase in all $(RC)_n$, $(TS)_n$, and $(PC)_n$ dipole moments as these structures adjust to the more polar environment. The effect on the O···H···O bond lengths is the most pronounced; inspection of the transition structures in Figures 6–9 shows that the degree of proton transfer is systematically smaller in water than in the gas phase.

The contribution of each active hydron to the overall solvent isotope effect is, therefore, systematically smaller in water. The cumulative isotope effects shown in Table 3 are the result of normal ($k_H/k_D > 1$) contributions from each active hydron and, usually, inverse ($k_H/k_D < 1$) contributions from each passive hydron. Figures 18–21 show the individual H/D isotope effects for $n = 1$, 2, 3 and 4, respectively. In each case the upper number refers to the gas phase and the lower number to water solvent.

Since each hydron transferred in the cooperative mechanism exhibits an isotope effect, it follows that *nucleophilic attack on the carbonyl group does not occur synchronously with a single hydron transfer.* Nevertheless, the movement of the hydrons is asynchronous, especially in water solvent, and the hydron attached to O_A, the oxygen which carries out the attack on the carbonyl carbon, makes the dominant contribution.

This effect can be seen more clearly in Figure 22, which is obtained by animation of the normal mode of the imaginary frequency of the transition structure for $n = 3$ in water.[32] The progress of the reaction can be seen as a sequence of instant structures along this normal mode. Figure 22a shows the three water molecules poised for reaction, with the carbonyl carbon beginning to pyramidalize. In Figure 22b the C–O_A bond has formed, and the active hydron attached to O_A, which exhibits the largest isotope effect, has begun to stretch; by Figure 22c, this hydron has transferred to O_B, and the active hydrons attached to O_B and O_C have begun to stretch. In Figure 22d, all of the new bonds have formed, and these have begun to relax in Figure 22e.

V. Proton Inventories. In the course of their work on the hydration of 1,3-dichloroacetone in aqueous dioxane, Bell and Critchlow[8] carried out experiments with HOH, DOD, and a 1:1 HOH–DOD mixture. The rate constant in the mixture was less than the average of the rate constants in HOH and DOD. This corresponds to a bowl-shaped proton inventory in Schowen's terminology.[14] The interpretation of solvent isotope effects in mixtures of HOH and DOD is based on equations first derived independently by Gross[33] and Butler[34] and applied to the study of organic reaction mechanisms by Kresge[35] and Gold[36] and to biochemical proton transfer.[37]

A proton inventory can yield information regarding the number of *active* exchangeable hydrogens in a reaction and, in principle, the kinetic isotope effects at each of the contributing sites. There is always ambiguity, however, because models generated from the inventory data are not necessarily unique,[14] and if more than one hydron is implicated and the isotope effects differ, it is not easy to specify which hydrogen is responsible for a given isotope effect. It was, therefore, of interest to approach this question from the opposite direction, i.e., to predict the proton inventories associated with $n = 1$, 2, 3, and 4, since all isotope effects had already been calculated and assigned.

The results are presented in Figures 23–26 together with the linear relationship, shown in each case as a dashed line. Figures 23a and 23b are the proton inventories of the one-water molecule mechanism in the gas phase and in water, respectively. The inventories are similar and generally characteristic of a one-proton mechanism;[14] however, the inverse isotope effect associated with the passive hydrogen leads to a slight upward curvature.

Figures 24a and 24b are the proton inventories of the two-water molecule mechanism in the gas phase and in water, respectively. Both are bowl-shaped, as is expected for a two-proton mechanism,[14] and both become nearly linear when the square root of V_n/V_0 is plotted against n in Figures 24c and 24d. The effect of the passive hydrogens is again seen in the slight upward curvature.

Figures 25a and 25b are the proton inventories of the three-water molecule mechanism in the gas phase and in water, and Figures 25c and 25d are obtained when the cube root of V_n/V_0 is plotted against n. Although, as discussed earlier, the active hydron attached to O_A makes the dominant contribution to the isotope effect, and its transfer to O_B is more advanced than the transfer of the active hydrons attached to O_B and O_C, this is not reflected in the proton inventories, which are clearly not characteristic of one-proton mechanisms. The inventories in the gas phase and in water are not the same: Figure 25a is bowl-shaped in the region $n = 0.5$, Figure 25b is dome-shaped, and the gas-phase result is more compatible with the aqueous dioxane experiments of Bell and Critchlow.[8] Both cube root plots are dome-shaped.

All of the features just described for the three-water molecule mechanism are repeated in the inventories of the four-water molecule mechanism shown in Figures 26a–26d.

VI. Basis Set Superposition Error and Pathways to the Transition States. A. In the Gas Phase, Water Tetramer or Water Octamer Supplies the Catalytic Water Molecules. As noted briefly in Section III, when a standard basis set is used to compute the energies of all species, the energy of formation of a complex A·B from its components A and B will be in error by the quantity E^{BSSE} (eq 24), where E^{BSSE} is the basis set superposition error, E_{A-B} is the energy of the complex, and E_A and E_B are the energies of A and B.

$$\Delta E = E_{A-B} - (E_A + E_B) + E^{BSSE} \qquad (24)$$

In the *counterpoise* correction,[30a] E^{BSSE} is calculated using

(32) This was done using the computer program Xmol (XMol; Version 1.3.1. Minnesota Supercomputer Center, Inc., Minneapolis, MN, 1993).

(33) Gross, P.; Steiner, H.; Krauss, F. *Trans. Faraday Soc.* **1936**, *32*, 877–879. Gross, P. *Z. Elektrochem.* **1938**, *44*, 299–301.
(34) Orr, W. J. C.; Butler, J. A. V. *J. Chem. Soc.* **1937**, 330–335.
(35) Kresge, A. J. *Pure Appl. Chem.* **1964**, *8*, 243–258.
(36) Gold, V. *Adv. Phys. Org. Chem.* **1969**, *7*, 259–331.
(37) Kresge, A. J. *J. Am. Chem. Soc.* **1973**, *95*, 3065–3067.

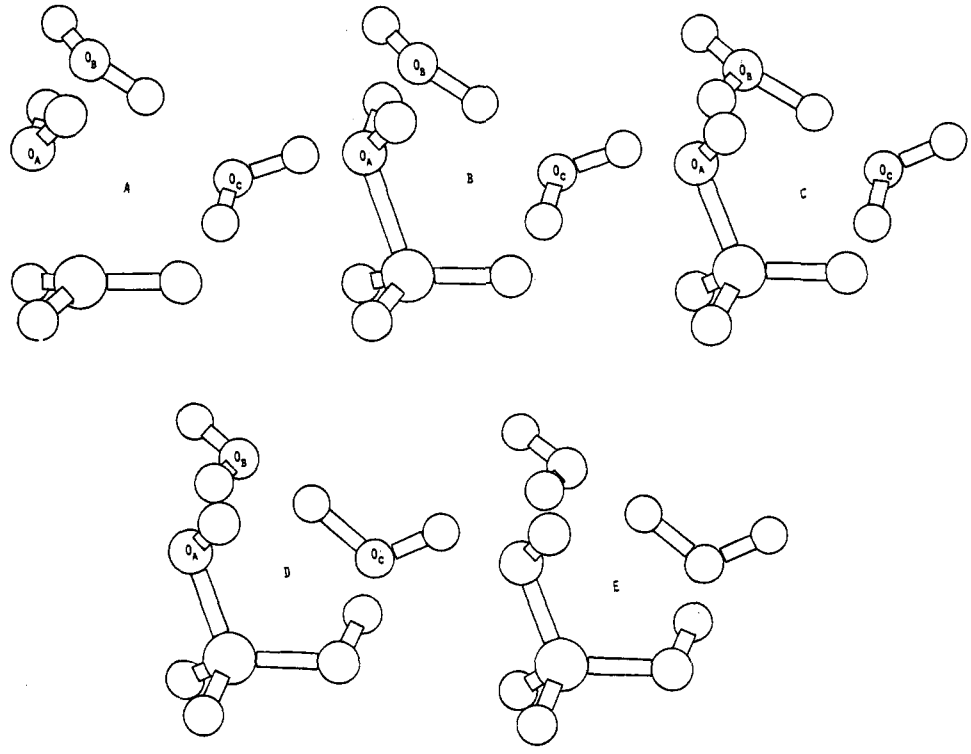


Figure 22. A sequence of instant structures which depict the normal mode of the imaginary frequency of $(TS)_3$ in water solvent. Progress from A to E refers to movement along the reaction coordinate from reactant to product.

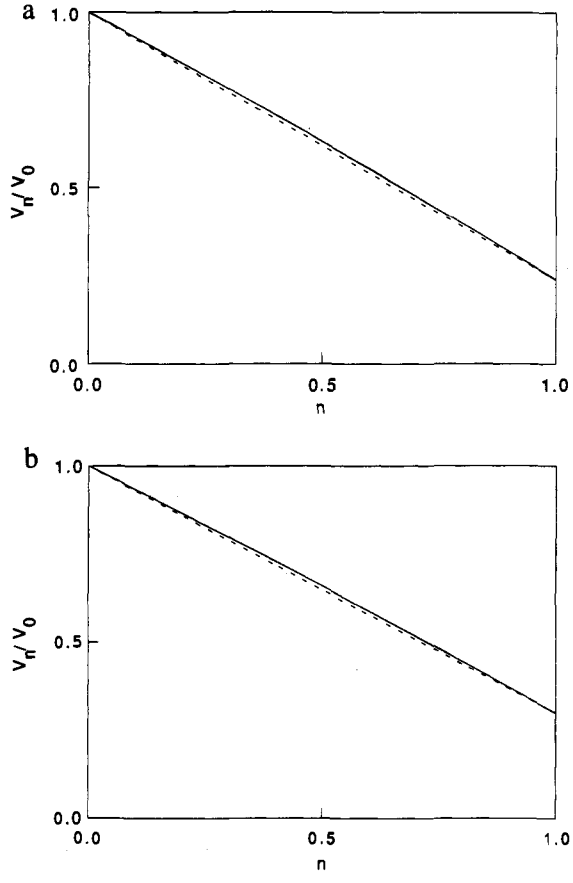


Figure 23. Proton inventories for the hydration of formaldehyde by one water molecule in the gas phase (a) and in water (b).

eq 25,

$$E^{\mathrm{BSSE}} = (E_{\mathrm{A}} - E^{\mathrm{c}}_{\mathrm{A(B)}}) + (E_{\mathrm{B}} - E^{\mathrm{c}}_{\mathrm{B(A)}}) \qquad (25)$$

where $E^{\mathrm{c}}_{\mathrm{A(B)}}$ and $E^{\mathrm{c}}_{\mathrm{B(A)}}$ are the energies of A and B in the presence of the empty (ghost) B and A orbitals, respectively, in the geometry of the complex.

An additional correction is necessary in the present case because of the changes in the geometries of CH_2O and H_2O when these are incorporated into a complex $(RC)_n$. For example, the insertion of formaldehyde into cyclic water trimer to form $(RC)_3$ (Path b of Figure 1) replaces a hydrogen bond between two water molecules by a hydrogen bond between a water molecule and a carbonyl oxygen. Concurrently, the internal angles of the structure change as the six-membered ring of water trimer is expanded to the eight-membered ring of the complex. Equation 25 should, therefore, be rewritten as eq 26, where

$$E^{\mathrm{BSSE}} = (E^{\mathrm{c}}_{\mathrm{A}} - E^{\mathrm{c}}_{\mathrm{A(B)}}) + (E^{\mathrm{c}}_{\mathrm{B}} - E^{\mathrm{c}}_{\mathrm{B(A)}}) \qquad (26)$$

$E^{\mathrm{c}}_{\mathrm{A}}$ and $E^{\mathrm{c}}_{\mathrm{B}}$ refer to the energies of A and B in the geometry of the complex. The difference between eqs 25 and 26, $(E^{\mathrm{c}}_{\mathrm{A}} - E_{\mathrm{A}}) + (E^{\mathrm{c}}_{\mathrm{B}} - E_{\mathrm{B}})$, is then a deformation energy.[38] There is still some ambiguity when more than two monomeric units are present in a complex, because different combinations of A with B lead to slightly different values of E^{BSSE}. In such cases we have averaged these values. As can be seen in Table 5, at MP2=FULL/6-31G* in the gas phase, the BSSE for the formation of $(RC)_n$ from CH_2O and $(H_2O)_n$ increases systematically as n increases.

The reactions shown from top to bottom at the left of Figure 1 constitute a second source of BSSE, because of the increase in the size of the water clusters. Table 6 summarizes the BSSE calculated at MP2=FULL/6-31G* for the stepwise addition of water to a water cluster. Taking the data of Tables 5 and 6 into account alters the free energies of Figure 12 to the BSSE-corrected values shown in Figure 27. The numbers shown in the second and third columns of this figure are derived from the appropriate thermochemical cycle.

It is possible, using Figure 27, to discuss the different reaction channels for the gas-phase hydration of formaldehyde by

(38) Mitchell, D. J.; Schlegel, H. B.; Shaik, S. S.; Wolfe, S. *Can. J. Chem.* **1985**, *63*, 1642–1648.

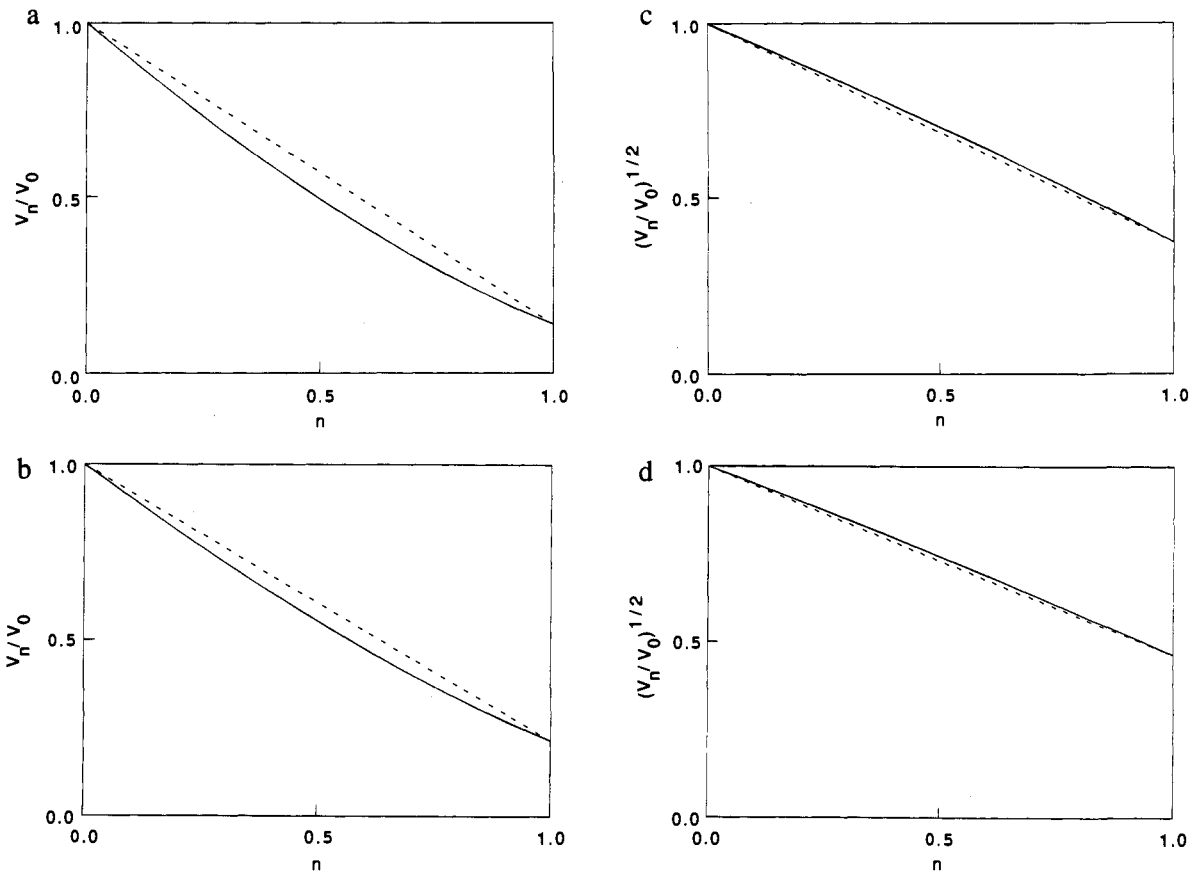


Figure 24. Proton inventories for the hydration of formaldehyde by a cooperative mechanism, with catalysis by one water molecule in the gas phase (a) and in water (b) and when the square root of V_n/V_0 is plotted in the gas phase (c) and in water (d).

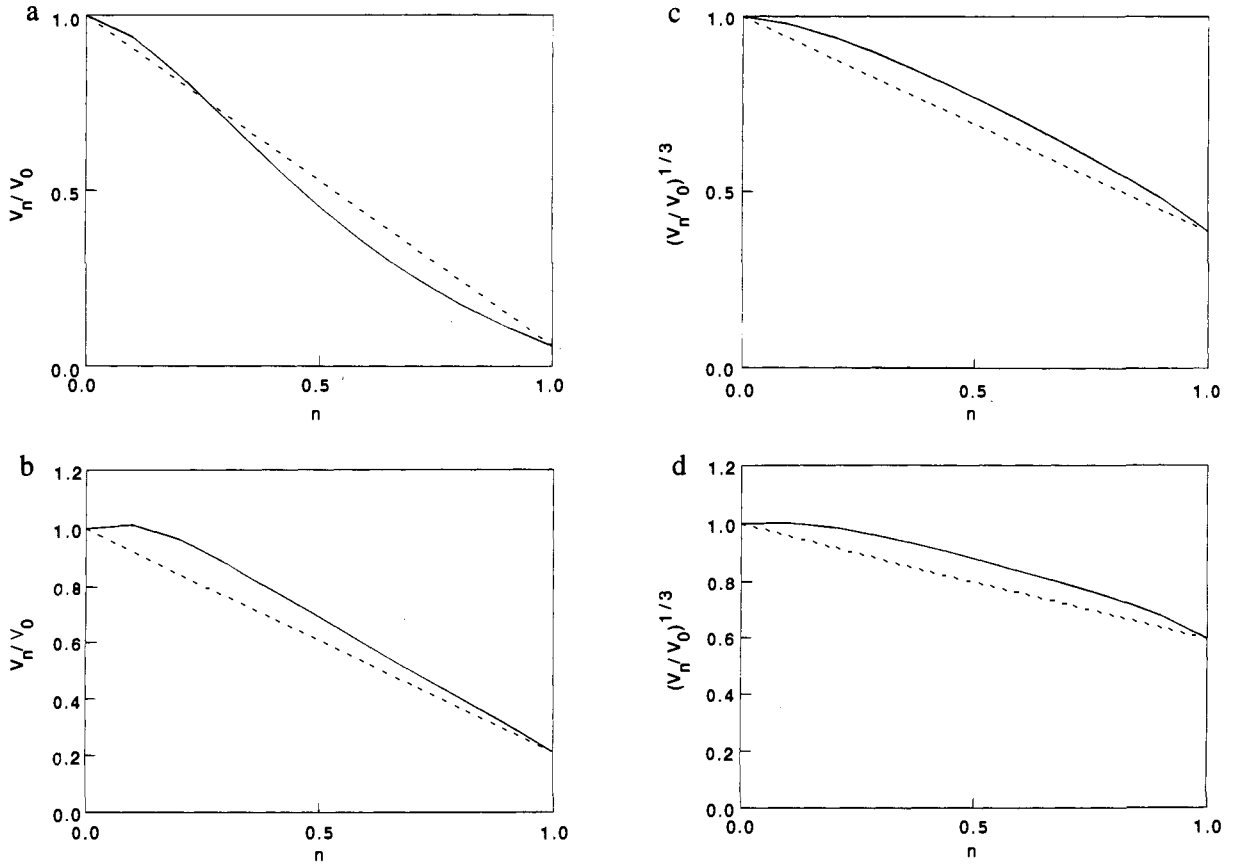


Figure 25. Proton inventories for the hydration of formaldehyde by a cooperative mechanism, with catalysis by two water molecules, in the gas phase (a) and in water (b) and when the cube root of V_n/V_0 is plotted in the gas phase (c) and in water (d).

different sized water clusters. To reach $(TS)_n$ from CH_2O and $(H_2O)_n$, i.e., proceeding from left to right in each row of Figure 27, the free energy changes are the following: $(TS)_1$, 50.82 kcal/mol; $(TS)_2$, 32.62 kcal/mol; $(TS)_3$, 29.53 kcal/mol; $(TS)_4$, 34.83

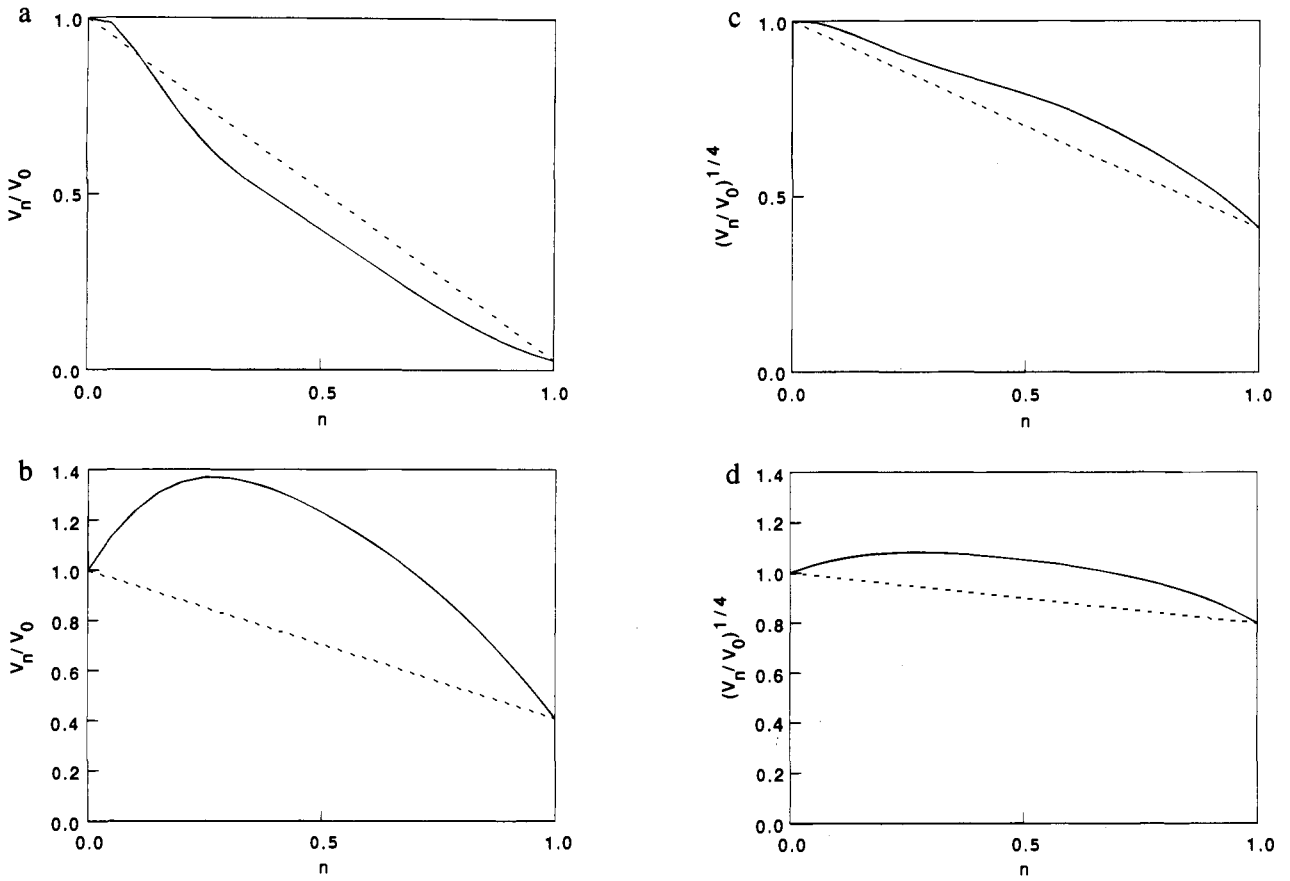
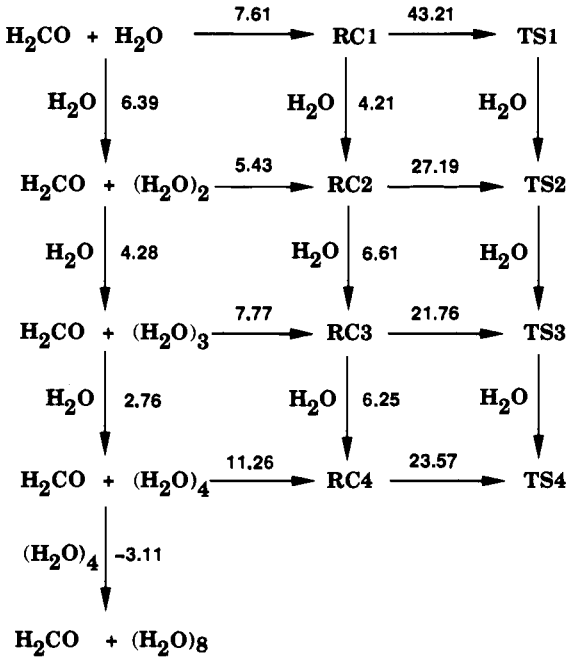


Figure 26. Proton inventories for the hydration of formaldehyde by a cooperative mechanism, with catalysis by three water molecules, in the gas phase (a) and in water (b) and when the fourth root of V_n/V_0 is plotted in the gas phase (c) and in water (d).



H_2CO + n H_2O Reaction in the Gas Phase with BSSE (MP2/6-31G*//MP2/6-31G*)

Figure 27. The free energy changes in the network of Figure 12 when basis set superposition error is taken into account.

kcal/mol. This suggests that the most favorable process is the path via $(RC)_3$ to $(TS)_3$. However, if the vertical pathways of Figure 27 are also taken into account, the situation is somewhat different, as summarized in Figure 28. The lowest energy pathway to $(TS)_3$ is now seen to begin with water tetramer. Since $(TS)_3$ is preceded in the network by $(RC)_3$, the latter is now seen to be formed preferentially by the displacement of one water molecule of the tetramer by formaldehyde (Figure 28, Path c of Figure 1).

Although Figure 28 also shows a 1.18 kcal/mol free energy preference for pathways leading to $(TS)_2$ via $(RC)_2$, consistent with the contention[3b] that $(TS)_2$ "*provides the major channel*

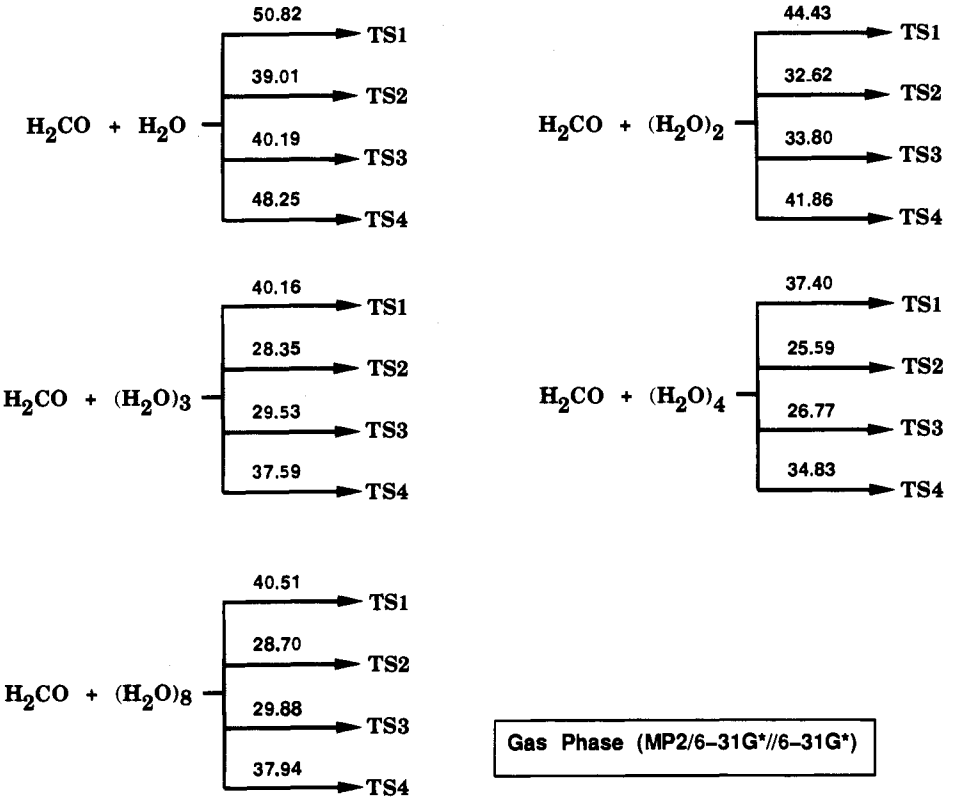


Figure 28. The free energy changes associated with the different ways in which CH_2O and $(H_2O)_n$ can reach the different $(TS)_n$ in the gas phase.

Table 7. Basis Set Superposition Error in the Reactions $CH_2O + (H_2O)_n \rightarrow (RC)_n$ in Water Solvent[a]

n	E^{BSSE} (kcal/mol)	n	E^{BSSE} (kcal/mol)
1	3.47	3	11.24
2	8.70	4	13.45

[a] 3-21G(SCRF)//3-21G(SCRF) calculations.

for hydration of formaldehyde in the gas phase", this conclusion required the assumption that the contribution of a process leading via $(RC)_3$ to $(TS)_3$ would be so disfavored entropically as to be negligible. In fact, the loss of entropy associated with the formation of $(RC)_3$ is almost exactly compensated by the gain in enthalpy in the process $(RC)_3 \rightarrow (TS)_3$.

It is useful to inquire at this stage why an 8-membered cyclic process should be enthalpically so preferred over a 6-membered cyclic process. The explanation seems to be found[3b,39] in the requirement for *linear proton transfer*[40] in a cyclic structure, and $(RC)_3$ and $(TS)_3$ are better able to accommodate this requirement. Thus, inspection of Figures 3–5 shows that the average gas-phase (solution) internal O···H···O angles in $(TS)_n$ increase, from 154.1° (152.4°) for $n = 2$, to 169.5° (165.3°) for $n = 3$, to 176.0° (171.7°) for $n = 4$.[41] If these angles are compared to the strainless ("standard") O···H···O angles of 167.3° (165.8°) of water dimer (Figure 6) it appears that only $(TS)_3$ and, interestingly, also $(RC)_3$ have this optimum geometry. When the same criterion is applied to the water clusters of Figures 7–9, we find that only cyclic water tetramer (168.7°, 166.2°) is unstrained. In the case of water octamer, Figure 9 shows two tetrameric units with local D_2 symmetry (oxygens a–b–c–d and e–f–g–h) and hydrogen bond angles of 158.9° (157.6°) and 171.0° (168.0°), held together by 171.0° (168.0°) hydrogen bonds.

These geometrical considerations are consistent with the conclusion that $(H_2O)_4$ and $(RC)_3$ are the most important species in the reaction network.

B. Entropic Effects and Reaction Pathways in Water Solvent. 99.9% of the Reaction Proceeds via $(RC)_3$, and the Experimental Pseudo-First-Order Rate Constant Is Reproduced. For the discussion of the preceding section to be extended to reactions in water solvent it was first necessary to check that the strategy used to estimate the gas-phase BSSE remains valid in the SCRF treatment and, in particular, whether the presence of ghost orbitals affects the radius of the cavity. We found that this radius varied by less than 0.05 Å in the presence of the ghost orbitals, and that, so long as the cavity radius of A was used to calculate E_A^c and $E_{A(B)}^c$, and the cavity radius of B was used to calculate E_B^c and $E_{B(A)}^c$, the effect on the BSSE was never greater than 0.3 kcal/mol. Because of the nature of the procedure only optimized energies are meaningful, and the results which follow refer to 3-21G(SCRF)//3-21G(SCRF) calculations.

Tables 7 and 8 are the BSSE corrections for the formation of $(RC)_n$ and $(H_2O)_n$, respectively, in water solvent. Figure 29 summarizes the BSSE-corrected free energies of the reaction network, and Figure 30 shows the free energy changes associated with the different pathways to $(TS)_n$ in water solvent.

(39) The authors are grateful to a Referee for drawing the work of Gandour (Gandour, R. D. *Tetrahedron Lett.* **1974**, 295–298) to our attention.

(40) Scheiner, S. *Acc. Chem. Res.* **1994**, *27*, 402–408.

(41) In related work,[42] we have found the following: (i) that the methanolysis of β-lactam rings by a cooperative one-step process involving proton transfer to *nitrogen* is *ca.* 10 kcal/mol *lower* in energy than the cooperative process, involving proton transfer to oxygen, which leads to a tetrahedral intermediate; (ii) in contrast to the current findings concerning the hydration of formaldehyde, there is no enthalpic difference in the catalysis of the preferred reaction by one water molecule (i.e., a 6-membered cyclic process) or by two water molecules (an 8-membered cyclic process); (iii) also in contrast to the results for formaldehyde hydration, the average O···H···O and O···H···N angles of the transition structures are almost the same in the 6-membered (158.6°) and 8-membered (156.5°) rings.

(42) Wolfe, S.; Kim, C.-K.; Yang, K. *Can. J. Chem.* **1994**, *72*, 1033–1043. For the application of these findings to the design of a novel antibacterial agent, see: Wolfe, S. *Can. J. Chem.* **1994**, *72*, 1014–1032. Wolfe, S.; Hoz, T. *Can. J. Chem.* **1994**, *72*, 1044–1050. Wolfe, S.; Jin, H.; Yang, K.; Kim, C.-K.; McEachern, E. *Can. J. Chem.* **1994**, *72*, 1051–1065. Wolfe, S.; Zhang, C.; Johnston, B. D.; Kim, C.-K. *Can. J. Chem.* **1994**, *72*, 1066–1075.

Table 8. Basis Set Superposition Error for the Formation of $(H_2O)_n$ by Addition of H_2O to $(H_2O)_{n-1}$ in Water Solvent[a]

n	E^{BSSE}
2	4.92
3	12.32
4	13.26

[a] 3-21G(SCRF)//3-12G(SCRF) calculations.

Inspection of Figure 30 suggests that the three-water molecule mechanism is almost 10 kcal/mol lower in energy than the four-water molecule mechanism, regardless of the starting point, but the free energy preference for the two-water molecule mechanism over the three-water molecule mechanism has increased from 1.18 kcal/mol in the gas phase to 1.41 kcal/mol in water. Although the calculated free energy change for the process $CH_2O + (H_2O)_4 \rightarrow (RC)_3 + H_2O \rightarrow (TS)_3$ is close to the estimated experimental free energy of activation for the neutral hydration of formaldehyde (16 kcal/mol), there is a much lower energy process which begins with water octamer.

To assess the *kinetic* relevance of these different pathways it is necessary to take the concentrations of the different $(RC)_n$ into account,[43] and this is possible if we make the reasonable assumption that CH_2O, $(H_2O)_n$, and $(RC)_n$ are in equilibrium in solution. The contributions of the different reaction channels to the overall pseudo-first-order rate constant for the hydration of formaldehyde will then depend upon these concentrations and on the individual free energies of activation $(RC)_n \rightarrow (TS)_n$.

The situation is summarized in Scheme 1, where w_1-w_8 refer to the water clusters and w_if are the different $(RC)_n$. In terms of this scheme, eqs 27 and 28 are the independent relationships

$$[w_2] = K_{12}[w_1]^2; \quad [w_3] = K_{13}[w_1]^3; \quad [w_4] = K_{14}[w_1]^4; \quad [w_8] = K_{18}[w_1]^8 \tag{27}$$

$$[RC1] = K_1[w_1][f]; \quad [RC2] = K_2[w_2][f]; \quad [RC3] = K_3[w_3][f]; \quad [RC4] = K_4[w_4][f] \tag{28}$$

among the concentrations of $(H_2O)_n$ and $(RC)_n$. Equations 29 and 30 relate these concentrations to the analytical concentra-

$$C_w = [w_1] + 2[w_2] + 3[w_3] + 4[w_4] + 8[w_8] + [RC1] + 2[RC2] + 3[RC3] + 4[RC4] \tag{29}$$

$$C_f = [f] + [RC1] + [RC2] + [RC3] + [RC4] \tag{30}$$

tions of water (C_w) and formaldehyde (C_f). Substitution of eqs 27 and 28 into eqs 29 and 30 leads to eqs 31 and 32 which, for

$$C_w = (1 + K_1[f])[w_1] + 2K_{12}(1 + K_2[f])[w_1]^2 + 3K_{13}(1 + K_3[f])[w_1]^3 + 4K_{14}(1 + K_4[f])[w_1]^4 + 8K_{18}[w_1]^8 \tag{31}$$

$$C_f = (1 + K_1[w_1] + K_2K_{12}[w_1]^2 + K_3K_{13}[w_1]^3 + K_4K_{14}[w_1]^4)[f] \tag{32}$$

a given C_w and C_f, can be solved for $[w_1]$ and [f]. The overall reaction rate can be written as a sum of partial rates (eq 33),

$$\begin{aligned} v &= v_1 + v_2 + v_3 + v_4 \\ &= k_1'[RC1] + k_2'[RC2] + k_3'[RC3] + k_4'[RC4] \\ &= k_1[w_1][f] + k_2[w_2][f] + k_3[w_3][f] + k_4[w_4][f] \\ &= (k_1[w_1] + k_2[w_2] + k_3[w_3] + k_4[w_4])[f] \end{aligned} \tag{33}$$

where $k_1 = K_1k_1'$, $k_2 = K_2k_2'$, $k_3 = K_3k_3'$ and $k_4 = K_4k_4'$. Substitution of eq 32 into eq 33 leads to the pseudo-first-order eq 34

$$v = kC_f \tag{34}$$

whose rate constant k is given in eq 35. The effective free

$$k = \frac{k_1[w_1] + k_2[w_2] + k_3[w_3] + k_4[w_4]}{1 + K_1[w_1] + K_2[w_2] + K_3[w_3] + K_4[w_4]} \tag{35}$$

energy of activation, $\Delta G^{\ddagger}$, can then be found from eq 36, where κ is the Boltzmann constant and h is Planck's constant.

$$\Delta G^{\ddagger} = RT\ln(\kappa T/h) - RT\ln k \tag{36}$$

The first column of Table 9 summarizes the results of this analysis, using $C_w = 55$ M and $C_f = 1$ M.[44] Although a 2/1 preference for reaction via $(RC)_3$ is now indicated, the calculated pseudo-first-order rate constant is orders of magnitude smaller than the observed rate constant of 9.2 s^{-1}.[3b] The reason for this low rate constant is that the ideal gas statistical mechanical treatment is used in Gaussian 92 to calculate thermodynamic parameters, not only in the gas phase, but also in the SCRF treatment (eq 37). Although this approximation is reasonable

$$\Delta G_{SCRF} = \Delta H_{\text{"gas"/SCRF}} - T\Delta S_{\text{"gas"/SCRF}} \tag{37}$$

for the description of the enthalpy of a reaction in solution (eq 38), entropic effects will be greatly overestimated.

$$\Delta H_{liq} \cong \Delta H_{gas} \tag{38}$$

In the gas phase, because vibrational effects are much smaller than translational and rotational effects, the major contribution to ΔS, as well as a significant contribution to the ΔG calculated for an association reaction $A + B \rightleftarrows AB$, arises from the loss of six degrees of freedom, representing unrestricted translations and free rotations of A and B, and the conversion of these into six vibrational modes of AB. In solution, however, rotational and translational motions are restricted and are better described by a set of low-frequency vibrations.[45] As a result, ΔS_{liq} represents only a fraction of $\Delta S_{\text{"gas"}}$ (eq 39). The use of eq 39 to estimate ΔS_{liq} leads to eq 40 for ΔG_{liq}.

$$\Delta S_{liq} = \lambda\Delta S_{\text{"gas"}} \quad (0 < \lambda < 1) \tag{39}$$

$$\Delta G_{liq} = \Delta G_{\text{"gas"}} - \lambda T\Delta S_{\text{"gas"}} \tag{40}$$

This equation is close to a model discussed by Williams,[3b,46] which is based on the observation by Wertz[47] of an empirical linear correlation between S_{gas} and S_{liq}.

The data shown in the second column of Table 9 are obtained

(43) The authors are indebted to a Referee for this suggestion.

(44) A computer program for the computation of these data is included with the Supplementary Data.

(45) Hirschfelder, J. O.; Curtis, C. F.; Bird, R. B. *Molecular Theory of Gases and Liquids*: Wiley, New York, 1954.

(46) Williams, I. H. *J. Am. Chem. Soc.* **1987**, *109*, 6299–6307.

(47) Wertz, D. H. *J. Am. Chem. Soc.* **1980**, *102*, 5316–5322. See also: Abraham, M. H. *J. Am. Chem. Soc.* **1981**, *103*, 6742–6744.

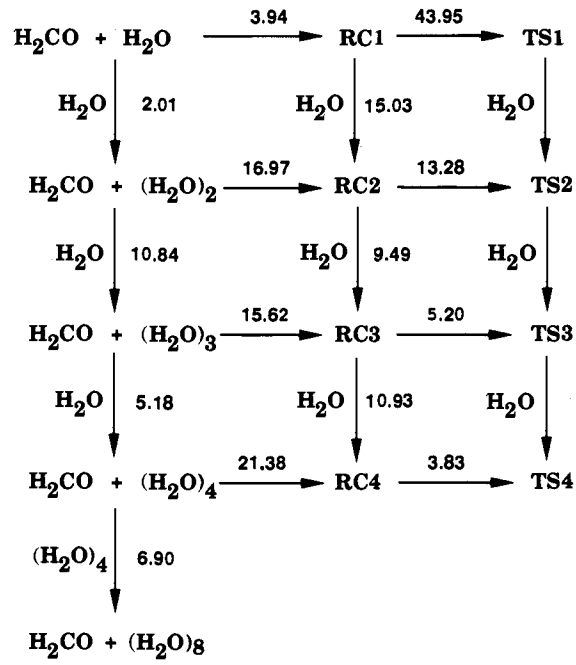


Figure 29. The free energy changes in the network of Figure 13 when basis set superposition error is taken into account.

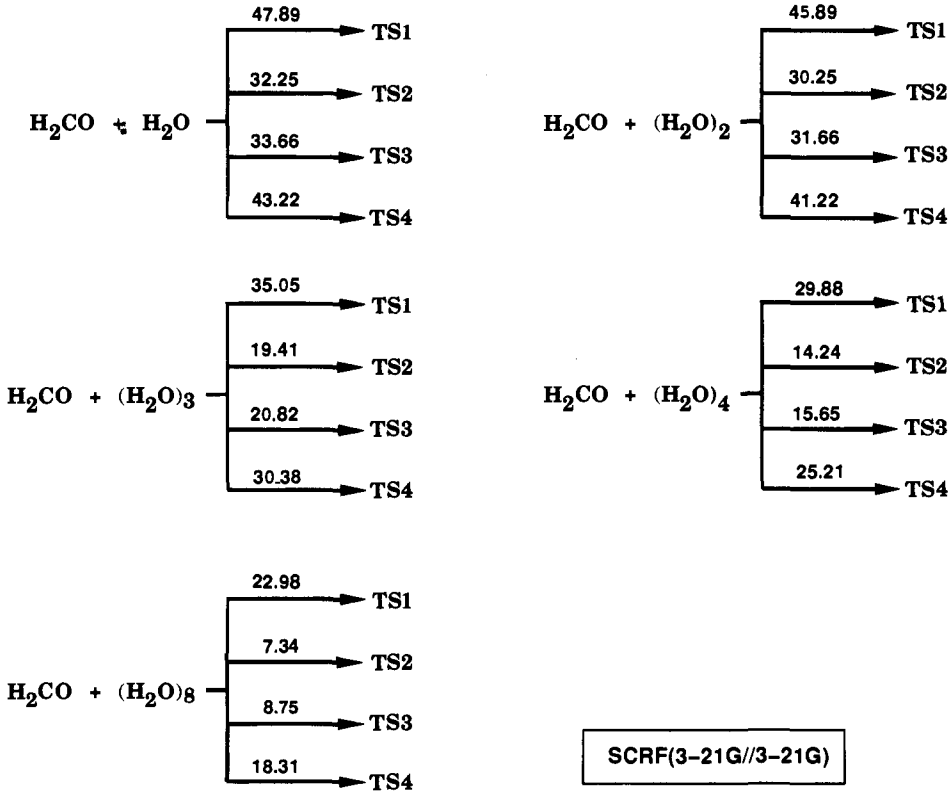


Figure 30. The free energy changes associated with the different ways in which CH_2O and $(H_2O)_n$ can reach the different $(TS)_n$ in water solvent.

with the assumption $\lambda = 0$ in eq 40. In this case 99.1% of the reaction proceeds via $(RC)_3$, but the calculated pseudo-first-order rate constant is two orders of magnitude *larger* than the experimental value. The data in the third column of Table 9 show that the experimental rate constant can be reproduced with $\lambda = 0.224$, and 99.8% of the reaction proceeds via $(RC)_3$.

In lieu of the empirical approach of eq 39, one could attempt to estimate ΔS_{liq} using the cell model of water.[16,45] In this model, bulk water is treated as a cubic lattice having a unit cell of size a given by eq 41,

$$a = (V_m/N_A)^{1/3} = 3.1 \text{ Å} \tag{41}$$

where V_m is the molar volume of water and N_A is Avogadro's number. A $3 \times 3 \times 3$ cubic cluster representing a fraction of this lattice is shown in Figure 31. If we assume that the low-frequency vibrations of a water molecule enclosed within this cluster can be represented by a single effective frequency ν_{eff}

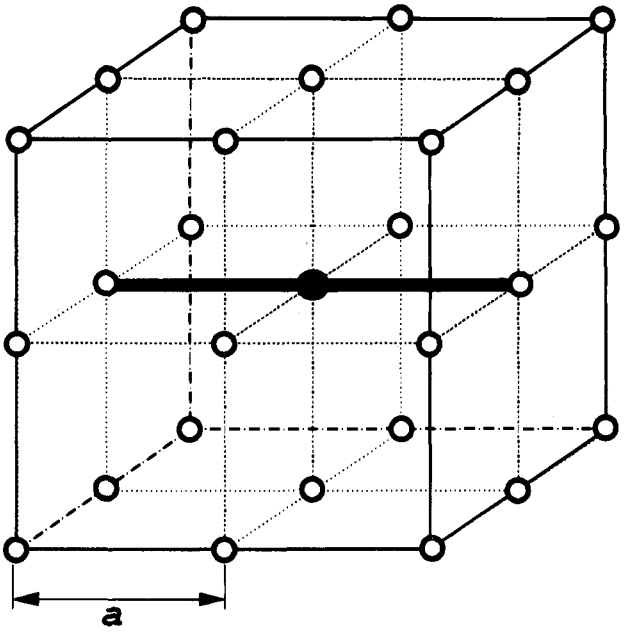


Figure 31. Cell model for water: the motion of the central water molecule (filled circle) is restricted by the neighboring molecules (open circles) forming the faces of the 3 × 3 × 3 cubic cluster. The bold line represents the one-dimensional model of Figure 32 (see text).

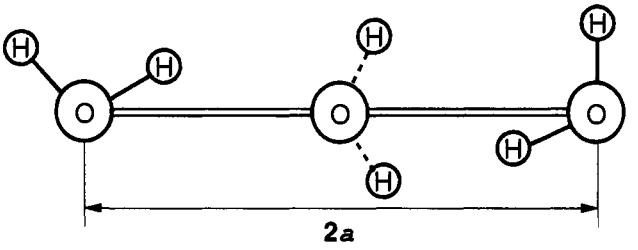
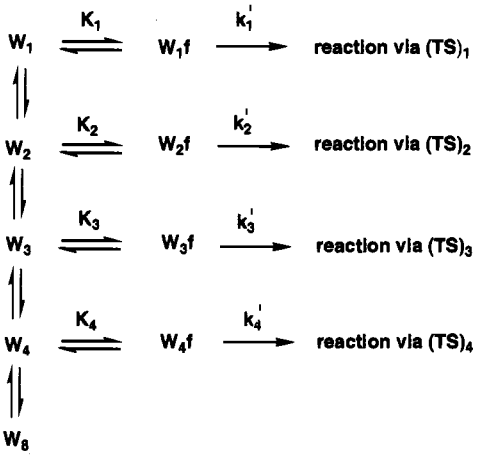


Figure 32. The linear cluster of three randomly oriented water molecules used to calculate the effective frequency ν. The centers of mass of the terminal molecules are constrained to a distance $2a = 6.2$ Å.

Scheme 1



$= \nu$ (the Einstein model[48]), the liquid phase entropy, S_{liq}, can be estimated from eq 42, in which the term S_v refers to the

$$S_{liq} = S_v + 6R\left[\frac{h\nu/\kappa T}{\exp(h\nu/\kappa T) - 1} - \ln\{1 - \exp(-h\nu/\kappa T)\}\right] \quad (42)$$

(48) Kubo, R. *Statistical Mechanics*; North Holland Publishers: Amsterdam, 1965.

contribution from higher frequency *intra*molecular vibrations, and the factor 6 refers to six low-frequency *inter*molecular vibrations.

If it is now supposed that ν is approximately the same for the separated A and B and for AB, the entropy of association can be written as eq 43, leading to eq 44, whose $\Delta H_{\text{"gas"}}$ and

$$\Delta S_{liq} = \Delta S_v - 6R\left[\frac{h\nu/\kappa T}{\exp(h\nu/\kappa T) - 1} - \ln\{1 - \exp(-h\nu/\kappa T)\}\right] \quad (43)$$

ΔS_v are obtained from the SCRF calculations.

$$\Delta G_{liq} = \Delta H_{\text{"gas"}} + T\Delta S_v - 6RT\left[\frac{h\nu/\kappa T}{\exp(h\nu/\kappa T) - 1} - \ln\{1 - \exp(-h\nu/\kappa T)\}\right] \quad (44)$$

To reduce the computational effort, the effective frequency ν was calculated from the one-dimensional model of Figure 32, corresponding to the bold line of Figure 31, in which only two neighboring water molecules were considered.

With the terminal oxygen atoms fixed at −3.1 and +3.1 Å from the central oxygen, the position of the latter was varied along the axis from −0.1 to +0.1 Å, using a step size of 0.02 Å. At each position, and with intramolecular O−H bond lengths and HOH angles maintained at the equilibrium values of water monomer, 1000 random orientations of the six hydrogens were generated and the 3-21G energies of these structures were calculated and averaged. The second derivative of the resulting potential curve was then calculated numerically. To check the degree of anharmonicity of this potential, differentiation was performed with step sizes of 0.02 and 0.04 Å. The resulting force constants k differed by less than 0.5%. The effective frequency ν was calculated using eq 45, where m is the mass

$$\nu = \frac{1}{2\pi}(k/m)^{1/2} \quad (45)$$

of a water molecule. The value found, after scaling by 0.89, was 115 cm^{-1}.

The data of Table 10 were obtained by solution of Scheme 1,[44] using $\nu = 115$ cm^{-1} in eq 44 to obtain the necessary free energies. The calculated pseudo-first-order rate constant and free energy of activation for the hydration of formaldehyde are 11.2 s^{-1} and 16.2 kcal/mol, in agreement with the experimental result, and also with the data in the third column of Table 9; in each case over 99.5% of the reaction proceeds via the three-water molecule complex $(RC)_3$.

The results of the two calculations are almost the same because the same free energies of activation, taken from Figure 13, were used in both cases, and the concentrations of $(RC)_2$, $(RC)_3$, and $(RC)_4$ are also very similar. However, the calculated concentrations of $(H_2O)_n$ are not the same: of the five species examined in the present work, the third column of Table 9 predicts the octamer, and Table 10 predicts the dimer to be the principal structure in liquid water. Neither of these predictions is consistent with Benson's tetramer−octamer model.[17] This is not too surprising. In Benson's analysis, higher order clusters were also suggested to play a role; these were beyond the scope of the present work. What is perhaps more important is that, with some hopefully reasonable assumptions, one of the central reactions of organic chemistry and biochemistry can be modeled, and experimental results reproduced.

Summary

The mechanism of a reaction is determined, typically, from measurements of rates, equilibria, reaction orders, activation

Table 9. Concentrations of CH_2O, $(H_2O)_n$, and $(RC)_n$, Contributions of the Different Reaction Channels, Pseudo-First-Order Rate Constants, and Free Energies of Activation for the Hydration of Formaldehyde in Water

	entropy = 1.0*S*		entropy = 0.0*S*		entropy = 0.224*S*	
species	conc[a]	contribution to reaction	conc	contribution to reaction	conc	contribution to reaction
$(H_2O)_1$	0.2228×10^{2}		0.1681×10^{-1}		0.1382	
$(H_2O)_2$	0.1665×10^{2}		0.2501×10^{1}		0.1041×10^{2}	
$(H_2O)_3$	0.4139×10^{-5}		0.8696×10^{-1}		0.4217×10^{-1}	
$(H_2O)_4$	0.1461×10^{-7}		0.2087×10^{1}		0.2317	
$(H_2O)_8$	0.3123×10^{-7}		0.5157×10^{1}		0.4122×10^{1}	
CH_2O	0.9721		0.2826		0.3925	
$(RC)_1$	0.2788×10^{-1}	0.000	0.7174	0.000	0.6075	0.000
$(RC)_2$	0.5751×10^{-11}	0.327	0.4228×10^{-5}	0.000	0.5946×10^{-6}	0.000
$(RC)_3$	0.1399×10^{-16}	0.673	0.4558×10^{-6}	0.991	0.9620×10^{-8}	0.998
$(RC)_4$	0.2938×10^{-23}	0.000	0.4055×10^{-9}	0.009	0.2168×10^{-11}	0.002
k (s^{-1})	1.98×10^{-6}		437		9.2	
$\Delta G^{\ddagger}$ (kcal/mol)	27.9		13.8		16.1	

[a] Mol/L

Table 10. Composition of a Solution of Formaldehyde in Water and Contributions of the Different Hydration Channels When Entropic Effects Are Calculated Using the Cell Model and an Effective Low-Frequency Vibration of 115 cm^{-1}

species	conc (M)	contribution to reaction
$(H_2O)_1$	0.9833	
$(H_2O)_2$	0.2621×10^{2}	
$(H_2O)_3$	0.1039×10^{-1}	
$(H_2O)_4$	0.1708	
$(H_2O)_8$	0.2785×10^{-9}	
CH_2O	0.1246	
$(RC)_1$	0.8754	0.0000
$(RC)_2$	0.7039×10^{-6}	0.0001
$(RC)_3$	0.1055×10^{-7}	0.9990
$(RC)_4$	0.1084×10^{-11}	0.0010

parameters, and isotope effects. When hydron transfer is involved, solvent isotope effects and proton inventories afford additional information. The application of all of these criteria led Bell and co-workers to conclude[4a,8] that the neutral hydration of a carbonyl group proceeds, as Eigen had proposed,[6] by a cyclic cooperative mechanism in which two additional water molecules provide the necessary catalysis. In the present work we have found that the same conclusion is reached when each of these probes is applied to the problem using ab initio molecular orbital theory. Consequently it has been possible to achieve insights into aspects of the reaction that cannot be examined experimentally, and to understand experimental results whose interpretation has sometimes required ambiguous assumptions. Thus, although the hydron transfers in the transition state have been found to be asynchronous, the use of proton inventories to demonstrate the multiple hydron transfers has been validated. The loss of entropy that attends the organization of four molecules into a reaction complex has been found to be almost exactly balanced by the enthalpic advantage associated with increased linearity of the hydron transfers, as suggested by Gandour[39] and by Scheiner.[40] The SCRF procedure[23,25] appears to work well; in conjunction with a protocol for the treatment of entropic effects in solution, this has allowed solvation to be taken into account for a reaction whose complexity may preclude the application of the free energy perturbation approach.[19,26] We can therefore proceed to an examination of the mechanism of carboxylic acid catalysis and the nature of rate–equilibrium and structure–reactivity relationships in the hydration of the carbonyl group, with some confidence that the results will be relevant to real experiments.

Acknowledgment. This work was supported by the Operating and Equipment Grants Programmes of the Natural Sciences and Engineering Research Council of Canada and by Simon Fraser University. The authors thank Professor Andrew Bennet for helpful discussions and the Referees for many insightful comments and suggestions.

Supplementary Material Available: Gaussian archive information for the structures calculated in this work, and a computer program for the analysis of Scheme 1 (13 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA941788M