Theoretical Study of Base-Catalyzed Amide Hydrolysis: Gas- and Aqueous-Phase Hydrolysis of Formamide

Dirk Bakowies[†] and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Abstract: Base-catalyzed hydrolysis of formamide in the gas phase and in aqueous solution has been studied using a combination of quantum chemical and statistical mechanical methods. A three-step procedure has been applied which comprises the determination of a gas-phase reaction path by high-level ab initio calculations, the calibration of empirical solute−solvent potentials, and classical Monte Carlo simulations of the solute immersed in a bath of solvent molecules. These simulations yield the solvent effect as a potential of mean force along the predetermined reaction coordinate. Each of the three consecutive steps of base-catalyzed hydrolysis has been analyzed in detail: the formation of a tetrahedral intermediate, its conformational isomerization, and the subsequent breakdown to products. The reaction is very exothermic in the gas phase and involves only moderate barriers for the latter two steps. Aqueous solvent, however, induces a significant barrier toward formation of the intermediate. On the other hand, it also facilitates conformational isomerization and produces a more product-like transition state for the breakdown step. Solvent effects, as expressed by differences in free energy of solvation, are found to reflect variations in the solute's charge distribution and are readily explained by the analysis of hydrogen bond patterns. The calculated free energy profile is in satisfactory agreement with available experimental data for the solution-phase reaction.

1. Introduction

The hydrolysis of amides serves as an important model reaction for the enzymatic cleavage of peptide bonds. In analogy to its biochemical counterpart, it involves the formation of a tetrahedral intermediate through nucleophilic attack at the carbonyl group. The details of the mechanism, however, depend on the experimental conditions applied.

Hydrolysis in neutral, aqueous solution has been observed,[1,2] but is known to be a very slow process that requires high temperatures.[1] A free energy of activation of 55 kcal/mol (at 25 °C) has been estimated by ab initio calculations[3,4] for the gas-phase addition of water to formamide (1), and similar values
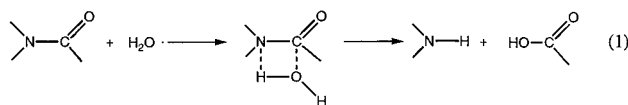
(48−51 kcal/mol) have been given for the aqueous-phase reaction.[5] Although there is some experimental evidence for reaction 1 to occur,[1] kinetic data favor a base-catalyzed mechanism (2) even in neutral aqueous solution.[1,2] This corresponds to a reaction first order in hydroxide concentration, a mechanism which has been suggested for the hydrolysis of
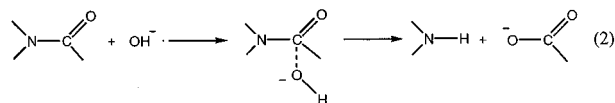
simple aliphatic amides, benzamides, and toluamides in alkaline media.[2,6−10] The free energy of activation has been measured experimentally and varies between 20 and 30 kcal/mol,[9,11] depending on substrate and temperature (typically 25−100 °C). The higher basicity of hydroxide thus provides for a substantial decrease in barrier height. Much of the experimental work has focused on the probability of oxygen exchange (^{18}O, ^{16}O) with respect to hydrolysis.[6−9,12] Assuming a rapid protonic equilibrium for the tetrahedral intermediate, these experiments yield valuable information about the barrier for the reverse reaction relative to that for C−N bond cleavage. A preference for hydrolysis was found for tertiary benzamides and toluamides, while the tetrahedral intermediates of primary amides tend to regenerate reactants.[6−9] This result was rationalized in terms of amine basicity and leaving ability.[7−9] The hydrolysis of amides involving amines of very low basicity (such as certain anilines) is known to follow rate laws of second order in hydroxide concentration, which is commonly explained with

[†] Current address: Laboratory of Physical Chemistry, Swiss Federal Institute of Technology (ETHZ), CH-8092 Zürich, Switzerland.

(1) Hine, J.; King, R. S.-M.; Midden, W. R.; Sinha, A. *J. Org. Chem.* **1981**, *46*, 3186−3189.

(2) Robinson, B. A.; Tester, J. W. *Int. J. Chem. Kinet.* **1990**, *22*, 431−448.

(3) Oie, T.; Loew, G. H.; Burt, S. K.; Binkley, J. S.; MacElroy, R. D. *J. Am. Chem. Soc.* **1982**, *104*, 6169−6174.

(4) Jensen, J. H.; Baldridge, K. K.; Gordon, M. S. *J. Phys. Chem.* **1992**, *96*, 8340−8351.

(5) Kallies, B.; Mitzner, R. *J. Mol. Model.* **1998**, *4*, 183−196.

(6) Bunton, C. A.; Nayak, B.; O'Connor, C. *J. Org. Chem.* **1968**, *33*, 572−575.

(7) Brown, R. S.; Bennet, A. J.; Slebocka-Tilk, H. *Acc. Chem. Res.* **1992**, *25*, 481−488.

(8) Slebocka-Tilk, H.; Bennet, A. J.; Keillor, J. W.; Brown, R. S.; Guthrie, J. P.; Jodhan, A. *J. Am. Chem. Soc.* **1990**, *112*, 8507−8514.

(9) Slebocka-Tilk, H.; Bennet, A. J.; Hogg, H. J.; Brown, R. S. *J. Am. Chem. Soc.* **1991**, *113*, 1288−1294.

(10) Liberles, A. *Introduction to Theoretical Organic Chemistry*; Macmillan: New York, 1968; pp 561−589.

(11) Guthrie, J. P. *J. Am. Chem. Soc.* **1974**, *96*, 3608−3615.

(12) Brown, R. S.; Bennet, A. J.; Slebocka-Tilk, H.; Jodhan, A. *J. Am. Chem. Soc.* **1992**, *114*, 3092−3098.

doubly negatively charged intermediates which expel the amine in its deprotonated form.[7,10,13]

Catalysis can also be achieved under acidic conditions. Two mechanisms may be considered: The first involves N-protonation, followed by addition of a water molecule, and C$-$N bond cleavage. The second mechanism assumes initial O-protonation and proton transfer to N either concurrently with or after the addition of water. Due to their zwitterionic character, amides prefer O-protonation[14,15] (by ca. 14 kcal/mol in the case of formamide[16]), which, however, also increases their C$-$N bond strength and makes them less susceptible to nucleophilic attack. Krug et al. have reported gas-phase ab initio calculations for both pathways and conclude that the hydrolysis of O-protonated formamide involves by far the higher energy barrier (24 vs 6 kcal/mol).[16] They argue, however, that the presence of a second water molecule might result in substantial stabilization of the transition state and thus favor the O-protonation pathway. This is supported by more recent ab initio and density functional studies which predict much lower activation barriers for a reaction involving a second water molecule as part of a six-membered ring transition state.[17,18] Unfortunately, these theoretical studies take limited or no account of the aqueous solvent. From experimental evidence, however, the O-protonation pathway is generally accepted for the aqueous-phase reaction.[7,19-21] The addition of water to the protonated amide is known to be the rate-limiting step.[16,19,22]

Even though acids can promote amide hydrolysis, base catalysis is generally more efficient.[1] In a number of studies on aliphatic amides, acid- and base-catalyzed hydrolyses have been compared directly, and activation energies were found to be 3$-$4 kcal/mol lower for base catalysis.[23,24] The higher efficiency is one reason, the closer analogy to the enzymatic process another, that makes the base-promoted pathway particularly interesting for further theoretical investigation. Only the base-promoted pathway affords the formation of a *negatively charged* tetrahedral intermediate, which is considered to be a characteristic feature of the reaction occurring in serine proteases (see Figure 1). The catalytic triad (Asp, His, Ser) is known to be the active site, initiating peptide cleavage by nucleophilic attack of Ser to the substrate. The His residue provides for deprotonation of Ser to achieve the required basicity, and Asp is assumed to stabilize the reactive center electrostatically.

The majority of the previously published theoretical work on amide hydrolysis considers neutral[3-5,16-18,25] or acid-catalyzed[16-18,25] hydrolysis, while most investigations of the base-promoted pathway are limited either to the gas-phase reaction[16,26] or to the formation of the tetrahedral intermediate,[25] or consider C$-$N bond cleavage without water assistance.[13,16,25,26]
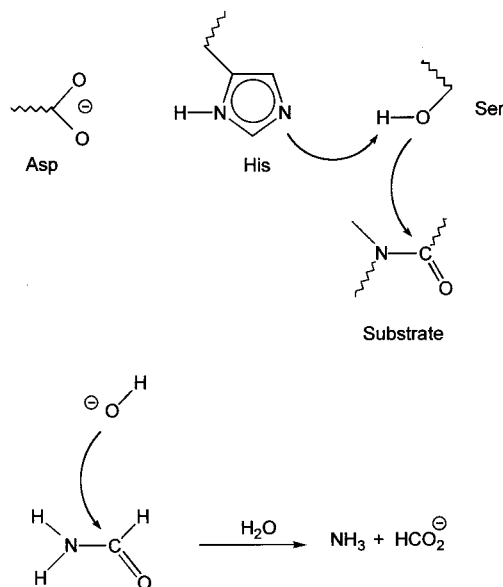

Figure 1. Schematic representation of amide hydrolysis in trypsin and under base catalysis.

which is known to be very unfavorable. Several authors have pointed out the substantial stabilization of transition states by ancillary water molecules. The catalytic effect of water in proton-transfer processes has been established for neutral[5,17,18] and acid-promoted amide hydrolysis[17,18] as well as for the addition of a water molecule to formaldehyde.[27,28] Surprisingly, such participation of a water molecule has rarely been considered for base-promoted amide hydrolysis. In 1985, Weiner et al. suggested that a water molecule catalyzes the C$-$N bond fission step by two consecutive proton transfers[29] (as shown in Figure 2). Due to computational constraints at that time, however, the authors only considered a few geometries on the reaction path without full optimization of transition states, and were further restricted to a very simple model based on molecular mechanical energy minimization to study the effect of bulk solvent.

Here we report a theoretical study of base-promoted hydrolysis of formamide, using high-level ab initio electronic structure theory combined with the statistical mechanical treatment of aqueous solvation. Following a procedure suggested by Jorgensen,[30] the potential of mean force for the solution-phase reaction is obtained from the gas-phase reaction path and from subsequent Monte Carlo simulations[31] of the solute immersed in a box of water molecules. The reaction studied in this work is outlined in Figure 2 (for atom numbering see Figure 3). The mechanism is analogous to that proposed by Weiner et al.[29] Our main objective is to study the aqueous-phase reaction, and to analyze the influence of bulk solvent. Solvent effects for related reactions have previously been studied with reaction field approaches[5,13,17,28] or point dipole models,[25] but the application of statistical mechanical techniques is clearly more appropriate in this regard. They not only provide a satisfactory description of specific solute$-$solvent interactions, but also permit a more

(13) Hori, K.; Kamimura, A.; Ando, K.; Mizumura, M.; Ihara, Y. *Tetrahedron* **1997**, *53*, 4317$-$4330.

(14) Gillespie, R. J.; Birchall, T. *Can. J. Chem.* **1963**, *41*, 148$-$155.

(15) Ou, M.-C.; Chu, S.-Y. *J. Phys. Chem.* **1995**, *99*, 556$-$562.

(16) Krug, J. P.; Popelier, P. L. A.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 7604$-$7616.

(17) Antonczak, S.; Ruiz-López, M. F.; Rivail, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 3912$-$3921.

(18) Antonczak, S.; Ruiz-López, M.; Rivail, J.-L. *J. Mol. Model.* **1997**, *3*, 434$-$442.

(19) O'Connor, C. *Q. Rev. Chem. Soc.* **1970**, *24*, 553$-$564.

(20) Smith, C. R.; Yates, K. *Can. J. Chem.* **1972**, *50*, 771$-$773.

(21) Kresge, A. J.; Fitzgerald, P. H.; Chiang, Y. *J. Am. Chem. Soc.* **1973**, *96*, 4698$-$4699.

(22) Bennet, A. J.; Slebocka-Tilk, H.; Brown, R. S.; Guthrie, J. P.; Jodhan, A. *J. Am. Chem. Soc.* **1990**, *112*, 8497$-$8506.

(23) Langlois, S.; Broche, A. *Bull. Soc. Chim. Fr.* **1964**, 812$-$816.

(24) DeRoo, M.; Bruylants, A. *Bull. Soc. Chim. Belg.* **1954**, *63*, 140$-$157.

(25) Voityuk, A. A.; Bliznyuk, A. A. *Bull. Acad. Sci. USSR Div. Chem. Sci.* **1989**, *38*, 1635$-$1641.

(26) Alagona, G.; Scrocco, E.; Tomasi, J. *J. Am. Chem. Soc.* **1975**, *97*, 6976$-$6983.

(27) Ventura, O. N.; Coitiño, E. L.; Lledos, A.; Bertran, J. *J. Comput. Chem.* **1992**, *13*, 1037$-$1046.

(28) Wolfe, S.; Kim, C.-K.; Yang, K.; Weinberg, N.; Shi, Z. *J. Am. Chem. Soc.* **1995**, *117*, 4240$-$4260.

(29) Weiner, S. J.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1985**, *107*, 2219$-$2229.

(30) Jorgensen, W. L. *Adv. Chem. Phys.* **1988**, *70*, 469$-$488.

(31) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, U.K., 1989.
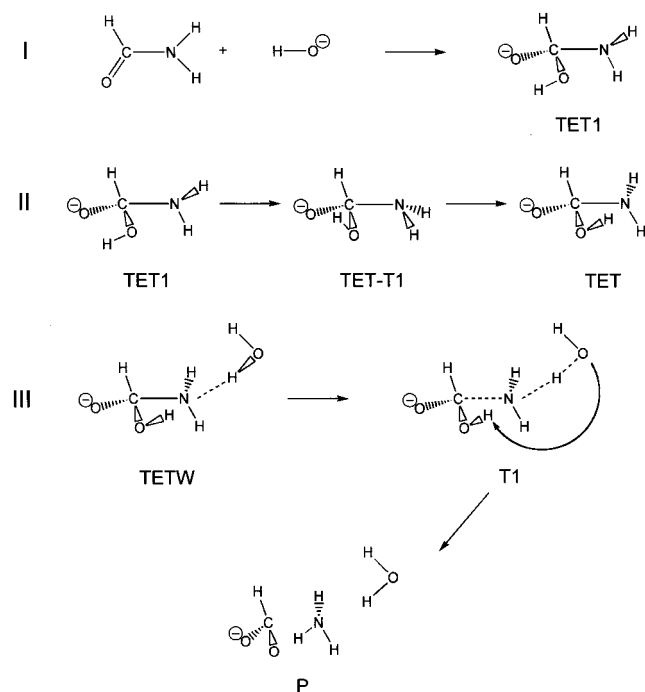
Figure 2. Schematic drawing of the formation (I), the conformational isomerization (II), and the breakdown (III) of the tetrahedral intermediate. For clarity, the product is not shown in its final geometry (see Figure 12).
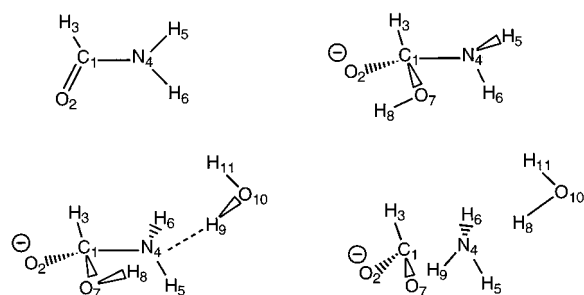

Figure 3. Numbering scheme used in the text, tables, and figures, shown for R, TET1, T1, and P. For clarity, atoms of classical water molecules are referred to by letters (HW and OW).

detailed analysis which can help to understand the mechanism by which solvent alters the energy profile.

Our paper begins with a description of the ab initio results for the gas-phase reaction. The following section describes our simulation protocol and discusses the custom-tailored design of appropriate force field models used in the Monte Carlo calculations. The main part of the paper is concerned with the presentation and analysis of the simulation results and is divided into three subsections, one for each of the consecutive steps of the reaction shown in Figure 2. The paper concludes with a discussion of our results and a comparison to available experimental data.

2. Ab Initio Calculations

The gas-phase reaction of base-catalyzed formamide hydrolysis has been studied using ab initio electronic structure theory[32,33] as implemented in the Gaussian suite of programs.[34,35] Stationary points have been located by geometry optimization and identified by subsequent force constant analysis. Whenever feasible,

intrinsic reaction coordinates[36-38] have been calculated which connect the transition states with the corresponding minima.

All geometries have been optimized at the MP2/6-31+G* level of theory.[32] The basis set includes diffuse functions on heavy atoms to ensure the appropriate description of the anionic reactants. Additional energy calculations have been performed at higher levels of theory to assess the effects of basis set truncation, basis set superposition, and higher order electron correlation. The largest basis set employed in our study (AUG-cc-pVTZ)[39,40] is of triple-ζ plus polarization quality, augmented with one diffuse function of each function type used. Electron correlation[41] has been studied at the MP2 and coupled cluster (CCSD)[42,43] levels of theory. All results are collected in Table 1.

The ab initio calculations predict a very exothermic formation of the tetrahedral intermediate (TET1) with no apparent barrier. The subsequent isomerization of TET1 to TET is moderately endothermic and involves a barrier of 6−7 kcal/mol. The water-assisted breakdown of TET proceeds over an early transition state and is even more exothermic than the first reaction. This sequence of three reactions involves a net gain in energy of ca. 50 kcal/mol, not counting the complex formation between TET and water. Nevertheless, formamide hydrolysis is unlikely to be observed in the gas phase, due to the preference for hydrogen abstraction ($H_2NCHO + OH^- \rightarrow OH_2\cdots HNCHO^-$, $\Delta E = -47.1$ kcal/mol at the MP2/6-31+G* level*).

The "best estimates" given in Table 1 are based upon counterpoise-corrected[44] MP2/AUG-cc-pVTZ energies and include adjustments for higher order electron correlation (see footnote *c* of Table 1). Fortunately, the lower level MP2/6-31+G* energies are reasonably close to these reference data, with differences typically being smaller than 3 kcal/mol. This good agreement with high-level ab initio calculations justifies the use of MP2/6-31+G* reaction coordinates and energies for the subsequent solution simulations.

3. Aqueous-Phase Simulations

To study the effects of aqueous solvation on base-catalyzed amide hydrolysis, we followed a protocol suggested by Jorgensen. For the purpose of this paper, we shall outline the procedure briefly and refer the reader to the original publications for a more thorough discussion.[30,45-49]

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

(33) Head-Gordon, M. *J. Phys. Chem.* **1996**, *100*, 13213−13225.

(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; 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/DFT*, Revision G4; Gaussian, Inc.: Pittsburgh, PA, 1993.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision D3; Gaussian, Inc.: Pittsburgh, PA, 1994.

(36) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363−368.

(37) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154−2161.

(38) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523−5527.

(39) Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(40) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(41) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem.* **1996**, *100*, 12960−12973.

(42) Purvis, G. D., III; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910−1918.

(43) Scuseria, G. E.; Schaefer, H. F., III. *J. Chem. Phys.* **1989**, *90*, 3700−3703.

(44) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873−1885.

Table 1. Gas-Phase Reaction: Ab Initio Results[a,b]

	R* → TET1	TET1 → TET-T1	TET-T1 → TET	TET + H_2O → TETW	TETW → T1	T1 → P
MP2/6-31+G*	−26.52 (+3.69)	7.43 (−0.54)	−2.65 (0.76)	−16.69 (+2.24)	8.45 (−1.25)	−42.57 (−0.43)
MP2/6-311+G(2d,p)	−26.24	6.46	−2.29	−15.64	9.17	−42.09
MP2/AUG-cc-pVTZ	−24.83	6.26	−2.12	−15.53	8.82	−39.70
MP2/AUG-cc-pVTZ+CP	−22.32			−14.63		
CCSD/6-311+G(2d,p)	−27.26	6.57	−2.54	−14.59	11.67	−44.70
best estimate[c]	−23.3	6.4	−2.4	−13.6	11.3	−42.3

[a] All energies refer to geometries optimized at the MP2/6-31+G* level of theory. Zero-point energy corrections are given in parentheses. R* = separated reactants; TET, TET1, TET-T1 = conformations of the tetrahedral intermediate; TETW = TET−water complex; T1 = transition state; P = product complex. See Figure 2. [b] CP = counterpoise correction. [c] The "best estimate" is based on the counterpoise-corrected MP2/AUG-cc-pVTZ result to which we add—in the spirit of complete basis set methods—the difference between the CCSD and MP2 energies for the 6-311+G(2d,p) basis set.

3.1. Theoretical Model. The procedure suggested by Jorgensen is based upon gas-phase reaction paths obtained from ab initio electronic structure theory. A suitable force field representation is then sought which accurately describes the interactions between the reacting solute and a single solvent molecule. Typically this requires separate calibrations of parameters for closely spaced points on the reaction coordinate. For reactions of charged solutes, solvent effects are likely to follow variations in the solute's charge distribution. The careful choice of a reliable charge model is thus a central focus of our study. In the second step, the custom-tailored force field is used in Monte Carlo (MC) simulations[31] of the solute immersed in a box of solvent molecules. Differences in free energy of solvation are calculated for adjacent points ("windows") on the reaction path, applying statistical perturbation theory[50] to the Monte Carlo generated ensembles. Repeating this procedure for all the predefined geometries from reactants to products yields the potential of mean force as a function of the reaction coordinate. The sum of this free energy of solvation (ΔG_{tot}) and the gas-phase internal energy (ΔE_{QM}) is used to estimate the free energy profile in solution (ΔG_{tot}):

$$\Delta G_{tot} = \Delta E_{QM} + \Delta G_{tot} \qquad (3)$$

The procedure outlined above combines the high-level quantum mechanical (QM) description of the gas-phase reaction with a completely classical simulation of solvent effects. Hence, it provides an alternative to the still unfeasible *fully* quantum mechanical simulation of chemical reactions in explicit solvent. Quantum mechanical potentials at sufficiently high levels of (ab initio) theory are currently too expensive to be used in simulations that require millions of single-point evaluations. More cost-effective semiempirical potentials employing MNDO-type Hamiltonians[51] (MNDO, AM1, PM3) are inappropriate, on the other hand, since they give a poor account of hydrogen bonding. In earlier studies on related reactions, semiempirical methods were found to be unable to reproduce the catalytic effect of ancillary water molecules[27] or to predict reasonable barriers for reactions involving proton transfer.[25] The procedure suggested by Jorgensen avoids these problems and appears to be the method of choice.

Despite its indisputable merits, however, the method has a few shortcomings related to the static description of the solute.

First, entropic effects are only considered for the solvent, but not for the solute. This may be regarded as a minor problem since the entropy of the solute is expected to change very little over the reaction coordinate.[52] Second, the geometry of the solute does not adjust to its environment. This precludes the study of solvent-induced changes of the reaction pathway and warrants the use of reaction coordinates which describe the solution pathway appropriately. As a rule of thumb, Jorgensen has suggested one should use gas-phase reaction coordinates only for reactions which do not involve charge separation.[53] While this requirement is fulfilled for the entire pathway of formamide hydrolysis, the first step of the gas-phase reaction is still not an appropriate approximation of the solution-phase reaction. There are two competing reactions between formamide and hydroxide:[10] proton abstraction of an NH and formation of a tetrahedral intermediate (TET). In the gas phase, the former is much preferred (by −20.6 kcal/mol at the MP2/6-31+G* level), disfavoring TET formation. In solution, however, the abstraction of a proton is a reversible process while TET formation initiates the hydrolysis reaction (see the Results). Regular ab initio geometry optimizations will always follow the (gas-phase) proton abstraction path and yield reaction coordinates that are inappropriate for subsequent solution simulations. Solution-phase reaction coordinates could—in principle—be obtained using ab initio SCRF (self-consistent reaction field) methods[54,55] which include a continuum dielectric description of the solvent. In practice, however, we experienced convergence problems arising from the diffuse charge distribution of the anionic solute. More recent developments in continuum solvent methodology address this problem of outlying charge distributions.[56,57] For the purpose of this study, however, we used an appropriately constrained gas-phase reaction path that showed a reasonable approach of the hydroxide ion to the carbonyl function of formamide. Further details concerning the choice of reaction coordinates are given in sections 4.1−4.3.

3.2. Force Field Model. As noted above, the solution simulations require an accurate description of the solute's charge distribution along the reaction coordinate. Several charge models have been tested in combination with standard OPLS/AA[58] and

(45) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 3049−3050.

(46) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154−163.

(47) Madura, J. D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 2517−2527.

(48) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184−189.

(49) Carlson, H. A.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1996**, *118*, 8475−8484.

(50) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(51) Thiel, W. *Adv. Chem. Phys.* **1996**, *93*, 703−757.

(52) Explicit calculations using ab initio geometries and vibrational frequencies have indeed shown that (gas-phase) thermal corrections normally change very little *over the reaction coordinate* of the solute. However, noticeable differences in thermal corrections are found for reactions which involve bond formation between *separated reactants*. The loss of translational (and rotational) entropy disfavors complex (or bond) formation and requires free energy to be spent to constrain the reactants (here formamide and hydroxide) to a productive reaction pathway. For details, see Table A1 in the Supporting Information.

(53) Lim, D.; Jorgensen, W. L. *J. Phys. Chem.* **1996**, *100*, 17490−17500.

(54) Cramer, C. J.; Truhlar, D. G. *Rev. Comput. Chem.* **1995**, *6*, 1−72.

(55) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.

(56) Klamt, A.; Jonas, V. *J. Chem. Phys.* **1996**, *105*, 9972−9981.

(57) Chipman, D. M. *J. Chem. Phys.* **1997**, *106*, 10194−10206.

TIP4P[59] van der Waals parameters for the solute and the well-established TIP4P potential for the solvent.[60] The force field energy is defined as a sum of all *inter*molecular interactions:

$$E = \sum_i \sum_k \left(c\frac{q_i q_k}{r_{ik}} + 4\epsilon_{ik}\left(\frac{\sigma_{ik}^{12}}{r_{ik}^{12}} - \frac{\sigma_{ik}^6}{r_{ik}^6}\right) \right) \quad (4)$$

where q_i and q_k, r_{ik} (Å), and c denote atomic charges, the interatomic distance, and a factor (332.0636) to convert to kilocalories per mole, respectively. The van der Waals parameters ϵ_{ik} and σ_{ik} are obtained from atomic parameters using standard combining rules: $\epsilon_{ik} = (\epsilon_i \epsilon_k)^{1/2}$ and $\sigma_{ik} = (\sigma_i \sigma_k)^{1/2}$.

The force fields have been tested for their accuracy to reproduce the hydrogen-bonding patterns of solute−solvent interactions. Reference data have been obtained from ab initio (MP2/6-31+G*) geometry optimizations of selected solute−water complexes. For consistency with the static force field description, only intermolecular degrees of freedom have been optimized using fixed solute (MP2/6-31+G*) and solvent (TIP4P) geometries.

Among the most reliable charge models currently used are those which fit the charge-generated electrostatic potential to ab initio reference data.[61] The simplest procedure (ESP) involves a linear least-squares fit for reference points on surfaces 1.4−2.0 times the van der Waals surface of the molecule.[62] This method is known, however, to suffer from statistically ill-defined charges of buried atoms and tends to show unreasonably large variations with geometry. A more refined protocol (RESP, restricted ESP) subjects the least-squares fit to certain restraints which reduce the geometry dependence of ill-defined charges.[63,64] The ESP and RESP models have been tested applying both RHF/6-31+G* and MP2/6-31+G* wave functions.

For the statistical evaluation, we selected 11 hydrogen-bonded complexes for reaction I (formation of TET), 5 for reaction II (isomerization of TET), and 31 for the more complex reaction III (breakdown of TET). All electrostatic potential based charge models reproduce qualitative trends as predicted by ab initio calculations. Hydrogen bond energies are typically overestimated by 5−10% with respect to MP2/6-31+G* reference data, and by 20−30% if counterpoise corrections are taken into account. ESP/MP2 and RESP/MP2 perform only marginally better than the RHF-based models, and reduce the average errors from 5−12% (uncorrected) and 23−29% (corrected) to 4−11% and 19−27%, respectively.[65] These observations reflect the well-known tendency of the 6-31G* (and 6-31+G*) basis set to overestimate electric dipole moments.[32,61,66] It should be noted, however, that a moderate increase in polarity compensates for the multibody polarization neglected in effective two-body potentials, making the 6-31+G* basis set a popular choice for

Figure 4. Hydrogen bond energies (kcal/mol) for structures along the TET breakdown path calculated ab initio (MP2/6-31+G*) and with model potentials. Model potentials are based on ESP/HF charges (top) or QEq/PD charges (bottom). Hydrogen bonding sites are O-2 and O-7 (squares, type I, cf. Table 6) or O-7 and O-10 (crosses, type II, cf. Table 7).

potential-derived charges in classical simulations.[67,68] The well-established TIP3P and TIP4P potentials for water, e.g., reproduce bulk solvent properties very accurately, but they do so at the expense of a 20% increase in the water dimer energy.[59] Also common to all our model potential calculations is a consistent underestimation of hydrogen bond lengths by about 0.2 Å. This is another characteristic of effective two-body potentials such as TIP3P and TIP4P and no major deficiency, since qualitative trends are well reproduced.

While the ESP and RESP models provided satisfactory descriptions of the molecular charge distribution, we also explored simpler and less time-consuming empirical models. The charge equilibration model (QEq) proposed by Rappé and Goddard[69] and later implemented in a semiempirical framework[70] was found to be appropriate for many simple organic molecules and has been successful in applications of combined QM/MM potentials.[71] The semiempirical variant QEq/PD was specifically parametrized to reproduce ESP/HF charges at the 6-31G* level.[70] Unfortunately it performs very poorly in treating hydrogen bonds of structures along the TET formation pathway. Figure 4 shows the correlation between (uncorrected) MP2/6-31+G* and model potential (ESP/HF- and QEq/PD-based) results for these hydrogen bond energies. While the ESP/HF-

(58) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(60) The solute van der Waals parameters (σ (Å), ϵ (kcal/mol)) are C1 (3.75, 0.105), N4 (3.25, 0.17), O2 and O7 (2.96, 0.21), O10 (3.15358, 0.15504), H3 (2.42, 0.015), and all polar hydrogen atoms (0.0, 0.0).

(61) Williams, D. E. *Rev. Comput. Chem.* **1991**, *2*, 219−271.

(62) Besler, B. H.; Merz Jr., K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.

(63) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(64) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.

(65) Unsigned averages have been taken of the data for each of the three reactions. The given numbers show the spread in these averages. For more details, see Tables A2/A3 in the Supporting Information.

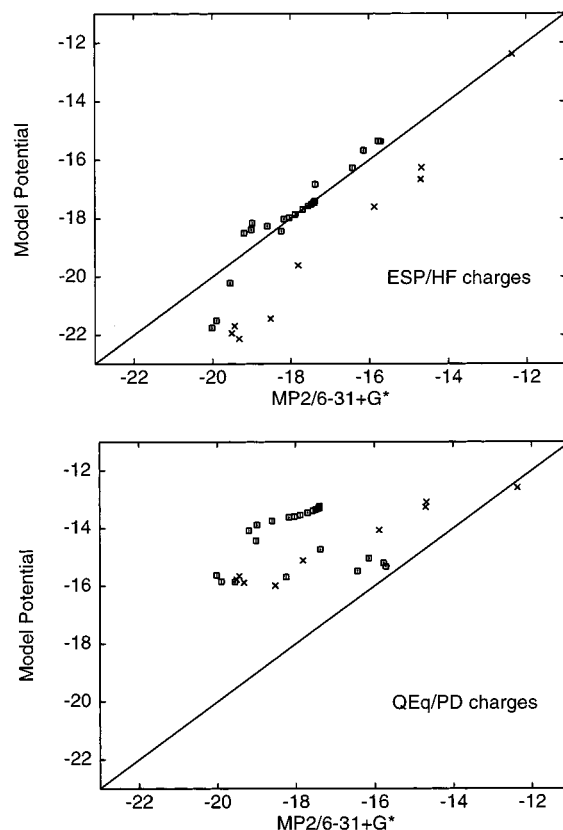(66) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304−323.

(67) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(68) Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461−469.

(69) Rappé, A. K.; Goddard, W. A., III. *J. Phys. Chem.* **1991**, *95*, 3358−3363.

(70) Bakowies, D.; Thiel, W. *J. Comput. Chem.* **1996**, *17*, 87−108.

(71) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.

based model potential follows the correct trends predicted by ab initio calculations, the correlation between QEq/PD and ab initio data shows considerable scatter. The failure of the QEq/PD-based model potential to reproduce qualitative trends can partly be attributed to an inappropriate distribution of excess charge over the atoms of ionic solutes. Simple reparametrizations of the charge model did not eliminate this problem. Hence, we decided to use the well-established ESP and RESP models which—as shown above—reproduce qualitative trends in hydrogen-bonding patterns. The use of restraints in charge derivation (RESP instead of ESP) was found to have little influence on the calculated hydrogen bond properties. The simpler ESP model was thus generally preferred, augmented by specific restraints only for reaction II where unphysical charge fluctuations produced considerable noise in the PMF (see section 4.2).

3.3. Details of the Simulations. The BOSS program (version 3.6)[72] and the force fields specified above were used to perform constant-pressure (NPT) Monte Carlo (MC) simulations for the solute immersed in a rectangular box of 260−505 water molecules (see sections 4.1−4.3). Periodic boundary conditions were applied together with a cutoff of 7.5−10.0 Å for all intermolecular interactions. In the MC simulations, new configurations were accepted or rejected on the basis of the Metropolis criterion, modified for preferential sampling in which the probability of attempting to move a solvent molecule was made proportional to $1/(r^2 + c)$.[31] Here c is a constant (150 Å^2) and r denotes solute−solvent distances as calculated to check the cutoff criterion (see sections 4.1−4.3). This modification affords enhanced sampling of solvent molecules which are in close proximity of the solute. Standard statistical perturbation theory[50] and double-wide sampling[73] were applied to the MC-generated ensembles to perturb between adjacent points ("windows") on the reaction coordinate and thus obtain the free energy of solvation as a potential of mean force (PMF). Data collection over 6M−24M configurations per window was preceded by equilibration over 1M configurations (first 0.5M under constant volume to prevent box expansion). The spacing between some of the windows afforded free energy differences of more than kT, which was considered to be too large for accurate calculations. In these cases, intermediate geometries were constructed by linear interpolation in **Z** matrix space. Following this procedure, all free energy differences between adjacent windows were below or around kT with standard deviations of less than 0.1 kcal/mol as evaluated by the batch means procedure (using sampling batches of 0.5M configurations).[31] The convergence of the calculated PMF was monitored by comparing simulations of different lengths (typically 2M, 4M, and 6M configurations per window) and by propagating the standard deviations over the entire reaction path. The latter method yielded errors of less than 0.3−0.4 kcal/mol for each of the reaction coordinates (assuming 6M configurations per window), which seemed too optimistic. The comparison of simulations of different lengths was thus given priority to judge the convergence of a calculation.

4. Results

This section is divided into three subsections discussing the MC results for the three consecutive reactions. Each of the subsections begins with a description of the reaction coordinate and of the specific details of the simulations.

(72) Jorgensen, W. L. *BOSS*, Revision 3.6; Yale University: New Haven, CT, 1995.

(73) Beveridge, D. L.; DiCapua, F. M. *Free Energy via Molecular Simulations: A Primer*; ESCOM Science Publishers: Leiden, 1989; pp 1−26.
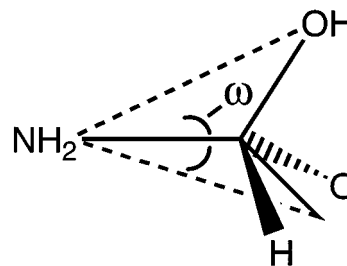

Figure 5. Direction of hydroxide attack in the formation of TET1. The angles ω and N4C1O7 have been fixed to values found for TET1.

Table 2. Formation of the Tetrahedral Intermediate: Hydrogen Bond Complexes with Water[a]

R_{C1-O7} (Å)	ab initio (MP2/6-31+G*) ref data[b]				model potential data[c]		
	E_{ref}	E_{ref} (CP)	R_{HW-O7}	R_{OW-H5}	E_{model}	R_{HW-O7}	R_{OW-H5}
1.538[d]	−15.55	−12.91	1.84	4.45	−16.63	1.63	2.25
1.6	−16.06	−13.40	1.83	4.38	−17.54	1.60	2.18
1.8	−17.84	−15.09	1.79	4.32	−20.88	1.56	2.09
2.0	−19.44	−16.93	1.76	4.26	−23.12	1.53	2.25
2.5	−23.51	−20.34	1.69	3.82	−30.72	1.50	2.05
3.0	−29.18	−25.37	1.66	2.05	−31.93	1.50	2.19
4.0	−30.99	−27.56	1.70	2.39	−31.44	1.50	2.67
5.0	−28.34	−25.57	1.74	3.10	−29.92	1.51	3.45
6.0	−26.85	−24.11	1.72	4.26	−29.13	1.50	5.06
7.0	−25.16	−22.30	1.71	9.87	−28.94	1.50	6.61
8.0	−25.23	−22.35	1.71	10.85	−28.79	1.50	7.92

[a] Only intermolecular degrees of freedom have been optimized, while MP2/6-31+G* and TIP4P geometries have been assumed for the reactant and the water molecule, respectively. Energies are in kilocalories per mole, and distances are in angstroms. [b] Hydrogen bond energies are reported with and without counterpoise (CP) correction. [c] Model potential data are based on ESP/HF-derived point charges using the 6-31+G* basis set. [d] TET1.

4.1. Formation of the Tetrahedral Intermediate. The formation of the tetrahedral intermediate between formamide and a hydroxide ion is in competition with the reversible proton abstraction from the amino group of formamide. Since TET1 formation is disfavored in the gas phase, a constrained reaction coordinate needs to be constructed which forces the hydroxide ion to attack the carbonyl group (see section 3.1). A series of geometry optimizations was thus performed for various fixed C1−O7 distances from 1.538 Å (as in TET1) to 8.0 Å, where the angles N4−C1−O7 and ω were constrained to the values calculated for TET1 (see Figure 5).

As discussed in the previous section, we used the ESP/HF model to generate individual sets of solute point charges for closely spaced points on the reaction coordinate. The custom-tailored model potential predicts hydrogen bond energies and geometries of solute−water complexes which are in satisfactory agreement with ab initio results (see Table 2). It reproduces the more favorable hydrogen bonding for intermediate C1−O7 distances around 3.0−4.0 Å, which is due to a strong HW−O7 (water−hydroxide) hydrogen bond supported by an attractive OW−H5 interaction. At longer C1−O7 distances, this additional stabilization is impossible, leaving the isolated hydroxide−water interaction as the most favorable hydrogen bond. At shorter C1−O7 distances, charge transfer from the hydroxide ion to formamide[74] causes the HW−O7 hydrogen bond to become much weaker and noticeably longer. The model potential, however, suggests a relatively close OW−H5 contact not only for intermediate, but also for short C1−O7 distances. This observation is not confirmed by the ab initio calculations which
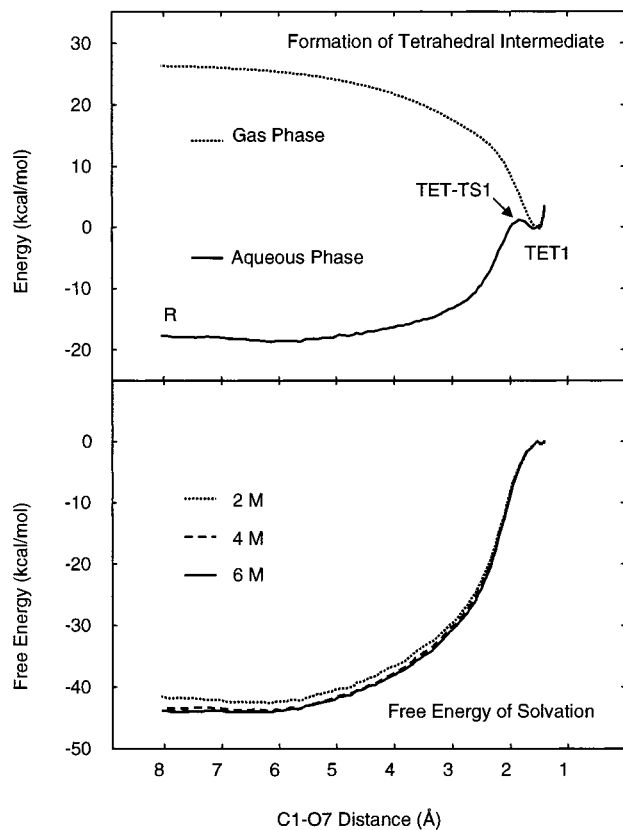
(74) See Table A4 in the Supporting Information.

Figure 6. Energy profile for the formation of the tetrahedral intermediate: (top) ab initio (MP2/6-31+G*) gas-phase energy and free energy in solution obtained from simulations over 6M configurations per window, (bottom) free energy of solvation obtained from 2M, 4M, and 6M configurations per window.

predict increasing OW−H5 distances for the latter part of the TET1 formation reaction. However, trends in hydrogen bond *energies* are in good agreement with the ab initio data, suggesting that the additional OW−H5 stabilization is less significant for short C1−O7 distances than it is for intermediate C1−O7 distances.

The free energy calculations performed for this reaction cover a total of 112 windows from $R_{C1O7} = 1.4$ Å to $R_{C1O7} = 8.0$ Å (box dimensions 20 Å × 20 Å × 30 Å, 395 water molecules, cutoff radius 7.5 Å). Good convergence is achieved after 4−6M configurations per window as can be seen from Figure 6. While TET1 formation is a steady downhill process in the gas phase, our results indicate a significant solvent-induced barrier of 20 kcal/mol in the aqueous phase. The apparent aqueous-phase transition state TET-TS1 is very close to the product TET1 and shows an almost entirely formed C1O7 bond (1.85 Å). In agreement with Hammond's postulate, TET-TS1 and TET1 are also very close in energy, differing by only 1.3 kcal/mol.

A closer analysis of the MC results sheds light on the mechanism by which solvation alters the energy profile of TET1 formation (see Figure 7 and Table 3[75]). The radial distribution functions (RDFs) for the O2−OW and O7−OW distances underline the clear preference for the reactants to form hydrogen bonds between the hydroxide ion and water. The first peak in the O7−OW RDF of the reactants integrates to 5.9 hydrogen bonds, compared to only 3.3 hydrogen bonds for TET1. Conversely, the product shows better solvation of the carbonyl oxygen O2 in formamide, with 3.9 hydrogen bonds integrated for the first peak, which compares to only 2.5 hydrogen bonds

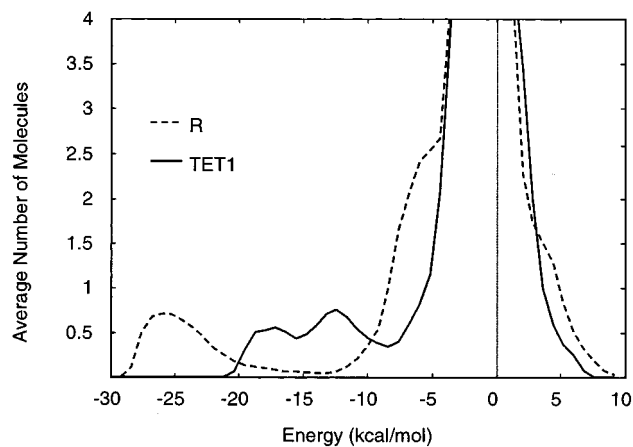(75) RDFs are also shown in Figure A1 (Supporting Information).

Figure 7. Calculated solute−solvent energy pair distributions for R (dashed) and TET1 (solid). Values are given in molecules per kcal/mol.

Table 3. Formation of the Tetrahedral Intermediate in Water: Analysis of the Monte Carlo Simulations[a,b]

			EPD		RDF	
					N_{HB}	N_{HB}
struct	ΔG_{solv}	ΔG_{tot}	N_{HB}^{min}	$N_{HB}^{-10.0}$	(OW−O2)	(OW−O7)
R	0.00	0.00	5.79	6.59	2.50	5.94
TET1	43.97	17.70	2.92[c]	7.11	3.85	3.31

[a] All results correspond to simulations performed with ESP/HF (6-31+G*) derived charges and are obtained from averaging over 6M configurations per window. ΔG_{solv} and ΔG_{tot} (see eq 3) are given in kilocalories per mole and refer to R_{C1-O7} distances of 8.0 Å (R) and 1.538 Å (TET1, gas-phase minimum), respectively. Using the solution-phase minimum of R at $R_{C1-O7} = 5.65$ Å results in a slightly higher free energy difference ($\Delta\Delta G_{tot} = 18.7$ kcal/mol). [b] N_{HB}^{min} and $N_{HB}^{-10.0}$ denote the number of hydrogen bonds calculated from the solute−solvent energy pair distribution (EPD) up to the first minimum and up to −10.0 kcal/mol, respectively. The first minimum appears at −15.6 kcal/mol for TET1 and at −14.0 kcal/mol for R. N_{HB}(OW − O2) and N_{HB}(OW − O7) give the number of hydrogen bonds calculated from the radial distribution functions (RDF) for OW−O2 and OW−O7 up to an O−O distance of 3.2 Å. [c] Only TET1 shows two minima in the energy pair distribution function. The area between the first and the second minima amounts to 4.91 hydrogen bonds, with a total of 7.83 hydrogen bonds up to the second minimum.

for the reactants. The N4−OW RDFs (not shown) depict a broad distribution and have much less structure, indicating that the amino group has little influence on the solvation pattern. This conclusion is supported by the overall similarity of the free energy profile with that for addition of hydroxide to *formaldehyde*.[47]

The solute−solvent energy pair distribution (EPD) function (Figure 7) confirms the importance of hydroxide−water interactions for the reactants. It shows a large peak at very high energies with noticeable distribution starting from −28.4 kcal/mol and with a maximum at −26.0 kcal/mol. The peak integrates to 5.8 hydrogen bonds, very similar to the number of hydrogen bonds in the first peak of the O7−OW RDF. This is clear evidence for all six strong hydrogen bonds being entirely formed between isolated hydroxide and water. A broad shoulder can be seen for the central peak of the reactant EPD, covering energies between −5 and −10 kcal/mol. Many of the interactions in this energy range are likely to correspond to hydrogen bonds between isolated formamide and water, which are expected to be 7−10 kcal/mol strong.[76] The EPD for TET1 shows a double peak centering around −17.2 and −12.4 kcal/mol which covers the stronger O2 and the weaker O7 hydrogen bonds (see Table 2; cf. section 4.2 and Table 4). Integration of the EPDs to −10.0

Base-Catalyzed Hydrolysis

J. Am. Chem. Soc., Vol. 121, No. 24, 1999 5719

Table 4. Isomerization of the Tetrahedral Intermediate: Hydrogen Bond Complexes with Water[a]

	ab initio (MP2/6-31+G*) ref data[b]				model potential data[c]		
		E_{ref}					
struct	E_{ref}	(CP)	R_{HW-O2}	$R_{HW-O7/N4}$	E_{model}	R_{HW-O2}	$R_{HW-O7/N4}$
TET1	−19.93	−17.03	1.88	2.50[d]	−20.98	1.62	2.60[d]
TET1	−18.71	−16.04	1.81	2.98[e]	−19.68	1.57	3.39[e]
TET1	−15.55	−12.91	3.67	1.84[e]	−14.88	3.84	1.66[e]
TET-T1	−21.21	−18.16	1.85	2.37[e]	−22.94	1.64	2.29[e]
TET	−21.12	−18.06	1.87	2.32[e]	−22.51	1.65	2.27[e]

[a] Only intermolecular degrees of freedom have been optimized, while MP2/6-31+G* and TIP4P geometries have been assumed for the reactant and the water molecule, respectively. Energies are in kilocalories per mole, and distances are in angstroms. [b] Hydrogen bond energies are reported with and without counterpoise (CP) correction. [c] Model potential data are based on combined fit RESP-derived point charges using the 6-31+G* basis set (see the text). [d] Hydrogen bond distance with N4. [e] Hydrogen bond distance with O7.

kcal/mol yields a total of 6.6 and 7.1 hydrogen bonds for the reactants R and for TET1, respectively. Apparently, the difference in solvation between R and TET1 is due only to hydrogen bond strengths and not to the number of hydrogen bonds. The localized negative charge in the hydroxide ion of the reactants enables the formation of six very strong hydrogen bonds. In TET1, however, this negative charge is distributed over the entire solute, causing much weaker hydrogen bonds with the hydroxide oxygen. This leads to much less favorable solvation of TET1 and is responsible for the solvent-induced activation barrier. These conclusions are completely in line with Madura and Jorgensen's MC study on the aqueous-phase reaction between formaldehyde and hydroxide.[47]

4.2. Isomerization of the Tetrahedral Intermediate. When the tetrahedral intermediate is formed in the gas phase, it adopts a synperiplanar O2C1O7H8 conformation which is stabilized by favorable dipole−dipole interaction (see Figures 2 and 8). Further stabilization is achieved through hyperconjugative p−σ* interaction between the nitrogen lone pair and the trans-oriented C1O7 bond, an interaction[77] commonly discussed as the origin of the anomeric effect.[78] While TET1 appears to be the lowest-energy conformer, it does not allow water-assisted C−N bond cleavage. To proceed, the reaction requires isomerization to a different conformer TET with antiperiplanar O2C1O7H8 and O2C1N4(lp) units (Figures 2 and 8). It is interesting to note that this new isomer not only provides the correct conformation for attack by water, but also shows the longer C−N bond (1.525 Å vs 1.490 Å). This observation agrees with an earlier analysis of stereoelectronic effects for hydrated formaldehyde,[79] which found polar bonds to be longer and weaker if they were in antiperiplanar (app) orientation to lone pairs, but shorter and stronger if one of their atoms had a lone pair oriented app with respect to an adjacent polar bond. The former of the isomers (TET1) shows a nitrogen lone pair app to C1−O7 and a shorter C−N bond, while the latter (TET) exhibits an O7 lone pair app to C−N and hence a longer C−N bond. Lehn and Wipff suggested that those tetrahedral intermediates with longer and weaker C−N bonds were kinetically more labile and thus more susceptible to hydrolysis.[79] The more facile hydrolysis of polar

(76) D. Bakowies and P. A. Kollman, unpublished results (MP2/6-31+G* calculations for formamide−water complexes, no counterpoise correction). For *N*-methylacetamide−water complexes, see also: Dixon, D. A.; Dobbs, K. D.; Valentini, J. J. *J. Phys. Chem.* **1994**, *98*, 13435−13439.

(77) Hoffmann, R.; Radom, L.; Pople, J. A.; Schleyer, P. v. R.; Hehre, W. J.; Salem, L. *J. Am. Chem. Soc.* **1972**, *94*, 6221−6223.

(78) David, S.; Eisenstein, O.; Hehre, W. J.; Salem, L.; Hoffmann, R. *J. Am. Chem. Soc.* **1973**, *95*, 3806−3807.

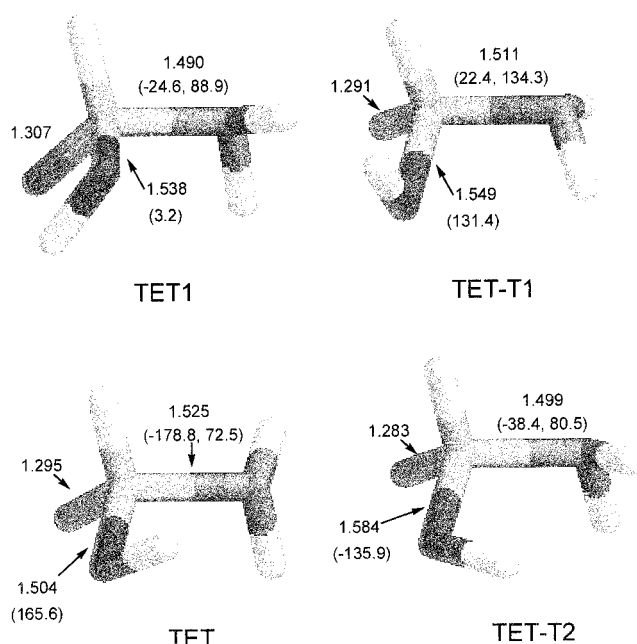(79) Lehn, J. M.; Wipff, G. *J. Am. Chem. Soc.* **1974**, *96*, 4048−4050.

Figure 8. Geometries of various conformations of the tetrahedral intermediate. Atom types are coded as follows: H (white), C (light gray), N (medium gray), O (dark gray). Selected MP2/6-31+G* bond lengths and dihedral angles (H8O7C1O2, H5/6N4C1O7) are given.

bonds weakened by antiperiplanar lone pairs has also been observed experimentally.[80]

In the gas phase, the conformational isomerization of TET1 to TET requires 4.8 kcal/mol and proceeds over a barrier of 7.4 kcal/mol (MP2/6-31+G* data, Table 1). The transition state TET-T1 shows rotations of the nitrogen lone pair and of the O7H8 bond, both of which adopt eclipsed orientations with respect to the C1H3 bond. The complicated nature of the isomerization involving two independent rotations precludes the use of a single torsional angle to monitor the reaction. Instead, we chose to use the intrinsic reaction coordinate (IRC) in mass-weighted coordinates.[38] Further investigation of the reaction coordinate revealed a second transition state TET-T2 showing eclipsed O7H8 and C1N4 bonds and an antiperiplanar orientation of the nitrogen lone pair with respect to C1O7 (Figure 8). As confirmed by IRC calculations, this transition state connects two identical copies of TET1 via an unsymmetrical pathway. Forming a closed thermodynamic cycle, this reaction path provides an excellent test for the convergence behavior of our PMF calculations.

Initial MC simulations using ESP/HF-derived solute charges were unsatisfactory due to considerable noise in the computed PMF. This problem resulted from unphysical charge fluctuations along the reaction path, which could be attributed to the artificial sensitivity of the ESP protocol to molecular geometry and orientation. The application of one constant set of point charges certainly eliminates the problem and seems justified for this particular reaction, which involves no changes in covalency. We decided to use RESP-derived point charges which were constrained to give the best simultaneous fit to the electrostatic potentials of TET and TET1. There were no significant differences in accuracy between the ESP- and RESP-based model potentials to describe solute−water interactions. Table 4 details some of the results obtained with the RESP-based model potential. Three different solute−water complex geom-

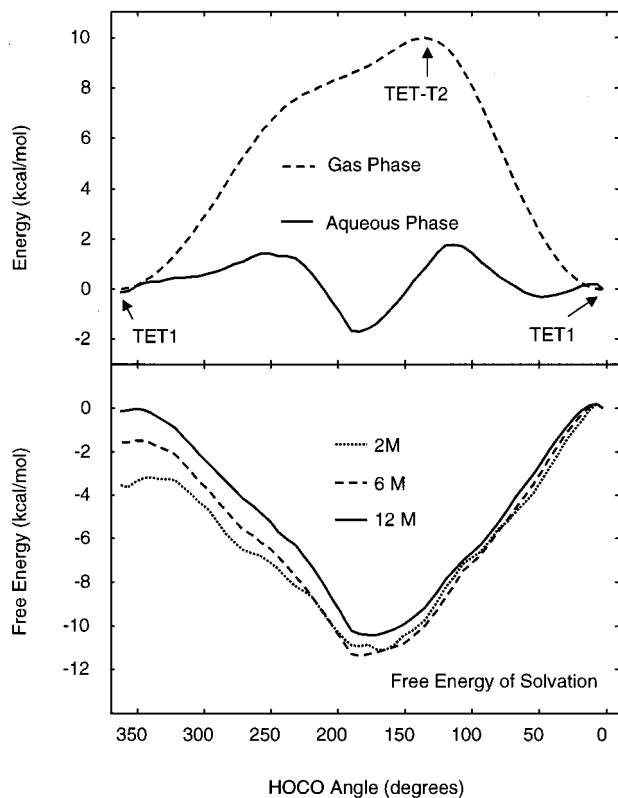(80) Deslongchamps, P.; Lebreux, C.; Taillefer, R. *Can. J. Chem.* **1973**, *51*, 1665−1669.

Figure 9. Energy profile for the pathway connecting two identical copies of TET1 over an unsymmetrical transition state TET-T2: (top) ab initio (MP2/6-31+G*) gas-phase energy and free energy in solution obtained from simulations over 12M configurations per window, (bottom) free energy of solvation obtained from 2M, 6M, and 12M configurations per window.

etries have been chosen for TET1, with a water hydrogen binding to either O2 or O7, and one complex each for TET-T1 and TET, with a water hydrogen binding to O2. As can be seen from Table 4, the model potential nicely reproduces trends in energies and geometries. Hydrogen bonds between the anionic O2 and HW are preferred over those between the hydroxyl O7 and HW. The strongest hydrogen bonds are formed by TET-T1 and TET, due to their O2C1O7H8 conformations which allow stabilizing contacts between the second water hydrogen and O7 (cf. Figure 8).

The agreement between model potential and ab initio results justified the use of RESP charges for all subsequent Monte Carlo simulations. Solvent effects were studied for the main path, leading from TET1 to TET, and, additionally, for the auxiliary path which connects two identical copies of TET1. The simulations were performed for solvent boxes containing 262 water molecules (box dimensions 20 Å × 20 Å × 20 Å, cutoff radius 7.5 Å), and covered a total of 20 and 30 windows, respectively.

Judging from the results for the auxiliary path (TET1 → TET1, Figure 9), the simulations converge slower than expected. Closure of the thermodynamic cycle to within 1 kcal/mol is achieved only after 8 million configurations per window.[81] A further increase to 12 million configurations per window, however, cuts the residual error by almost an order of magnitude. The free energy differences between TET1 and TET-T1 or TET on the main path vary much less with the length of the simulation, which can be attributed to the shorter pathway and the smaller gas-phase barrier. However, the data do not suggest

---

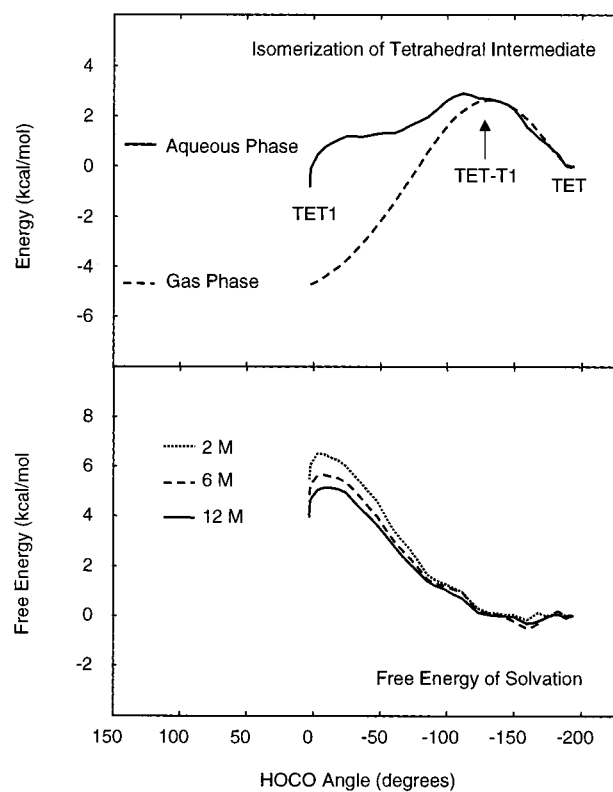(81) See Table A8 in the Supporting Information.

Figure 10. Energy profile for the isomerization of TET1: (top) ab initio (MP2/6-31+G*) gas-phase energy and free energy in solution obtained from simulations over 12M configurations per window, (bottom) free energy of solvation obtained from 2M, 6M, and 12M configurations per window.

a fully converged number for the free energy difference between TET1 and TET. We can only estimate the overall accuracy for a given length of simulation by judging from the results obtained for the auxiliary pathway. With this in mind, we should expect reasonable results for simulations of 6 million configurations per window, but accurate results only for simulations twice as long.

Figure 10 summarizes our results. Solvation by water favors TET over TET1, diminishing the gas-phase preference for TET1. However, the computed free energy curve shows a steep slope in the vicinity of TET1, indicative of significant structural changes upon solvation. The true minimum in solution is apparently not covered by the gas-phase reaction coordinate. Still it seems reasonable to assume that optimization of only the hydroxy group and nitrogen lone pair orientations suffices to determine the solution-phase structure. We have thus performed additional free energy calculations on a two-dimensional grid with the torsional angles H8O7C1O2 and H6N4C1O2 defining the two grid axes.[82]

The results are summarized in contour diagrams of the (free) energy around the gas- and solution-phase minima of TET1 (Figure 11). An interesting observation is the high degree of complementarity between the gas-phase energy and the free energy of solvation. Regions of low gas-phase energy correspond to less favorable solvation (H6N4C1O2 close to 0°), while regions of high gas-phase energy correspond to more favorable solvation (H6N4C1O2 close to −90°). The resulting free energy landscape in aqueous solution is relatively flat but shows two minima for intermediate H6N4C1O2 angles around −50°, one corresponding to positive H8O7C1O2 angles, the other corresponding to negative H8O7C1O2 angles. The precise orientation of the hydroxy group appears to be less critical than
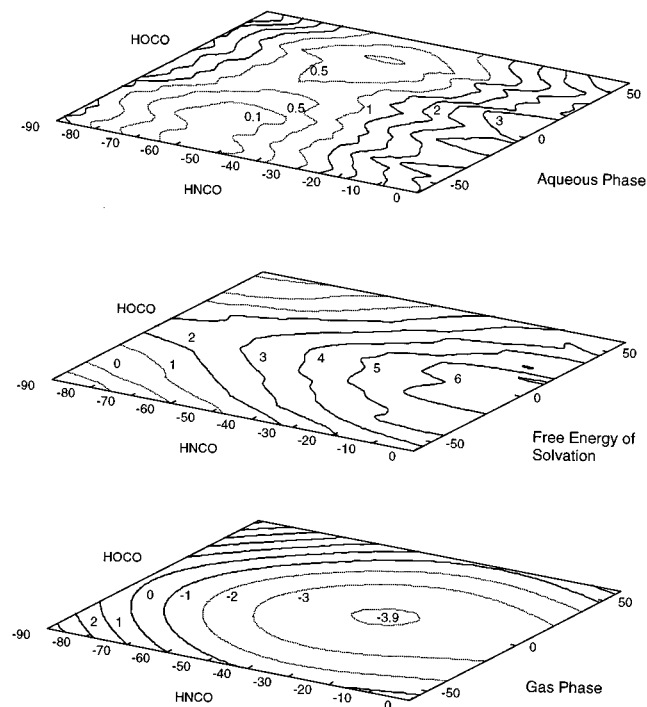
Figure 11. Energy contour diagrams for the tetrahedral intermediate. Gas-phase energy (bottom), free energy of solvation (middle), and free energy in solution (top) for conformations of the tetrahedral intermediate with H6N4C1O2 and H8O7C1O2 in the range of −90° to 0° and −60° to +60°, respectively. Spacings between adjacent contour lines are 0.5 kcal/mol for the top graph and 1.0 kcal/mol for the middle and bottom graphs. Dotted contour lines are used for low-energy regions.

the orientation of the nitrogen lone pair, which, compared to the gas-phase structure, is more perfectly antiperiplanar with respect to the C1O7 bond. The preference for this geometry is likely due to the more facile formation of bidentate complexes with water (O2 and N4 as acceptors) and to reduced sterical congestion (caused by the amino hydrogens).

The solvent-related geometric distortion stabilizes TET1 by about 1 kcal/mol. This result is in good agreement with ab initio geometry optimizations which treat the solvent as a dielectric continuum ($\epsilon = 78.4$).[83] Taking the geometric relaxation into account, the isomerization of TET1 to TET becomes slightly more endothermic (ca. 2 kcal/mol), but it is still more favorable

---

(82) Ab initio (MP2/6-31+G*) geometry optimizations have been carried out with these two angles constrained to their grid values. The resulting grid with a 20° spacing in each direction was then refined by interpolation between adjacent grid points, yielding a final spacing of 5°. Solution simulations have been carried out for given H8O7C1O2 angles to obtain one-dimensional potentials of mean force (PMFs) along the H6N4C1O2 coordinate. A total of three one-dimensional PMFs along the H8O7C1O2 coordinate were calculated at given H6N4C1O2 angles to be combined with the former results and to obtain a two-dimensional PMF. Strictly speaking, only one such "orthogonal" one-dimensional PMF is required to construct the two-dimensional grid, and we have chosen to use a value of H6N4C1O2 close to the expected minimum (−60°). The calculation of two additional orthogonal PMFs at either end of the H6N4C1O7 scale, however, allows us to evaluate the accuracy to which the thermodynamic cycles are closed. The solution simulations have been performed for 6 million configurations per grid point and cover the H8O7C1O2 and H6N4C1O2 angles from −90° to +30° and from −60° to +60°, respectively. On average, thermodynamic cycles formed from adjacent points on the HOCO axis and from reference (−60°) and end points on the HNCO axis are closed to within 0.3−0.4 kcal/mol (maximum error 1.3 kcal/mol). This allows for a semiquantitative determination of free energy minima and is sufficient for our purposes.

(83) A COSMO[84] geometry optimization has been carried out at the MP2/ 6-31+G* level, starting from the gas-phase structure of TET1. The energy difference between the initial and optimized structures is 1.0 kcal/mol. The optimized structure corresponds to a H6N4C1O2 angle of −51°, in close agreement with the simulation result.

Table 5. Isomerization of the Tetrahedral Intermediate in Water: Analysis of the Monte Carlo Simulations[a]

| | | | EPD | | RDF | |
| | | | | | N_{HB} | N_{HB} |
struct	ΔG_{solv}	ΔG_{tot}	N_{HB}^{min}	$N_{HB}^{-10.0}$	(OW−O2)	(OW−O7)
TET1	0.00	0.00	2.93	6.40	3.54	2.83
TET-T1	−3.90	3.53	3.73	6.82	3.81	3.49
TET	−3.95	0.83	3.30	6.69	3.82	2.99
TET-T2	−9.17	0.82	4.00	7.06	4.05	3.28

[a] All results correspond to simulations performed with charges from a combined RESP fit (see the text) and are obtained from averaging over 12M configurations per window. ΔG_{solv} and ΔG_{tot} (see eq 3) are given in kilocalories per mole and refer to optimized gas-phase geometries. N_{HB}^{min} and $N_{HB}^{-10.0}$ denote the number of hydrogen bonds calculated from the solute−solvent energy pair distribution (EPD) up to the first minimum and −10.0 kcal/mol, respectively. The first minimum appears at −15.6 kcal/mol for TET, at −14.8 kcal/mol for TET1 and TET-T2, and at −14.0 kcal/mol for TET-T1. N_{HB}(OW−O2) and N_{HB}(OW−O7) give the number of hydrogen bonds calculated from the radial distribution functions (RDF) for OW−O2 and OW−O7 up to an O−O distance of 3.2 Å.

than in the gas phase (4.8 kcal/mol; see above). Hence, the general picture remains unchanged: Solvation reduces the gas-phase preference for TET1 and lowers its barrier to isomerization. The antiparallel orientation of the C1O2 and O7H8 dipoles in TET1 affords internal stabilization, but also reduces hydrogen bond strengths due to steric congestion of O7 as the second acceptor in a bidentate (O2 and O7) solute−water complex. This can be seen from Table 4, which lists the hydrogen bond energies and geometries of various optimized solute−water complexes. The generally most favorable complex geometry shows a strong hydrogen bond between a water hydrogen and O2 of the solute and a weaker interaction between the other water hydrogen and O7. The latter distance is considerably longer for TET1 than for TET-T1 and TET, rationalizing the lower hydrogen bond energy of TET1. The generally reduced hydrogen bond strength for complexes of TET1 may explain its less favorable solvation. An analysis of our MC simulations supports this assumption (see Table 5): The solute−solvent energy pair distribution (EPD) functions show fewer strong hydrogen bonds for TET1 (2.9, integrated up to the first minimum) than for TET (3.3) and TET-T1 (3.7). There are also fewer short OW−O2 or OW−O7 contacts for TET1, as indicated by the radial distribution functions (RDFs).

Of all structures considered, TET-T2 is the highest in gas-phase energy and the most favorably solvated. Both the EPD and RDF data indicate very strong hydrogen bonds, reflecting the simultaneous exposure of all three hydrogen bond acceptors (O2, N4, O7). Such geometry is unfavorable in the gas phase, but enables the formation of *two different* bidentate complexes with water (O2, O7 and O2, N4). In summary, we may conclude that the solution geometry of the tetrahedral intermediate is determined by the balance between two competing effects: internal stability (electrostatic and orbital overlap) and solvent accessibility.

4.3. Breakdown of the Tetrahedral Intermediate. In our gas-phase model, a single water molecule catalyzes the breakdown of TET. The reaction is initiated by proton donation to the nitrogen atom, followed by C−N bond cleavage and proton abstraction from the carboxyl group. The water molecule is thus regenerated (Figures 2 and 12). Interestingly, all three steps are covered by a single reaction coordinate, as confirmed by IRC calculations. There is only one transition state for the entire reaction, and it corresponds to a structure with an almost formed N4−H9 bond (1.139 Å, MP2/6-31+G*) and a still intact C1−
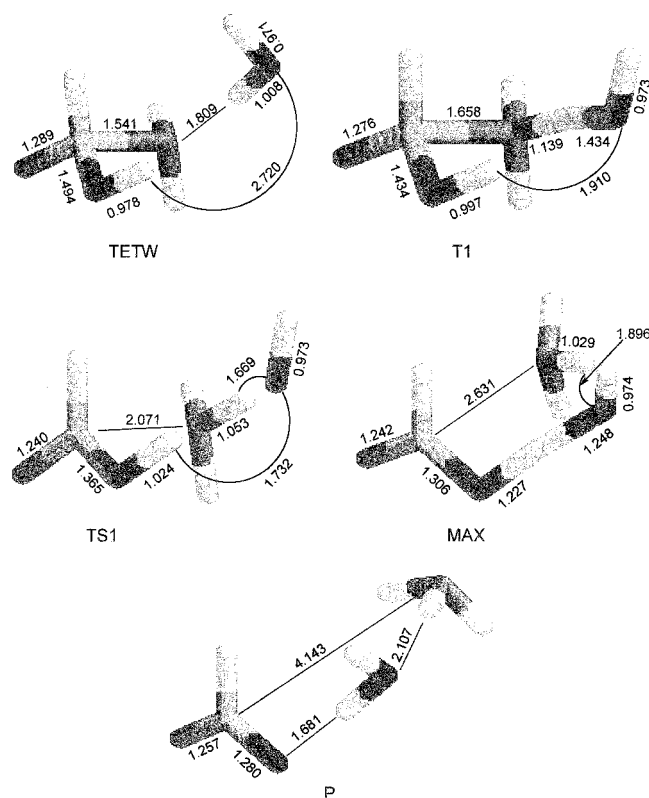
Figure 12. Geometries of various structures along the TET breakdown pathway. Atom types are coded as follows: H (white), C (light gray), N (medium gray), O (dark gray). Selected MP2/6-31+G* bond lengths are given.

N4 bond (1.658 Å, MP2/6-31+G*). The transition state may be considered early as both the C−N bond cleavage and the subsequent proton transfer correspond to the downhill section of the reaction coordinate. In line with Hammond's postulate, the overall reaction is very exothermic.

Such a complex reaction cannot be described by monitoring a single internal coordinate, and the IRC[37,38] would be the natural choice for a suitable reaction coordinate. Unfortunately, parts of the reaction coordinate are so flat that many numerical problems are encountered when the IRC is calculated. We were successful at low levels of theory (HF/3-21+G), but obtained only part of the IRC at the desired level (MP2/6-31+G*). The calculated sections of the IRC were sufficient, however, to obtain the corresponding minima by extrapolation and subsequent geometry optimization. Finally, we chose a combination of the N4−H9 coordinate (TET → T1) and of the Cartesian IRC (T1 → P) to define the reaction coordinate used for the solution simulations.

ESP/HF-derived point charges were used to describe the solute's charge distribution as a function of the reaction coordinate. The quality of the force field was assessed in calculations of two different hydrogen-bonded complexes with water: Type I refers to complexes with O2 and O7 as hydrogen bond acceptors, while type II refers to complexes with O7 and O10 as hydrogen bond acceptors. The results are compared with predictions at the MP2/6-31+G* level (Figure 4, Tables 6 and 7). The model potential data are in close agreement with the raw ab initio results, but as expected, they overestimate the counterpoise-corrected reference values. They also tend to overestimate the stabilization of a principal hydrogen bond by a second contact. This is apparent from the very short HW−O7 distances for type II hydrogen bonds, but the effect is only moderate in energy (Table 7). The overall performance of the

Table 6. Breakdown of the Tetrahedral Intermediate: Type I Hydrogen Bond Complexes with Water[a,b]

| | ab initio (MP2/6-31+G*) ref data[c] | | | | model potential data[d] | | |
struct	E_{ref}	E_{ref} (CP)	R_{HW-O2}	R_{HW-O7}	E_{model}	R_{HW-O2}	R_{HW-O7}
TETW	−20.02	−16.88	1.85	2.43	−21.76	1.65	2.31
T1	−19.55	−16.09	1.86	2.46	−20.22	1.66	2.38
TS1	−16.43	−13.33	1.96	2.41	−16.27	1.74	2.45
MAX	−17.37	−14.16	2.07	2.23	−16.83	1.78	2.38
P	−17.39	−14.87	2.01	2.31	−17.41	1.80	2.29

[a] Only intermolecular degrees of freedom have been optimized, while MP2/6-31+G and TIP4P geometries have been assumed for the reactant and the water molecule, respectively. Energies are in kilocalories per mole, and distances are in angstroms. [b] All complexes show one major hydrogen bond involving one water hydrogen atom and a further stabilizing contact involving the other water hydrogen atom. [c] Hydrogen bond energies are reported with and without counterpoise (CP) correction. [d] Model potential data are based on ESP/HF-derived point charges using the 6-31+G* basis set.*

Table 7. Breakdown of the Tetrahedral Intermediate: Type II Hydrogen Bond Complexes with Water[a,b]

| | ab initio (HF/6-31+G*) ref data[c] | | | | model potential data[d] | | |
struct	E_{ref}	E_{ref} (CP)	R_{HW-O7}	R_{HW-O10}	E_{model}	R_{HW-O7}	R_{HW-O10}
TETW	−12.36	−10.04	6.28	1.87	−12.39	5.44	1.71
T1	−18.52	−15.57	4.27	1.77	−21.44	2.03	1.66
TS1	−19.51	−16.64	3.96	1.77	−21.94	2.18	1.64
MAX	−17.81	−14.72	2.26	1.97	−19.61	2.16	1.72
P	−14.68	−12.12	1.96	2.56	−16.26	1.78	2.00

[a–d] See footnotes a–d in Table 6.

model potential is still satisfactory as important trends observed for hydrogen bond strength and geometry are well reproduced.

Monte Carlo simulations have been performed for a total of 81 windows covering the reaction coordinate from TETW to P (box dimensions 20 Å × 20 Å × 20 Å, 260 water molecules, cutoff radius 7.5 Å). The results are summarized in Figure 13. Apparently, solvation has rather little influence on the energy profile. The reaction energy is reduced from −34 kcal/mol (MP2/6-31+G*, Table 1) to −26 kcal/mol, but the activation barrier is almost unaffected. More conspicuously, the transition state becomes rather product-like, which explains the reduced exothermicity. The solution-phase transition structure (TS1; see Figure 12) corresponds to an intermediate stage between proton donation (O10−H9 = 1.67 Å) and proton abstraction (O10−H8 = 1.73 Å), and it shows a completely broken C1−N4 bond (2.07 Å). Interestingly, the gas-phase maximum is the most favorably solvated structure and appears to be a minimum in solution.

The Monte Carlo calculations show satisfactory convergence (Figure 13), and the results are quite insensitive to the details of the simulation. Neither the use of ESP/MP2-derived point charges (instead of ESP/HF) nor the increase in box size and cutoff radius changes the results significantly.[85] The latter observation is quite important since errors in absolute free energies of solvation are expected to be large for charged solutes if finite cutoffs are applied. The results indicate, however, that relative free energies are almost unaffected due to efficient error cancellation.

(84) Baldridge, K.; Klamt, A. *J. Chem. Phys.* **1997**, *106*, 6622−6633.
(85) Two additional simulations have been performed, one with ESP/MP2 charges (instead of ESP/HF) and one with a larger solvent box (box dimensions 25 Å × 25 Å × 25 Å, 505 instead of 260 water molecules, cutoff radius 10.0 Å instead of 7.5 Å). The calculated free energies are very similar to the ones reported here, with deviations of less than 1 kcal/mol (TETW → TS1) and 3 kcal/mol (TS1 → P), respectively. For further details, see Table A9 in the Supporting Information.
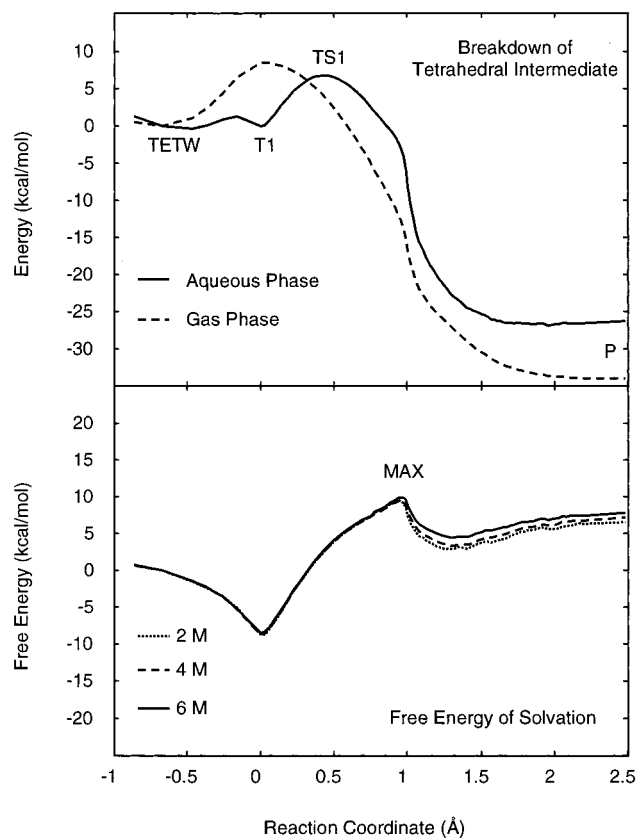
Figure 13. Breakdown of the tetrahedral intermediate in the gas phase and in aqueous solution: (top) ab initio (MP2/6-31+G*) gas-phase energy and free energy in solution obtained from simulations over 6M configurations per window, (bottom) free energy of solvation obtained from 2M, 4M, and 6M configurations per window. The reaction coordinate is expressed as $R = -r_{\mathrm{N4H9}} + r^0_{\mathrm{N4H9}}$ for $R < 0$ and as $R = r_{\mathrm{C1N4}} - r^0_{\mathrm{C1N4}}$ for $R > 0$ where r^0_{N4H9} and r^0_{C1N4} denote the N4H9 and C1N4 distances of the transition structure T1.

The solvent effects seen in Figure 13 are readily explained with changes in the solute's electronic structure: The first proton transfer is complete as soon as the reaction coordinate approaches its maximum, producing an isolated hydroxide ion for both the gas- and solution-phase transition states T1 and TS1. The second proton transfer follows TS1 and is partially complete for MAX (O7−H8 = 1.23 Å, H8−O10 = 1.25 Å). Consequently, there is a high charge density at O10 for both T1 and TS1,[86] which results in strong and short hydrogen bonds with water (type II complexes; see Table 7). This is reflected in the O10−OW radial distribution functions which show a higher peak at short distances for T1 than for TETW or MAX.[87]

The C−N bond cleavage accounts for a second important change in the solute's electronic structure: In the beginning, the carbonyl oxygen O2 carries a formal negative charge which diminishes in size as soon as the C−N bond cleaves and a formally neutral formic acid unit is formed. This simple valence bond picture corresponds to a decrease in charge transfer with increasing C−N distance, which results in a less polar charge distribution for the carboxyl unit. The C−N bond is still intact for T1, but it is cleft for TS1, MAX, and P. Consequently, the O2 charge (ESP/HF) decreases sharply from T1 to TS1,[86] resulting in much weaker type I hydrogen bond complexes for TS1 (Table 6). Upon formation of a carboxylate anion (in MAX and P), the O2 charge increases again, but it does not reach its

(86) For more details, see Table A6 in the Supporting Information.
(87) RDFs are shown in Figure A2 (Supporting Information).

Table 8. Breakdown of the Tetrahedral Intermediate in Water: Analysis of the Monte Carlo Simulations[a]

			EPD		RDF		
					N_{HB}	N_{HB}	N_{HB}
struct	ΔG_{solv}	ΔG_{tot}	$N^{\mathrm{min}}_{\mathrm{HB}}$	$N^{-10.0}_{\mathrm{HB}}$	(OW−O2)	(OW−O7)	(OW−O10)
TETW	0.00	0.00	3.33[b]	6.19	4.00	2.98	3.70
T1	−8.48	−0.02	7.79	7.45	3.58	2.45	3.38
TS1	2.44	6.74	8.39	6.88	3.39	2.17	3.96
MAX	9.87	−2.05	8.18	6.65	3.53	2.21	3.28
P	7.84	−26.28	8.46	6.62	3.42	2.16	2.51

[a] All results correspond to simulations performed with ESP/HF (6-31+G*) derived charges and are obtained from averaging over 6M configurations per window. ΔG_{solv} and ΔG_{tot} (see eq 3) are given in kilocalories per mole and refer to optmized gas-phase geometries (TETW, T1, P), and maxima of ΔG_{solv} (MAX) and of ΔG_{tot} (TS1), respectively. $N^{\mathrm{min}}_{\mathrm{HB}}$ and $N^{10.0}_{\mathrm{HB}}$ denote the number of hydrogen bonds calculated from the solute−solvent energy pair distribution (EPD) up to the first minimum and up to −10.0 kcal/mol, respectively. The first minimum appears at −14.0 kcal/mol for TETW, at −9.2 kcal/mol for T1, and at −7.6 kcal/mol for TS1, MAX, and P. N_{HB}(OW−O2), N_{HB}(OW−O7), and N_{HB}(OW−O10) give the number of hydrogen bonds calculated from the radial distribution functions (RDF) for OW−O2, OW−O7, and OW−O10 up to an O−O distance of 3.2 Å. [b] Only TETW shows two minima in the energy pair distribution function. The area between the first and the second minima amounts to 4.88 hydrogen bonds, with a total of 8.21 hydrogen bonds up to the second minimum.

initial value. Type I hydrogen bonds are thus strong for TETW and T1, weak for TS1, and of intermediate strength for MAX and P (Table 6). This analysis is supported by the O2−OW radial distribution functions which show preferred hydrogen bonding for TETW and T1.[87]

The role of O7 as a potential hydrogen bond acceptor is less significant than that of O2 and O10. Judging from the radial distribution functions, there are only 2−3 water molecules in the first solvation shell of O7, compared to 3−4 for O2 and O10 (see Table 8). Geometry optimization of solute−water complexes generally yields O2 and O10 as primary binding sites, with much longer hydrogen bond distances for O7 (see Tables 6 and 7). Only in the late stages of the reaction, there is noticeable competition between O2 and O7 as hydrogen bond acceptor, reflecting the balanced charge distribution in isolated formate anions. Like O7, the nitrogen atom also plays a minor role as hydrogen-bonding site. The N4−OW RDFs exhibit little structure except for the product which contains an isolated ammonia molecule. In the earlier stages of the reaction, however, the nitrogen atom is buried, giving rise to a very broad plateau in the RDF (not shown).

In summary, O2 is a particularly strong hydrogen bond acceptor for TETW and T1, while O10 shows strong hydrogen bonding in T1 and TS1. The gas-phase transition state T1 is the only structure characterized by two strong hydrogen-bonding sites, rationalizing its more favorable solvation. The energy pair distributions for TETW, T1, and MAX support this explanation (Figure 14): MAX is the least favorably solvated structure and exhibits only one peak at relatively low energies (−12.4 kcal/mol), which covers both type I and type II hydrogen bond complexes (Tables 6 and 7). TETW is more favorably solvated and shows two peaks at −18.0 and −10.0 kcal/mol, which most likely correspond to strong type I and weaker type II hydrogen bonds. T1 is the most favorably solvated structure with one peak at −15.6 kcal/mol covering both type I and type II hydrogen bond complexes. The peak for T1 is at slightly lower energies than the first peak for TETW, but covers more hydrogen bonds (integrated 7.8 vs 3.3). The total number of hydrogen bonds up to the lowest energy minimum is very similar for each of the structures (around 8), suggesting that the strength rather than the number of hydrogen bonds determines the solvent effect.
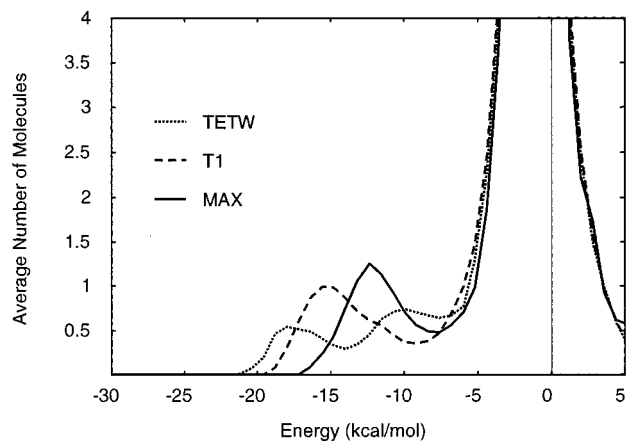
Figure 14. Calculated solute−solvent energy pair distributions for TETW (dotted), T1 (dashed), and MAX (solid). Values are given in molecules per kcal/mol.

5. Discussion and Conclusions

The base-catalyzed hydrolysis of formamide has been studied for the gas phase and for aqueous solution, following a protocol previously suggested by Jorgensen. The theoretical approach is very simple in concept, and combines the quantum chemical treatment of a gas-phase reaction with the completely classical simulation of solvent effects. This allows for a high-level ab initio treatment of the solute and a rigorous statistical mechanical description of solvent effects, which would not be feasible with hybrid quantum and molecular mechanical approaches.[88] We have thus been able to study the solution-phase reaction with higher accuracy and in much more detail than was possible in the past.

Our gas-phase model comprises a sequence of three consecutive reactions: formation, isomerization, and breakdown of a tetrahedral intermediate. The last reaction is catalyzed by a single solvent molecule which transfers a proton from the carboxyl oxygen to the amino nitrogen. It is remarkable that this double proton transfer and the C−N bond cleavage are covered by a single reaction with only one transition state.

Figure 15 shows the gas- and aqueous-phase energy profiles for the complete sequence from reactants to products. The offset between the gas-phase reactions of isomerization and breakdown accounts for the formation of a hydrogen bond between TET and water (TET + H_2O → TETW; see Table 1). This distinction between TET and TETW should become unnecessary in aqueous solution. In practice, however, the single water molecule in TETW has been constrained and is thus treated in a manner different from that of the water molecules surrounding TET and TETW. This gives rise to a constraint or "cratic" free energy contribution which has not been considered so far.[89] We can expect, however, that this free energy contribution is fairly small since TETW is a strong hydrogen-bonded complex which would remain close to its ideal geometry even in an unconstrained simulation. On the other hand, the numerical determination of
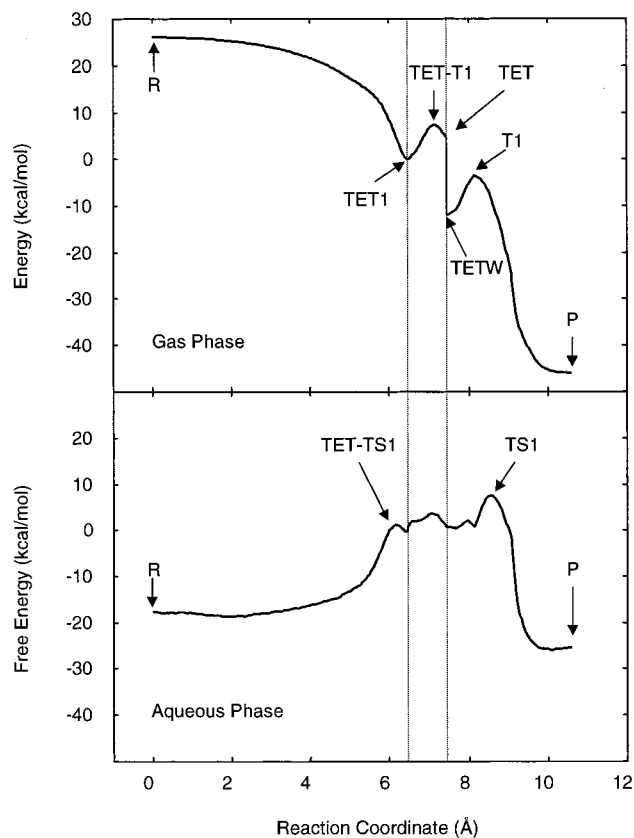

Figure 15. Overall energy profile for formamide hydrolysis in the gas phase and in aqueous solution. The graphs for reactions I (formation of TET1), II (isomerization), and III (breakdown) are separated by vertical lines. One unit on the abscissa corresponds to 1 Å (reactions I and III) or 100° (reaction II). See the text for discussion.

the constraint free energy is far from straightforward and probably not accurate enough to improve the quality of our results. Hence, we decided to neglect this correction.

According to Figure 15, the gas-phase reaction is very exothermic and involves two barriers of approximately 8 kcal/mol each (MP2/6-31+G*; see Table 1). The initial formation of a tetrahedral intermediate is the only process without a barrier and proceeds with a net energy gain of ca. 27 kcal/mol (MP2/6-31+G*; see Table 1). Nevertheless, this first step is in competition with the even more favorable (by 20.6 kcal/mol, MP2/6-31+G*) acid−base reaction between formamide and hydroxide. In the gas phase, the latter process would trap the reactants and thus prevent the *thermodynamically preferred* hydrolysis reaction. The situation is different in solution, however. First, both processes involve significant solvent-induced barriers, and the calculations indicate a diminished preference for proton abstraction by hydroxide.[90] Second, and

(88) For recent reviews on hybrid quantum mechanical and molecular mechanical approaches see: Mordasini, T. Z.; Thiel, W. *Chimia* **1998**, *52*, 288−291. Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298−305. Gao, J. *Rev. Comput. Chem.* 1996, 7, 119−185.

(89) Strictly speaking, there is also an incompatibility between the quantum chemical and classical treatments of intramolecular (water) and intermolecular (TET−water, water−water) degrees of freedom. On the other hand, we should note that the classical description of intermolecular degrees of freedom was calibrated to reproduce quantum chemical results. The remaining "error" should be well within the expected accuracy of the overall approach.

(90) Monte Carlo simulations have been performed to perturb TET1 into the complex (OH$_2$···HNCHO$^-$). Using ESP/HF-derived point charges and 6M configurations per window on a reaction path generated in **Z** matrix coordinates, a free energy difference of −4.9 kcal/mol was obtained for aqueous solution. Solvation apparently reduces the gas-phase preference (MP2/6-31+G*, −20.6 kcal/mol) for proton abstraction over TET1 formation. This result is in qualitative agreement with ab initio reaction field calculations using the COSMO model[84] to include the effect of water as bulk solvent. COSMO MP2 energy differences between the two structures are −12.3 kcal/mol (MP2/6-31+G*) and −14.9 kcal/mol (MP2/DZV++-(2d,p)), respectively, as evaluated for gas-phase geometries, and −11.9 kcal/mol (−13.6 kcal/mol) for COSMO-optimized geometries (Baldridge, K. K., private communication, 1998). Even though the agreement between reaction field calculations and Monte Carlo simulations is not quantitative, it seems clear that solvation reduces the preference for proton abstraction over TET1 formation.

more importantly, the abundance of water makes this acid—base reaction an equilibrium process. Hence, there should always be sufficient amounts of formamide to enable the formation of a tetrahedral intermediate which is then hydrolyzed. The large barrier of the reverse reaction (−33 kcal/mol; see Table 8 and Figure 15) provides for the kinetic stability of the hydrolysis products and for the removal of the tetrahedral intermediate from the acid—base equilibrium.

In summary, the solvent-induced barrier toward formation of the tetrahedral intermediate is the most significant solvent effect seen in our calculations. The gas- and aqueous-phase energy profiles for the C—N bond cleavage reaction are rather similar, but the solution reaction shows a somewhat later transition state and, consequently, a somewhat smaller activation barrier for the reverse process. The preceding isomerization of TET1 to a reactive geometry (TET) is facilitated by the solvent: TET1 and TET are almost isoenergetic in aqueous solution, and the barrier between the two structures is reduced significantly. It is thus possible that TET1 is not even formed in solution, and that the reaction proceeds directly to the reactive geometry TET. Furthermore, there is no necessity for the participation of *only one* water molecule in the breakdown of TET. It is equally plausible to assume that one water molecule provides a proton and a different water molecule abstracts a proton.

With these limitations in mind, our model performs reasonably well in a direct comparison with experimental results. For base-promoted hydrolysis of *N,N*-dimethylformamide, Guthrie reports values of −4.7, 17.9, and 22.6 kcal/mol, respectively, for the free energy of reaction, the free energy of TET formation, and the free energy of activation.[11] For *N,N*-dimethylacetamide, he gives similar values of −6.9, 19.0, and 24.1 kcal/mol. Our calculated free energies for formamide hydrolysis[91] (R → P, −5.7; R → TET1, 18.7; and R → TS1, 27.3 kcal/mol) are thus in satisfactory agreement with the experimental data. Experiments on the ratio of ^{18}O exchange and hydrolysis of benzamide[6] and toluamide[8] further indicate that the tetrahedral intermediates of amides with primary amine groups tend to regenerate reactants rather than form products. This is compatible with our calculations which suggest a higher barrier for TET breakdown (TET → TS1, 6.7 kcal/mol) than for reversal of TET formation (TET → TET-T1, 0.8 kcal/mol; TET1 → TET-TS1, 1.3 kcal/mol), although the difference in free energies is probably overestimated.[92]

We may also compare our results to an earlier theoretical investigation. Weiner et al. have studied the same reaction using a much simpler approach based on energy minimizations rather than free energy simulations.[29] These authors have obtained a similar energy profile and have estimated barriers of 22 and 13 kcal/mol for TET formation and TET breakdown, respectively. Their results differ in a number of details, however, as, e.g., in a very *early* solution-phase transition state for the breakdown of TET. Still, there is reasonable overall agreement with the results obtained from our more thorough free energy simulation. This indicates that solvent effects for reactions of charged solutes are largely captured by solute—solvent and solvent—solvent interaction energies and that entropic effects are of lesser

importance. It should be kept in mind, however, that more accurate results as well as meaningful analyses of solvent effects still require statistical averaging over representative ensembles.

The origin of the various solvent effects seen in Figure 15 has been studied and discussed on the basis of partial charges, hydrogen bond energies, radial distribution functions, and energy pair distributions. From our analyses, variations in hydrogen bond strengths appeared to be the single most important factor in determining differences in free energy of solvation. An examination of the strongest hydrogen bonds between the solute and surrounding water molecules was thus sufficient to explain most of the effects of solvation. For the isomerization of TET, variations in hydrogen bond strength reflect changes of the solute geometry. For the other reactions (TET formation and breakdown), however, they reflect changes in the solute's charge distribution along the reaction path. The accurate representation of the solute's electrostatic potential was thus very important; hence, we decided to use ESP or RESP atomic charges calibrated individually for closely spaced points on the reaction coordinate. Initially, we also considered the application of a much simpler empirical charge scheme (QEq),[70] which, however, failed to reproduce subtle variations in the solute's charge distribution.

The present study was stimulated by our interest in enzymatic peptide cleavage. In a recent paper, we applied similar methodology to study the formation of a tetrahedral intermediate between the serine protease trypsin and a small peptide.[93] In this enzymatic reaction, the hydroxyl oxygen of Ser195 serves as the nucleophile which—deprotonated by His57—forms a bond with the scissile carbon of the substrate to initiate the subsequent C—N bond cleavage (see Figure 1). While our present model assumes a reaction under basic conditions, the enzyme still needs to deprotonate the hydroxy group to achieve the necessary nucleophilicity. This in turn requires proper geometric arrangement of the substrate, Ser195, and His57, which (in the absence of the rest of the protein) was estimated to cost 11 kcal/mol of cratic free energy. The total free energy for TET formation was calculated to be 31 kcal/mol for a model containing only the active site and the solvent. The protein, however, provides the proper preorganization from the beginning, and further stabilizes the tetrahedral intermediate by favorable electrostatic interactions (by ca. 5 kcal/mol). In summary, the enzyme-like formation of TET with a neutral hydroxyl group and a neighboring proton acceptor requires significantly more energy than the corresponding model reaction under basic conditions, but the geometric preorganization and electrostatic stabilization of the protein provide for substantial catalysis, which is likely to favor the enzymatic process over the corresponding base-catalyzed reaction.

From a methodological point of view, we may conclude that Jorgensen's conceptually simple approach has been successfully applied to study the rather complex hydrolysis of formamide in aqueous solution. On the basis of a quantum chemical determination of the gas-phase reaction path, this approach treats solvent effects applying statistical mechanics in combination with calibrated classical potentials. Normally, such a procedure requires that the solvent does not alter the pathway. As shown in this paper and in related work,[93] however, it can be applied more generally, provided that reasonable assumptions about the solution-phase path are implemented as constraints to the calculation. Ultimately, it would be desirable to include the solvent reaction field in the quantum chemical calculation of

(91) Taking into account the additional free energy necessary to constrain the reactants (formamide and hydroxide) to a productive reaction pathway,[52] we expect somewhat larger free energies of activation and of TET formation. For further details, see Table A1 in the Supporting Information.

(92) For toluamide at 100 °C, a k_{ex} (^{18}O exchange) to k_{hyd} (hydrolysis) ratio of 3.7 has been measured, which corresponds to a difference of 1.5 kcal/mol in free energies of activation at that temperature. See: Slebocka-Tilk, H.; Bennet, A. J.; Keillor, J. W.; Brown, R. S.; Guthrie, J. P.; Jodhan, A. *J. Am. Chem. Soc.* **1990**, *112*, 8507−8514.

(93) Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448−3457.

the reaction path, and to continue with a statistical mechanical treatment for a more thorough description of solvent effects. Such calculations have been reported by Lim and Jorgensen for a *polar* [2 + 2] cycloaddition reaction.[53] The reaction field treatment of *anionic* solutes is more difficult, however, due to problems arising from diffuse charge distributions. Recent developments in reaction field theory[56,57] are very promising in this regard and should allow for a more general applicability of the combined quantum chemical and statistical mechanical approach.

Supporting Information Available: Thermal corrections of gas-phase energies, statistical evaluation of hydrogen bond energies calculated with model potentials, atomic charges for structures discussed in this paper, potential of mean force convergence data, and radial distribution functions (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9837349