Ab Initio Evaluation of the Potential Surface for General Base-Catalyzed Methanolysis of Formamide: A Reference Solution Reaction for Studies of Serine Proteases

Marek Štrajbl,[†] Jan Florián, and Arieh Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062

Received July 12, 1999. Revised Manuscript Received March 1, 2000

Abstract: To elucidate the catalytic power of enzymes it is crucial to have clear information about the corresponding reference reactions in solution. This is needed since catalysis is defined by comparing enzymatic reactions to the relevant uncatalyzed reactions. Unfortunately, the energetics of the reference reactions of many important classes of enzymatic reactions have not been fully determined by experimental studies. In many cases it is hard to determine whether the given reaction involves a stepwise or a concerted mechanism. It is also hard to estimate the activation barrier for steps which are not rate determining. Fortunately, it is possible now to use computational approaches to augment the available experiments and to elucidate the shape of free energy surfaces of various reference reactions. Here we present a systematic study of the reference solution reaction for studies of serine proteases, i.e., the base-catalyzed and general base/acid catalyzed methanolysis of formamide. The present work is based on the use of combined ab initio/Langevin dipoles calculations and on a careful comparison to available experiments. The applied ab initio methodologies involve nonlocal density functional (B3LYP/AUG-cc-pVDZ) calculations and G2 theory. The construction of the relevant free energy surfaces involves partial geometry optimizations for the ammonia-catalyzed methanolysis of formamide. Subsequently, energy corrections based on the appropriate experimental pK_a values are applied to construct interpolated free energy surfaces for the water- and histidine-catalyzed reactions. Crucial points on the free energy surface for the water catalyzed reaction are also evaluated independently. We start by exploring the first step of the alcoholysis reaction, which involves a proton transfer from ROH to a base, a nucleophilic attack of RO^- on the amide, and a formation of a tetrahedral intermediate anion (TI). The interpolated free energy surface for the water-assisted alcoholysis involves a least energy path where the proton transfer is concerted with the nucleophilic attack. The corresponding activation barrier is ∼32 kcal/mol. The independently calculated surface for this reaction involves an activation barrier of ∼34 kcal/mol. These results are in a good agreement with the corresponding experimentally observed barrier (30−32 kcal/mol). The interpolated free energy surface for the histidine-catalyzed reaction involves a stepwise path, with a shallow surface that can also allow for a concerted path. This free energy surface is quite different than the fully concerted surface obtained in previous theoretical studies. The calculated activation barrier for the hisitidine-catalyzed reaction is around 26 kcal/mol. To examine the next step of the reaction we evaluated the basicities of the O and N atoms of the TI. These values were found to be 14 and 8 pK_a units, respectively. The calculated pK_a of the N atom indicates that the leaving group is protonated prior to the cleavage of the CN bond. The activation free energy for the CN bond cleavage is predicted to be 22 kcal/mol at 298 K. This barrier is independent of the nature of the general base. It is concluded that the nucleophilic attack is the rate-determining step for the acylation reaction studied. After this step, the reaction surface is rather flat, with only small barriers separating anionic and N-protonated TI from the product valley. The TI may become stabilized by the formation of the O-protonated form. The presented potential surface for the reaction with histidine as a base should provide a useful way for validating quantum mechanical studies of serine proteases. This surface should also allow the calibration of semiempirical approaches that can be used in studies of these enzymes.

1. Introduction

The enzyme-catalyzed formation and cleavage of peptide bonds represents a fundamental biochemical reaction that is utilized in many cellular metabolic processes. A majority of these reactions involve the nucleophilic attack of alcohol or thiol groups that are present in the active sites of the serine and cysteine proteases. The catalytic mechanisms of these enzymes have been the subject of intensive experimental studies for over 50 years (see refs 1−4). Since solution experiments can provide valuable information about reaction mechanisms and shed light on the corresponding enzymatic reactions, the analysis of the kinetics and thermodynamics of amide alcoholysis in solution has attracted considerable attention.[5−9] These studies have been

* Corresponding author. E-mail: warshel@invitro.usc.edu.
[†] Permanent address: Institute of Physics, Charles University, Prague, Czech Republic.

(1) Polgár, L. Mechanisms of Protease Action; CRC Press: Boca Raton, FL, 1989.
(2) Storer, A. C.; Ménard, R. *Methods in Enzymol.* **1994**, *244*, 486−501.
(3) Bender, M. L. *Chem. Rev.* **1960**, *60*, 53.
(4) Bruice, T. C.; Benkovic, S. J. *Bioorganic Mechanisms*; W. A. Benjamin, Inc.: New York, 1966; Vol. 1.
(5) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105−6109.
(6) Fersht, A. R. *J. Am. Chem. Soc.* **1971**, *93*, 3504−3515.
(7) Guthrie, J. P. *J. Am. Chem. Soc.* **1978**, *100*, 5892−5904.
(8) Guthrie, J. P. *J. Am. Chem. Soc.* **1974**, *96*, 3608−3615.

10.1021/ja992441s CCC: $19.00
Published on Web 05/18/2000

found eventually to be of an additional special value, since the origin of the catalytic power of enzymes can be understood completely only in reference to (and in contrast to) the uncatalyzed reaction in aqueous solution. Yet, several important mechanistic issues, such as the exact nature of the reaction coordinates or the rate-determining steps, remained unresolved.[6] This is because acyl-transfer reactions in solution represent a complex system with multitude of possible pathways that are difficult to separate by varying the experimental conditions. For example, it is difficult to assess individual rate constants for reactions in which imidazole serves as a general base, nucleophile, or both.[4]

In principle, the feasibility of individual reactive pathways can be examined by the methods of quantum chemistry. However, previous theoretical studies left major open questions. Empirical valence bond (EVB) studies used consensus experimental information[10,11] to generate solution potential surfaces. This procedure depends, of course, on the uniqueness of the analysis of the corresponding experimental information, which is still not fully established. The nature of the reference potential for the catalytic reaction of serine proteases has not yet been resolved by semiempirical or ab initio calculations since most of these calculations were restricted to the gas phase (rather than to solution) environment.[12−22] Unfortunately, gas phase ab initio calculations tend to overstabilize the concerted reaction pathways. For example, the gas-phase energy for a proton transfer from serine to imidazole[16] is much larger than the corresponding energy in aqueous solution. This is because the large hydration energy of the ion pair is neglected in gas-phase calculations. Thus, one of the mechanistic questions that remain open is whether the nucleophilic attack by serine involves a concerted or stepwise mechanisms in the nucleophilic attack/proton transfer step and in the C−N bond cleavage/proton transfer step. This problem cannot be resolved in a unique way despite the observation of kinetic isotope effects both in solution[23,24] and proteins.[25] That is, although these effects are usually interpreted as evidence for concerted pathways, they are also expected from shallow potential surfaces where the activation barriers for the stepwise and concerted pathways are similar. It is also clear that the isotope effects cannot tell us about the actual difference between the activation energies of the concerted and stepwise pathways.

It is obvious that a clearer understanding of the reference solution reaction requires one to move from gas phase to solution calculations. This fact was recognized in our early EVB studies[11,26,27] as well as by Kollman's[28] and Jorgensen's[29] groups, who studied the OH^- attack on formamide. The studies of Kollman and co-workers used a hybrid QM/MM approach that took into account the solute polarization by the solvent but evaluated average energies rather than the actual free energy profile. The study of Jorgensen and co-workers evaluated the free energy profile using gas-phase charges and structures. Such a procedure does not take into account the solute polarization by the solvent. This deficiency is not a major concern in studies of reactions that do not involve a large charge separation. This is probably the reason recent simulations of the attack of OH^- on formamide,[30] which used gas-phase charges, provided a free energy surface that is consistent with the measured rate of this reaction.

In contrast, the hydrolysis of amides by a water molecule and the general base catalysis of amide hydrolysis, which is directly relevant to the reaction of serine proteases, involve a major charge separation. Here, it is essential to use a solvation model that describes properly the solute−solvent coupling. Such models include implicit and dipolar solvation models, which, when properly calibrated, represent robust tools in the arsenal of the physical organic chemistry[31−34] and hybrid quantum mechanics/molecular mechanics (QM/MM) approaches (see refs 28, 35, and 36), which should, of course, involve a proper free energy evaluation.

Models that consider the solute polarization by the solvent have been used in studies of the hydrolysis of formamide in aqueous solution. These studies include the ellipsoidal cavity continuum model of Rivail and co-workers[37] that has been used to study the hydrolysis of formamide assisted by a single water molecule.[38] This study, as well as a later reinvestigation of this model system by Kallies and Mitzner,[39] who used the polarized continuum model,[40,41] provided a detailed discussion of the transition state structures involved. However, the calculated free energies for these transition states were about twice as large as the experimental activation free energies of the related reactions.[5]

The comparison of the alcoholysis of an amide in aqueous

(9) Gregory, M. J.; Bruice, T. C. *J. Am. Chem. Soc.* **1967**, *89*, 2121−2127.
(10) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
(11) Warshel, A.; Russell, S. T. *J. Am. Chem. Soc.* **1986**, *108*, 6569−6579.
(12) Scheiner, S.; Kleier, D. A.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2606−2610.
(13) Oie, T.; Loew, G. H.; Burt, S. K.; Binkley, J. S.; MacElroy, R. D. *J. Am. Chem. Soc.* **1982**, *104*, 6169−6174.
(14) Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2225−2229.
(15) Taira, K.; Gorenstein, D. G. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 3625.
(16) Daggett, V.; Schröder, S.; Kollman, P. *J. Am. Chem. Soc.* **1991**, *113*, 8926−8935.
(17) Jensen, J. H.; Baldridge, K. K.; Gordon, M. S. *J. Phys. Chem.* **1992**, *96*, 8340−8351.
(18) Krug, J. P.; Popelier, P. L. A.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 7604.
(19) Harrison, M. J.; Burton, N. A.; Hillier, I. H.; Gould, I. R. *Chem. Commun.* **1996**, 2769−2770.
(20) Dive, G.; Dehareng, D.; Peeters, D. *Int. J. Quantum Chem.* **1996**, *58*, 85−107.
(21) Antonczak, S.; Ruiz-Lopez, M.; Rivail, J.-L. *J. Mol. Model* **1997**, *3*, 434−442.
(22) Baeten, A.; Maes, D.; Geerlings, P. *J. Theor. Biol.* **1998**, *195*, 27−40.
(23) Jencks, W. P.; Carriuolo, J. *J. Biol. Chem.* **1959**, *234*, 1272.
(24) Jencks, W. P.; Carriuolo, J. *J. Biol. Chem.* **1959**, *234*, 1280.
(25) Schowen, R. L. Molecular Structure and Energetics. In *Principles of Enzyme Activity*; Liebman, J. F., Greenberg, A., Eds.; VCH Publisher: Weinheim, FRG: 1988; Vol. 9.
(26) Warshel, A.; Russell, S.; Weiss, R. M. In *Chemical Approaches to Understanding Enzyme Catalysis: Biomimetic Chemistry and Transition-State Analogues, Proceedings of the 26th OHOLO Conference. Studies in Organic Chemistry*; Green, B. S., Ashani, Y., Chipman, D., Eds.; Elsevier Scientific Publishing Company: Amsterdam, 1981; Vol. 10.
(27) Warshel, A.; Sussman, F.; Hwang, J.-K. *J. Mol. Biol.* **1988**, *201*, 139−159.
(28) Weiner, S. J.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1985**, *107*, 2219−2229.
(29) Madura, J. D.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1986**, *108*, 2517−2527.
(30) Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 5712−5726.
(31) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(32) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
(33) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719−734.
(34) Florián, J.; Åqvist, J.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 11524.
(35) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York: 1995; Vol. 7.
(36) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(37) Dillet, V.; Rinaldi, D.; Rivail, J. L. *J. Phys. Chem.* **1994**, *98*, 5034.
(38) Antonczak, S.; Ruiz-Lopez, M. F.; Rivail, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 3912−3921.
(39) Kallies, B.; Mitzner, R. *J. Mol. Model.* **1998**, *4*, 183−196.
(40) Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1981**, *55*, 117.
(41) Miertus, S.; Tomasi, J. *J. Chem. Phys.* **1982**, *65*, 239.

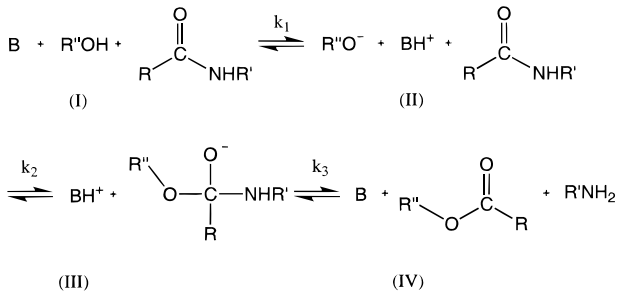
solution and in the active site of serine proteases has been the subject of several ab initio studies. The study of Stanton et al.[43] followed the approach developed by Jorgensen and co-workers[45] and used gas-phase ab initio calculations and a classical force field that was fitted to the gas-phase geometries. This classical force field or, alternatively, a set of fixed gas-phase geometries were subsequently used in free energy perturbation (FEP) calculations. As stated above this approach does not take into account the solute polarization by the solvent. This can lead to major problems in studies of processes that involve a large charge separation, such as proton transfer and S_N1 type processes[46] (see below). Furthermore, the classical force field used in the FEP studies is based on gas-phase calculations and thus may not reflect correctly the solvent effect on the solute geometry. These inherent problems could have been solved by fitting EVB surface and charges (rather than a classical force field) to the gas-phase ab initio results (e.g., refs 46−49), since the EVB approach provides a consistent way for transferring gas-phase potential surfaces and charges to solution studies. However, such an approach was not used in ref 43. Muller et al.[42] and Bentzien et al.[44] introduced a consistent ab initio/MM FEP approach for studies of enzymatic reactions. This approach involves the use of an EVB surface as a reference potential for the FEP ab initio calculations. This appears to give convergent results for the self-consistent ab initio solvation free energies (the substrate−protein interaction) along the reaction path that involved the solute fluctuations. However, a large statistical error was found for the contribution of the substrate intramolecular energy (note that no other ab initio approach has considered the fluctuations of the ab initio intramolecular solute energy). The challenging step of obtaining more stable intramolecular free energies was left for subsequent studies.

In view of the above discussion, it is not completely obvious what is the optimal strategy for obtaining a potential surface of the reference solution reaction for studies of serine proteases. It is also not clear what is the optimal way of comparing such a surface to the surface of the corresponding enzymatic reactions. It appears that ab initio QM/MM-FEP methods do not yet provide such an optimal approach.

An appealing strategy is provided by calibrating EVB surfaces and charges by fitting them to the corresponding gas-phase ab initio calculations[46−49] and using the calibrated EVB model in solution studies. The EVB approach provides a QM/MM model that captures correctly the solute−solvent coupling while still allowing very efficient FEP/umbrella sampling studies.[46] However, we feel that it is preferable to parametrize the EVB model using solution, rather than gas phase, ab initio results. This should allow one to use direct experimental information in validating the results of different basis sets and to provide a safeguard against cases where the EVB off-diagonal elements are not fully transferable from the gas-phase to solution studies.

The difficulties associated with the current use of all-atom solvent models may be partially overcome by using less rigorous simplified as well as implicit solvent models (e.g., refs 31−34). Such models provide converging results with high level ab initio wave functions. Here it is crucial to have a proper calibration and validation using available experimental reaction free energies and activation free energies. Of course, the use of simplified solvent models requires one to introduce some approximations about entropic contributions and other factors (see Theoretical Methods), but this treatment is justified in view of the above-mentioned difficulties with all-atom approaches. In fact, the open questions about entropic contributions can be eventually examined using EVB surfaces that are calibrated on the results of the simplified solvent model. Thus, we chose to use in this work a combination of high level ab initio calculations and the Langevin dipoles (LD) simplified solvent model.[50]

The ab initio/LD model is used in the present work in a systematic study of the potential surface of the reference solution reaction for studies of serine proteases. Since the formation of the acyl−enzyme is the rate-determining step in the serine protease catalyzed hydrolysis of a peptide bond,[51] we focus here on the reference solution reaction for the first (acylation) step of the reaction of serine proteases. This reaction can be described schematically as



where B denotes a general base (H_2O, NH_3, or histidine were considered in this study) and BH^+ denotes a general acid (NH_4^+, histidineH^+, or H_3O^+).

Any current method for calculations of reactions in enzymes and aqueous solution should be either calibrated or verified by considering relevant experiments in solution.[11,27,33] This is, of course, true for the EVB method, but it is also true for ab initio calculations, whose reliability in studies of solution chemistry is not fully guaranteed. Thus, we consider in section 3.1 the relevant experimental information about reactions of amides with alcohols in aqueous solution. The concertedness of the nucleophilic attack step (eq 1, I → III) and the basis set effects on the calculated energies are discussed in section 3.2. Sections 3.3 and 3.4 are devoted to the investigation of the stability of the tetrahedral intermediate III and to the breakdown of this intermediate (III → IV). Finally, it is argued that the resulting surface should provide a reliable starting point for studying catalytic effects of the corresponding enzymes.

2. Theoretical Methods

To evaluate free energy surfaces for the reference solution reactions, we used a self-consistent hybrid approach that combines ab initio quantum chemical calculations with the Langevin dipoles (LD)[36] and polarized continuum (PCM)[40,41] solvation models. This is done using our recent version of the LD model that has been carefully calibrated considering experimental hydration free energies of both neutral and ionic solutes.[50] Furthermore, this solvation model has been used

(42) Muller, R. P.; Florián, J.; Warshel, A. Semiempirical and Ab Initio Modeling of Chemical Processes: From Aqueous Solution to Enzymes. In *Biomolecular Structure and Dynamics*; NATO ASI Series; Vergoten, G., Theophanides, T., Eds.; Kluwer Academic Press: Norwell, MA, 1997.
(43) Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. *J. Am. Chem. Soc.* **1998**, *120*, 3448−3457.
(44) Bentzien, J.; Muller, R. P.; Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 2293−2301.
(45) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 154.
(46) Hwang, J.-K.; King, G.; Creighton, S.; Warshel, A. *J. Am. Chem. Soc.* **1988**, *110*, 5297.
(47) Lobaugh, J.; Voth, G. A. *J. Chem. Phys.* **1996**, *104*, 2056.
(48) Vuilleumier, R.; Borgis, D. *Chem. Phys. Lett.* **1998**, *284*, 71.
(49) Vuilleumier, R.; Borgis, D. *J. Phys. Chem. B* **1998**, *102*, 4261.
(50) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 5583−5595.
(51) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331−358.

extensively in studies of chemical processes in solution.[33,34,52,53] The actual implementation of the method is described below.

2.1. Gas-Phase Calculations. The gas-phase geometries were calculated by a gradient optimization at the Hartree−Fock (HF)/6-31G(d) level. The nature of the transition states was evaluated by subsequent calculations of the HF/6-31G(d) harmonic vibrational frequencies. The HF/6-31G(d) level was chosen on the grounds of computational efficiency and also for consistency with the LD solvation model (see below). The gas-phase energies were determined for the HF/6-31G(d) geometries by single point B3LYP density functional calculations, using the 6-31G(d) and AUG-cc-pVDZ basis sets. G2 theory[54,55] was used for the calculation of the gas-phase basicity of the key tetrahedral intermediate. Ab initio B3LYP/6-31G(d)//HF/6-31G(d) and B3LYP/AUG-cc-pVDZ//HF/6-31G(d) methods are further abbreviated as B3LYP and B3LYP$^+$, respectively. All ab initio calculations were carried out with the Gaussian 94 program.[56]

The gas-phase reaction coordinates were evaluated using the intrinsic reaction coordinate (IRC) calculation in mass-weighted internal coordinates at the B3LYP$^+$ level. The relevant searches on the corresponding solution surfaces are considered below.

2.2. Solvation Free Energies. Solvation free energies (i.e., the free energies of transfer of the solute molecule from a 1 M concentration in the gas phase to a 1 M concentration in aqueous solution at 298 K) were calculated using the LD solvation model.[50] The LD model evaluates an average polarization of the solvent molecules surrounding the solute by using a discrete dipolar representation of the solvent. This model uses potential-derived PCM HF/6-31G(d) atomic charges to represent the charge distribution of the solute molecules. That is, our model takes into account the polarization of the solute by the solvent and the corresponding energy contributions. This self-consistent treatment uses the solute charges evaluated by the PCM model[40,41] rather than by the LD model in order to have a convenient external coupling with the Gaussian program which implements the PCM method. Note, in this respect, that earlier LD models[57] involved a direct coupling of the solvent polarization to the solute Hamiltonian. Besides the electrostatic and induction contributions mentioned above, the calculated LD solvation free energies involve parametrized terms that represent dispersion and hydrophobic solute−solvent interactions. Default Pauling's atomic radii multiplied by a standard factor of 1.2 and a dielectric constant $\epsilon = 80$ were used for the PCM calculations. The LD and PCM calculations were carried out using the programs ChemSol 1.13[50] and Gaussian 94.[56]

2.3. The Solvent Cage Concept. To compare the calculations of a specific reaction in an enzyme and in solution it is very convenient[10,58] to divide the activation free energy of the reaction in solution into two parts: the free energy, ΔG^{cage}_{bind}, which involves assembling of the reacting fragments into a single solvent cage, and the activation barrier, $\Delta g^{\ddagger}_{cage}$, for the reaction within the cage. This approach has the following advantages: First ΔG^{cage}_{bind} can be compared to the binding free energy of the enzyme, ΔG^{enz}_{bind}, which corresponds to the enzyme−substrate dissociation constant. Second, $\Delta g^{\ddagger}_{cage}$ can be compared to the activation energy $\Delta g^{\ddagger}_{cat}$, which corresponds to the rate constant k_{cat} for the given enzymatic reaction. This partitioning allows one to examine the origin of the reduction of the difference ($\Delta g^{\ddagger}_{cat} - \Delta g^{\ddagger}_{cage}$), while conveniently separating it from the much better understood factors that are involved in the binding process. One may wonder at this point about the entropic factors associated with bringing the reacting fragments to the same cage,[30] but these are obtained from the well-known 55 M concentration correction or closely related estimates (see Chapter 9 in ref 10). That is, the cage can be defined by giving each reactant an effective volume V_{eff} that corresponds to a 55 M concentration (i.e., $\Delta S^{cage}_{bind} = R \ln(V_{eff}/V_0) \simeq R \ln(1/55)$, where $V_0 \simeq 1660$ Å^3 is the molar volume accessible to one molecule in the 1 M state).

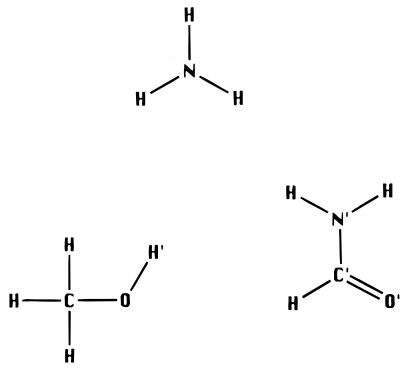


Figure 1. The system used for the evaluation of the potential surface for the NH_3-assisted methanolysis of formamide and for interpolation the surfaces for the H_2O and histidine reactions.

The transformation of the relevant experimental data about the activation barrier to the data corresponding to the solvent cage is straightforward for a stepwise mechanism. For example, the "experimental" free energy for the stepwise mechanism is obtained by combining the experimental information about each separate step in the stepwise mechanism for a 55 M concentration of the reacting fragments (e.g., combining the information about the proton transfer and nucleophilic attack steps of eq 1). Comparing the calculated energetics of the stepwise mechanism to the above "experimental" estimate is used to validate the calculations even for cases when the stepwise mechanism is not the one with the lowest activation barrier.

In calculating Δg_{cage} we used a reference state that corresponds to the reactants in the configuration characterized by the O···C′ distance of 3.5 Å (Figure 1). Analogously, the calculated free energies of the studied products correspond to the weakly bound complex characterized by the 3.5 Å distances between the C′···N′ atoms (coincidentally, the same 3.5 Å distance was chosen for the reference C′···O distance by Stanton et al.[43] based on their PMF calculations). The purpose of defining the reference state of the solution reaction in terms of the weakly bound complex (instead of using infinitely separated fragments) was to eliminate inaccuracies associated with the gas-phase and solvation energies of small ions, such as OH^-, CH_3O^-, or NH_4^+. Also, this approach is more directly related to our use of a reference solvent cage in calculations of enzyme catalysis.[63] Finally, avoiding the definition of the reference state in terms of the isolated (infinitely separated) reactants enables us to assume that the gas-phase entropy contributions do not modify in a major way the shape of the reaction surfaces obtained using the approximation of eq 2 (see also below).

(52) Florián, J.; Sponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884−892.
(53) Florián, J.; Strajbl, M.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 7959−7966.
(54) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
(55) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys*. **1993**, *98*, 1293.
(56) Frisch, M.; et al. Gaussian 94, Revision D.2, Gaussian, Inc. 1995.
(57) Luzhkov, V.; Warshel, A. *J. Comput. Chem*. **1992**, *13*, 199.
(58) Warshel, A. *Biochemistry* **1981**, *20*, 3167.

2.4. Free Energy Surfaces and Entropic Considerations. The dependence of the reaction free energy on the reaction coordinate should be considered (see ref 33) as a potential of mean force and was thus expressed as

$$\Delta g_{\text{cage}}(r) = \Delta E^{\text{gas}}(r) + \Delta\Delta G_{\text{solv}}(r) \qquad (2)$$

where r is the solute coordinate relative to a chosen reference coordinate (see below), ΔE^{gas} is the B3LYP$^+$ gas-phase energy, and $\Delta\Delta G_{\text{solv}}(r)$ is the change in solution free energy relative to its value in the reference point. Note that $\Delta\Delta G_{\text{solv}}$ includes the effect of the polarization of the solute by the field of the solvent (see section 2.2). A more rigorous version of eq 2 is given by $\Delta g_{\text{cage}}^{\text{total}} = \Delta g_{\text{cage}} - T\Delta S' + \Delta H_{\text{vib}}$, where $\Delta S'$ is the solute entropic contribution (the solvent contribution is included in $\Delta\Delta G_{\text{solv}}$) and ΔH_{vib} is the vibrational enthalpy contribution. Thus, the use of the approximation of eq 2 involves the assumption that the solute entropy and vibrational enthalpy do not modify the shape of the calculated free energy surfaces in a major way. The ground state of the system is expected to have a larger translational and rotational mobility than its transition state. Consequently, neglecting $\Delta S'$ should result in an underestimating of $\Delta g_{\text{cage}}^{\ddagger}$. However, the entropic contribution to $\Delta g_{\text{cage}}^{\ddagger}$ is much smaller than what is usually implied in related studies of enzymatic reactions (e.g., refs 43, 59, and 60). This is because in these studies, the entropic contribution to catalysis is implicitly assumed to be only due to the ground-state motions in the solvent cage.[61] However, most of the degrees of freedom of the reacting fragments have similar mobility in the ground and transition state (e.g., the base in eq 1) and their entropic contributions cancel.[61] Furthermore, the entropic contribution to $\Delta g_{\text{cage}}^{\ddagger}$ cannot be evaluated without specialized simulation approaches, and such approaches have not been fully developed (although a significant progress in this direction is being made[62]). One might still try to obtain a rough estimate of $(\Delta S_{\text{cage}}^{\ddagger})'$ by using[33] $(\Delta S_{\text{cage}}^{\ddagger})' = \beta\Delta S_{\text{gas}}^{\ddagger}$, where the scaling constant β is obtained by forcing $(\Delta S_{\text{cage}}^{\ddagger})'$ to reproduce the observed $\Delta S_{\text{cage}}^{\ddagger}$ (see Supporting Information). However, in our case the addition of $(\Delta S_{\text{cage}}^{\ddagger})'$ (and also the relatively small $\Delta H_{\text{vib}}^{\ddagger}$) to $\Delta g_{\text{cage}}^{\ddagger}$ results in an overestimation of the observed activation barriers. In fact, it seems that the neglect of $(\Delta S_{\text{cage}}^{\ddagger})'$ coincidentally cancels the overestimate of the ab initio activation energies, since the calculated $\Delta g_{\text{cage}}^{\ddagger}$ agrees well with the corresponding observed value. Thus, we confined the present study to the simple approximation of eq 2. Finally, it should be noted that the purpose of the present paper is to provide a reasonable estimate for different activation barriers and thus to serve as a basis for calibrating EVB or related potential surfaces. The EVB approach can then be used in a proper estimate of the entropic contributions to $\Delta g_{\text{cage}}^{\ddagger}$ and $\Delta g_{\text{cat}}^{\ddagger}$, and such studies are now in progress in our laboratory.

2.5. Calibration of the Calculated Free Energy Surface. To validate and calibrate our calculated free energy surfaces we used the corresponding differences in pK_a values. That is, we required that the calculated free energy, $\Delta g_{\text{PT}}^{\text{AH/B}}$ (kcal/mol), for a proton transfer (PT) between a donor (AH) and acceptor (B), be equal to the corresponding experimental estimate. This was done using the expression[58]

$$\Delta g_{\text{PT}}^{\text{AH/B}}(r) \simeq \Delta G_{\text{PT}}^{\text{AH/B}} + \Delta G_{\text{QQ}}(r) \simeq 2.3RT(\text{p}K_a(\text{AH}) - \text{p}K_a(\text{BH})) + \Delta G_{\text{QQ}}(r) \qquad (3)$$

where r is the distance between the donor and acceptor, R denotes the universal gas constant, T is the absolute temperature, and ΔG_{QQ} is the change in electrostatic interactions between the donor and acceptor during the proton-transfer processes. The ΔG_{QQ} term can be estimated using Coulomb's law with a dielectric constant $\epsilon_{\text{eff}} \geq 30$ ($\Delta G_{\text{QQ}}(r) = 332\ (Q_aQ_b/\epsilon_{\text{eff}}r)$). This approximation was found to describe quite reliably interactions between charged molecules in solution even at short distances ($r \geq 3$ Å).[64] Here, however, we decided to use the value of ΔG_{QQ} at 4 Å for $r \leq 4$ Å. This was done since we assumed that the interaction at shorter distances might involve charge-transfer effects that would reduce the corresponding $|\Delta G_{\text{QQ}}|$.

Equation 3 allowed us to construct reaction free energy surfaces for general bases that have not been studied by explicit ab initio calculations. More specifically, our actual ab initio calculations were carried out only for a model system that involved NH_3 as a base. However, the pK_a corrections of eq 3 were used in constructing the potential surfaces for H_2O and histidine as general bases. (The presented results will of course apply for similar bases of the same pK_a, e.g., imidazole.) In this way we were able to keep the size of the systems studied within the limit manageable by high level ab initio calculations, while effectively evaluating the free energy surface for the reference solution reaction for studies of serine proteases (in this case histidine is the general base). Thus, in constructing free energy surfaces for H_2O and histidine as bases, we determined the values of the free energies at the corners of the surfaces, where the bases are fully protonated, by adjusting the corresponding values calculated for ammonia as a base using the relationships

$$\Delta g_{\text{PT}}(r)^{\text{CH}_3\text{OH/His}} = \Delta g_{\text{PT}}(r)^{\text{CH}_3\text{OH/NH}_3} + \Delta G_{\text{PT}}{}^{\text{NH}_4^+/\text{His}} \qquad (4)$$

$$\Delta g_{\text{PT}}(r)^{\text{CH}_3\text{OH/H}_2\text{O}} = \Delta g_{\text{PT}}(r)^{\text{CH}_3\text{OH/NH}_3} + \Delta G_{\text{PT}}{}^{\text{NH}_4^+/\text{H}_2\text{O}} \qquad (5)$$

where the ΔG_{PT} values were determined using eq 3.

(59) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678.

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

(61) The real open question about entropic effects is associated with the restriction of the degrees of freedom of the reacting fragments upon moving from the solvent cage to the enzyme active site or, more specifically, the difference between the entropic contributions to the ground and transition state in the enzyme and solvent cage. The evaluation of the entropic contributions to catalysis, i.e., the difference between these contributions to $\Delta g_{\text{cage}}^{\ddagger}$ and $\Delta g_{\text{cat}}^{\ddagger}$, is very challenging. To the best of our knowledge, the only reported study that involved the actual potential surfaces in both the enzyme and solution is given in Chapter 9 of ref 10. This study indicated that the entropic contribution to catalysis is small, although more rigorous studies are clearly needed. Thus, it is quite likely that the entropic contributions to catalysis were overestimated in early works.[59,60] In this respect it is also useful to point out that a recent theoretical estimate of entropic contributions to the catalysis of serine proteases[43] overestimated this effect. That is, this study was limited to the estimate of the so-called cratic free energy without attempting to evaluate the relevant $\Delta g_{\text{cage}}^{\ddagger}$ and $\Delta g_{\text{cat}}^{\ddagger}$. For example, no estimate was made for the entropy at the transition state and no relevant simulation study was performed in the enzyme active site. The entropy at the transition state would reduce the total contribution to $\Delta g_{\text{cage}}^{\ddagger}$, since for example the histidine molecule has a similar mobility in the ground and transition state and its entropic contribution cancels.[62] Also note that the use of the full $\Delta S_{\text{gas}}^{\ddagger}$ (i.e., $\beta = 1$) in estimating $\Delta S_{\text{cage}}^{\ddagger}$ (as done, for example, by Kuhn and Kollman [Kuhn, B.; Kollman, P. A. *J. Am. Chem. Soc.* **2000**, *122*, 2586]) does not reproduce the observed activation barriers.

(62) Strajbl, M.; Sham, Y. Y.; Villà, J.; Chu, Z.-T.; Warshel, A. *J. Phys. Chem. B* **2000**, *104*, 4578.

(63) Warshel, A.; Florián, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950−5955.

(64) Warshel, A.; Russell, S. T. *Q. Rev. Biol.* **1984**, *17*, 283.

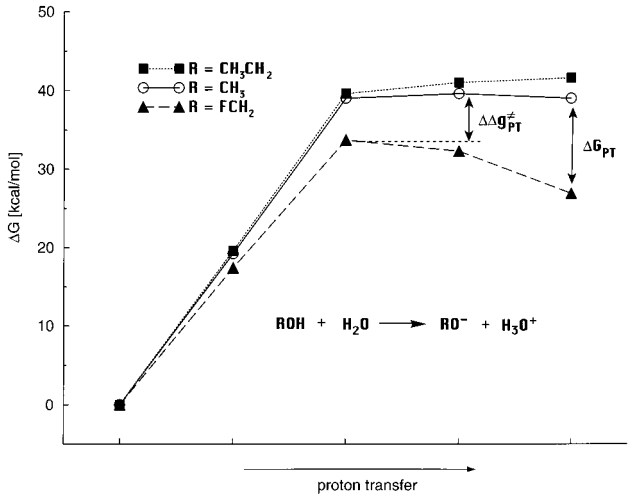


Figure 2. The free energy profiles (B3LYP$^+$+LD) for proton transfers from ROH to H_2O used in estimating of the α coefficient in the LFER relationship ($\alpha = \Delta\Delta g^{\ddagger}_{PT}/\Delta G_{PT}$, see eq 6). The donor and acceptor oxygen atoms were constrained at a 3 Å separation.

For the transition state regions of the proton-transfer reaction surfaces we used the Marcus relationship[65] that provides a linear correlation between the change in the reaction free energy and the change in the corresponding activation barrier. This relationship can be written as

$$(\Delta g^{\ddagger}_{PT})^{AH/B_2} = (\Delta g^{\ddagger}_{PT})^{AH/B_1} + \alpha \Delta G_{PT}^{B_1H/B_2} \qquad (6)$$

where B_1 and B_2 are two different proton acceptors and α is a linear free energy relationship parameter, which is given by

$$\alpha = (\Delta G_{PT}^{CH_3OH/NH_3} + \lambda)/2\lambda \qquad (7)$$

where λ is the reorganization energy for a proton transfer from CH_3OH to NH_3. The value of λ was estimated as 50 kcal/mol from the EVB calculations of ref 66. This gave $\alpha \simeq 0.62$ for $\Delta G_{PT} = 12$ kcal/mol. Alternatively, we evaluated α by ab initio calculations of the free energy surface for the reaction $CH_2XOH + H_2O \rightarrow CH_2XO^- + H_3O^+$, where X is H, CH_3, and F, respectively (Figure 2). These calculations yielded α in the $0.4 < \alpha < 0.7$ range. In view of this estimate, we chose $\alpha = 0.6$. In addition, we examined this assumption by using actual ab initio calculations to evaluate key points on the surface of the water-assisted reaction (see section 3.2). The application of the correction given by eq 6 to the whole reaction surface corresponds to the assumption of a linear free energy relationship between ΔG_{PT} and $\Delta g^{\ddagger}$. The validity of such relationships in solution reactions has been confirmed by numerous experimental[67−69] and theoretical works.[70,71] Furthermore, the validity of this relationship for the specific case of the water-assisted methanolysis of formamide is justified in this paper (see section 3.2).

The experimental pK_a values[72] used for NH_4^+, protonated histidine, H_3O^+, and methanol were 9.3, 6.0, −1.0, and 15.5, respectively. The pK_a of the tetrahedral intermediate was calculated using the relationship:[50,64]

$$pK_a(AH) = \frac{B_g(A) + \Delta G_{solv}(H^+) + \Delta G_{solv}(A) - \Delta G_{solv}(AH)}{2.303RT} \qquad (8)$$

where $B_g(A)$, $\Delta G_{solv}(AH)$, and $\Delta G_{solv}(A)$ are the gas-phase basicity and the solvation free energy of a protonated solute (AH) and its conjugate base (A), respectively. The solvation free energy of the proton, $\Delta G_{solv}(H^+)$, was taken as −259.5 kcal/mol.[50] The gas-phase basicities were calculated at the G2 level as

$$B_g(A) = G_{gas}(A) + G_{gas}(H^+) - G_{gas}(AH) \qquad (9)$$

where G_{gas} denotes the free energy of a gas-phase solute at 1 M concentration.

2.6 Geometry Search on the Solution Free Energy Surface. The present work focuses on the energetics of the amide hydrolysis in solution rather than on the location of the exact transition state (TS), which is less crucial for estimating activation barriers because the transition state region is frequently flat. To explore the solution free energy surface near the TS region, we performed one initial search along the gas-phase intrinsic reaction coordinate (IRC) and determined free energies along this coordinate. This search was followed, in most cases, by a systematic mapping of the solution free energy surface. This mapping was carried out using a sequence of partial geometry optimizations. The coordinates used in the different mapping studies are summarized below.

In studying the general base catalyzed surface we relaxed our model system (Figure 1) by the gas-phase HF/6-31G* optimization, while keeping the N−O and N−N′ distances frozen at 3.0 Å. This initial geometry optimization was carried out in order to find a suitable starting configuration for subsequent calculations. The resulting geometry was used to define the mutual orientation of the reacting fragments. This orientation was determined by setting the bond angles C−O−C′ and O−C′−N′ to 119.7° and 98.1°, respectively, and setting the improper torsions C−O−C′−N′ and O−C−C′−N′ to 232.4° and 331.0°, respectively. These degrees of freedom were kept frozen during the subsequent calculations along the reaction coordinate described below. The interacting molecules depicted in Figure 1 were driven from the reactant to the product state of the general base/acid catalyzed methanolysis of formamide (see eq 1) by changing the position of the H′ atom and by changing the O−C′ and C′−N′ distances. The proton-transfer coordinate of the general base-catalyzed step was mapped for five different positions of the proton: H′ bonded to O, $r(OH') = 1 + (r(ON) - 2)/4$, $r(OH') = 1 + (r(ON) - 2)/2$, $r(NH') = 1 + (r(ON) - 2)/4$, and H′ bonded to N. Similarly, the following proton positions were used for the general acid catalyzed step: H′ bonded to N, $r(NH') = 1 + (r(NN') - 2)/4$, $r(NH') = 1 + (r(NN') - 2)/2$, $r(N'H') = 1 + (r(NN') - 2)/4$, and H′ bonded to N′. The nucleophilic attack coordinate, $r(OC')$, and the coordinate for the breakdown of the tetrahedral intermediate, $r(C'N')$, were mapped, as shown in Tables 1 and 3, at 10 points. These calculations were performed for fixed N−O and N−N′ separations ranging from 2.3 to 4.0 Å in order to find the lowest

(65) Marcus, R. A. *J. Chem. Phys.* **1968**, *72*, 891.
(66) Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218.
(67) Hammet, L. P. *Physical Organic Chemistry*; McGraw-Hill: New York, 1990.
(68) Kreevoy, M. M.; Truhlar, D. G. In *Investigation of Rates and Mechanisms of Reactions*; Bernasconi, C. F., Ed.; John Wiley Sons: New York: 1986; Vol. 6.
(69) Albery, W. J. *Annu. Rev. Phys. Chem.* **1980**, *31*, 227.
(70) Kong, Y. S.; Warshel, A. *J. Am. Chem. Soc.* **1995**, *117*, 6234−6242.
(71) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Chem.* **1991**, *20*, 267−298.
(72) Serjeant, E. P.; Dempsey, B. *Ionisation Constants of Organic Acids In Aqueous Solution*; Pergamon Press Inc.: New York, 1979.

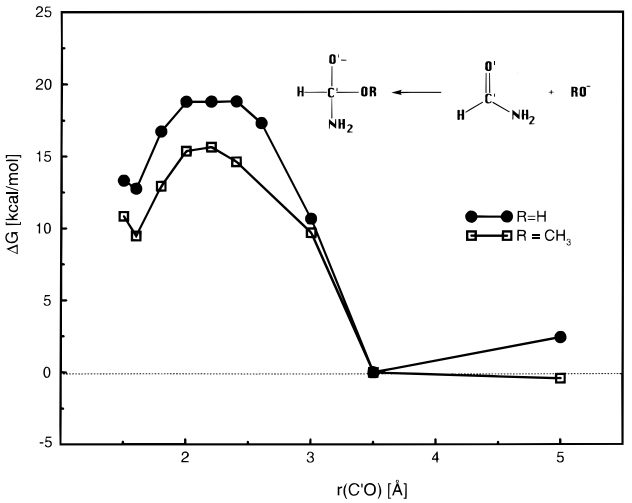


Figure 3. The calculated reaction profiles for the unassisted attack of OH^- and CH_3O^- on formamide.

energy pathway. The values of all other degrees of freedom were determined by gas-phase HF/6-31G* optimizations. In addition, the reference reactant (I in eq 1) and product structures (IV in eq 1) were determined by gas-phase HF/6-31G* optimizations while only $r(OC') = 3.5$ Å and $r(C'N') = 3.5$ Å, respectively, were constrained.

In studying the uncatalyzed bimolecular reaction of CH_3OH and formamide we determined three gas-phase HF/6-31G* structures on the IRC pathway. The reference reactant state configuration was calculated with the O−C′ distance fixed to 3.5 Å, all remaining degrees of freedom were optimized. The transition state structures presented in this work were fully optimized. Reaction profiles for the unassisted attack of RO^- on formamide were calculated by mapping the OC′ coordinate while other degrees of freedom were optimized.

3. Results and Discussion

3.1. Analysis of Relevant Experiments. The acylation reaction (eq 1) proceeds extremely slowly in aqueous solution at ambient temperature. Fortunately, it is also strongly endothermic,[73,74] so that the kinetic parameters can be determined by combining the experimental ΔG_0 with the experimental rates for the reverse reaction.[6,75] The analysis of Fersht[6] indicated that the rate constant k_2 (M^{-1} s^{-1}) for the nucleophilic attack of $R''O^-$ on amide (when the base B is a water molecule) increases with the decreasing basicity of the amine moiety according to the relationship

$$\log k_2 = 7.2 - 0.67\ pK_a(R'NH_2) \qquad (10)$$

Note that k_2 in eq 10 corresponds to k_{-4} in the notation of ref 6. It was also established that the overall rate constant for the reaction I → III when R″OH is methanol and the substrate is formylhydrazine is, in water at neutral pH, 4.3×10^{-10} M^{-1} s^{-1} and 4.3×10^{-8} M^{-2} s^{-1}. The two values for this rate constant correspond to the reaction that proceeds in the absence and presence of a base of $pK_a(BH) = 7$, respectively. Because for formylhydrazine the reaction III → IV is fast,[76] the overall rate constant for the methanolysis of formylhydrazine (representing reaction I $\xrightarrow{k}$ IV) can be estimated as $k = 4 \times 10^{-10}$ M^{-1} s^{-1}.

The experimental kinetic information mentioned above can be transformed into the activation free energy $\Delta g^{\ddagger}_{cage}$ of the pertinent reaction step occurring in the solvent cage (see section 2.3) using absolute rate theory.[77] This theory is a very good approximation for reactions with significant barriers in condensed phase.[10] Thus, for the bimolecular reactions occurring at ambient temperature, we can write

$$55k \simeq F \times 6 \times 10^{12} \exp(-\Delta g^{\ddagger}_{cage}/RT) \qquad (11)$$

where $\Delta g^{\ddagger}_{cage}$ is the activation barrier (in kcal/mol), F is the transmission factor, R is the universal gas constant, T is the thermodynamic temperature ($RT \simeq 0.6$ kcal/mol for $T = 298$ K), and the rate constant k is given in s^{-1} mol^{-1}. A reviewer has questioned the validity of transition state theory (TST) and the prominent role played by the activation free energy of eq 11. This impression may be due to the impact of Kramers' model[78] and many subsequent studies. These works provided a rigorous description of diffusive processes and the behavior of the rate constant in cases of low activation barriers. However, for chemical reactions in solution where the activation barriers are higher than the diffusive limit (i.e. higher than ∼10 kcal/mol), TST is an excellent approximation. More specifically, it is now well accepted (e.g. refs 84−86) that all true dynamical effects are contained in the transmission factor F of eq 11. This factor reflects the number of recrossing of the transition state region by productive trajectories and the chance that a trajectory that reaches the transition state region will continue to the product state.[46,84−86] Computer simulations of chemical reactions in solutions (see, e.g., Figure 15 in ref 46 and ref 87) have established repeatedly that $0.1 < F < 1$. Furthermore, simulations of reactions in enzymes[10,27,88] have established that F has similar magnitudes in enzymes and solutions. This implies that dynamical factors do not play a major role in reactions of the type studied here. This is not to say, of course, that atoms are not moving but to clarify that the probability for a productive trajectory is given by the corresponding $\exp(-\Delta g^{\ddagger}/RT)$.

Using eqs 10 and 11 and assuming $pK_a(RNH_2) = 10$, the barrier for the attack of CH_3O^- on an amide becomes $\Delta g^{\ddagger}_{cage} \sim 15$ kcal/mol. Similarly, for the attack of methanol on amides, i.e., for the I → IV reaction, we obtain $\Delta g^{\ddagger}_{cage} \sim 28$ kcal/mol. Interestingly, for the related reaction of OH^- attack on dimethylformamide, a considerably higher experimentally based estimate of $\Delta g^{\ddagger}_{cage} \sim 22$ kcal/mol has been reported.[8] However, the rate constant for the base-catalyzed hydrolysis of formamide at 80 °C (0.21 M^{-1} s^{-1})[89] implies a somewhat lower value of $\Delta g^{\ddagger}_{cage}$ (18 kcal/mol) for the OH^- attack at 298 K. Therefore, we adopt here an average value of 20 kcal/mol as the consensus experimental estimate of $\Delta g^{\ddagger}_{cage}$ for a OH^- attack on an amide.

Additional important information is obtained by analyzing the reaction when the nucleophile is H_2O. The uncatalyzed attack

(73) Fersht, A. R.; Requena, Y. *J. Am. Chem. Soc.* **1971**, *93*, 3499.
(74) Jencks, W. P.; Gilchrist, M. *J. Am. Chem. Soc.* **1964**, *86*, 4651.
(75) Fersht, A. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1970**, *92*, 5442.
(76) Blackburn, G. M.; Jencks, W. P. *J. Am. Chem. Soc.* **1968**, *90*, 2638.
(77) Eyring, H.; Polanyi, M. Z. *Phys. Chem. B* **1931**, *12*, 279.
(78) Kramers, H. A. *Physica* **1940**, *7*, 284.
(79) Frauenfelder, H.; Wolynes, P. G. *Science* **1985**, *229*, 337.
(80) Truhlar, D. G.; Hase, W. L.; Hynes, J. T. *J. Phys. Chem.* **1983**, *87*, 2664.
(81) Grote, R. F.; Hynes, J. T. *J. Chem. Phys*. **1980**, *73*, 2715.
(82) Northrup, S. H.; Hynes, J. T. *J. Chem. Phys*. **1980**, *73*, 2700.
(83) Hynes, J. T.; Kapral, R.; Torrie, G. M. *J. Chem. Phys.* **1980**, *72*, 177.
(84) Anderson, J. B. *J. Chem. Phys*. **1973**, *58*, 4684.
(85) Bennet, C. H. In *Algorithms for Chemical Computations*; Christofferson, R. E., Ed.; American Chemical Society: Washington, DC, 1977.
(86) Keck, J. C. *Adv. Chem. Phys.* **1966**, *13*, 85.
(87) Bergsma, J. P.; Gertner, B. J.; Wilson, K. R.; Hynes, J. T. *J. Chem. Phys.* **1987**, *86*, 1356.
(88) Neria, E.; Karplus, M. *Chem. Phys. Lett.* **1997**, *267*, 23.

by water on the peptide bond has been studied by Radzicka and Wolfenden[5] in neutral aqueous solution at elevated temperatures. The measured rate constants were then extrapolated to the room-temperature by assuming that the activation entropy is constant over the considered temperature range. For the hydrolysis of acetylglycylglycine *N*-methylamide by neutral water, Radzicka and Wolfenden obtained the rate constant 3.6×10^{-11} s^{-1} at 298 K. This rate constant corresponds to an activation free energy $\Delta g^{\ddagger}_{cage} = 32$ kcal/mol. The rate constant of 8×10^{-8} s^{-1} for the hydrolysis of formamide by neutral water at 80 °C determined by Hine and co-workers[89] is also consistent with $\Delta g^{\ddagger}_{cage} \simeq 30$ kcal/mol at 298 K.

3.2. The Formation of the Tetrahedral Intermediate. The RO$^-$ Nucleophilic Attack. Our task is to obtain reliable free energy surfaces for base-catalyzed reactions of amides with alcohols in aqueous solution. As a first step in this analysis, we examine the simplest and cleanest related reaction: the nucleophilic attack of RO$^-$ on amides. The reaction profiles for the unassisted nucleophilic attack of CH_3O^- and OH^-, calculated at the B3LYP$^+$+LD level are compared in Figure 3 (see section 2.1 for notation of methods). The calculated activation barriers of 16 and 19 kcal/mol, respectively, are in reasonable agreement with the corresponding experimental estimates of 15 and 20 kcal/mol (see section 3.1). The transition state (TS) for the OH^- attack was found at a C···O distance of around 2.2 Å. This distance indicates that the TS for the OH^- attack on formamide in aqueous solution should occur earlier than implied previously by Bakowies and Kollman,[30] who calculated a C···O distance of 1.85 Å at the TS of this reaction.

The General Base Catalysis. Although the rate constants for acyl transfer reactions proceeding via an alkoxide attack are larger by several orders of magnitude than those for the attack by a neutral alcohol (see section 3.1), the latter mechanism is more relevant for neutral aqueous solutions and as a reference system for the serine protease reaction. This is especially true when the reaction of ROH and an amide is catalyzed by a general base (eq 1). It is generally believed that the formation of the tetrahedral intermediate occurs as a result of the nucleophilic attack on the amide.[90] As a next step, the protonated base donates its proton to the amide nitrogen and the CN bond breaks. Following this scheme, we first investigated the concertedness of the nucleophilic attack and the proton transfer from ROH to the base B. The calculations were carried out for the model system comprising formamide, methanol, and ammonia.

The effects of a base on the reaction surfaces were explored for two important bases, i.e., water and histidine. The study of the water-assisted reaction enables us to compare our results with the corresponding experimental observations, whereas the study of the histidine-assisted reaction provides a reference system for studies of serine proteases. To obtain free-energy surfaces for these bases, the ab initio surface calculated with ammonia as the base had to be corrected for the pK_a difference between ammonia and H_2O or histidine. These corrections consist of the full ΔpK_a differences in the regions where the bases are in their protonated forms, whereas the ΔpK_a corrections modified by the coefficient α were utilized for the transition state regions of proton-transfer steps (eqs 3−7). Note that these corrections are not employed to compensate for deficiencies of the ab initio methods, but they rather reflect the fact that the calculations were carried out with an ammonia as a base. Thus, the pK_a corrections enable us to build a surface for another base that has a different basicity.

Table 1. The Free Energy Surface for the Base-Catalyzed Formation of the Tetrahedral Intermediate in the Methanolysis of Foramide in Aqueous Solution[a] (I → III in eq 1)

	proton configuration[c]				
r(OC′)[b] (Å)	A	B	C	D	E
Ammonia as a Base (B = NH_3)					
5.0	1	3	4	8	9 [I1][d]
3.5	0 [R][d]				15
3.0	−1	13	18	18	16
2.5	2	12	18	19	20[e]
2.3	7	14	18	21	21[e]
2.2	9	16	19	22	23[e]
2.1	11	18	22	24	24[e] [TS1]
2.0	14	20	24	25	23
1.8	26	25	27	28	20
1.5	41	31	28	22	18 [I2]
Histidine as a Base (B = His)					
5.0	1	3	5	9	11 [I1]
3.5	0 [R]				17
3.0	−1	13	19	19	18
2.5	2	12	19	20	22[e]
2.3	7	14	19	22	23[e]
2.2	9	16	20	23	25[e]
2.1	11	18	23	25	26[e] [TS1]
2.0	14	20	25	26	25
1.8	26	25	28	29	22
1.5	41	31	29	23	20 [I2]
Water as a Base (B = H_2O)					
5.0	1	5	8	15	20
3.5	0 [R]				26[f]
3.0	−1	15	22	25	27
2.5	2	14	22	26	31[e]
2.3	7	16	22	28	32[e]
2.2	9	18	23	29	34[e]
2.1	11	20	26	31	35[e]
2.0	14	22	28	32 [TS1]	34
1.8	26	27	31	35	31
1.5	41	33	32	29	29 [I2]

[a] The table gives the pK_a-corrected free energies (kcal/mol). The system depicted in Figure 1 was used for the calculations. [b] The distance between the O and C′ (see Figure 1). The r(C′N′) distance was kept constant (1.4 Å), except for the structure with r(C′O) = 1.5 Å, where r(C′N′) was 1.5 Å. [c] A: The hydrogen H′ was bonded to the oxygen O, r(NO) = 3.0 Å, r(NN′) = 4.0 Å. B: r(OH′) = 1.125 Å, r(NO) = 2.5 Å, r(NN′) = 4.0 Å. C: r(OH′) = 1.25 Å, r(NO) = 2.5 Å, r(NN′) = 4.0 Å. D: r(NH′) = 1.125 Å, r(NO) = 2.5 Å, r(NN′) = 4.0 Å. E: The hydrogen H′ was bonded to the nitrogen N, r(NO) = 3.0 Å, r(NN′) = 4.0 Å. [d] Energies that were used to construct the free energy profile of Figure 8 are denoted with the corresponding symbols (R, I1, I2, TS1) in brackets. [e] r(NO) = 2.5 Å.

The calculated free energy surfaces are presented in Table 1. The structures of the gas-phase transition states for the reactions with NH_3 as a base and with water as a base are depicted in Figure 4. Note that the CN bond tends to become longer in the course of the nucleophilic attack. In amides, both CO and CN bonds have partial double bond character. However, when the nucleophilic attack is complete, the amide oxygen is negatively charged, the resonance is lost, and the CN bond becomes a single bond. The corresponding lengthening of the CN bond occurs gradually as the tetrahedral intermediate is being formed.

The actual shape of the reaction surface, which determines the concertedness of the reaction, can be viewed more conveniently in the contour representation of Figure 5. Clearly, the first step of the reaction with a water molecule as a base has a concerted character. The concerted pathway is favored by about 4 kcal/mol compared to the stepwise mechanism. While the reaction with water as a general base is of interest, our main point is to establish a reliable surface for the reaction where

(89) Hine, J.; King, R. S.-M.; Midden, W. R.; Sinha, A. *J. Org. Chem.* **1981**, *46*, 3186−3189.
(90) Bender, M. L.; Thomas, R. J. *J. Am. Chem. Soc.* **1961**, *83*, 4183−4189.

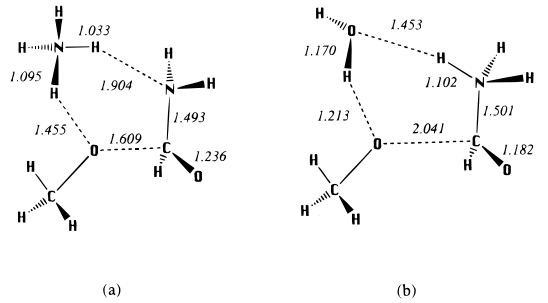


Figure 4. Gas-phase (HF/6-31G*) transition states of the general base-catalyzed methanolysis of formamide with (a) ammonia as a base and with (b) water as a base. Numbers in italics denote interatomic distances (Å).

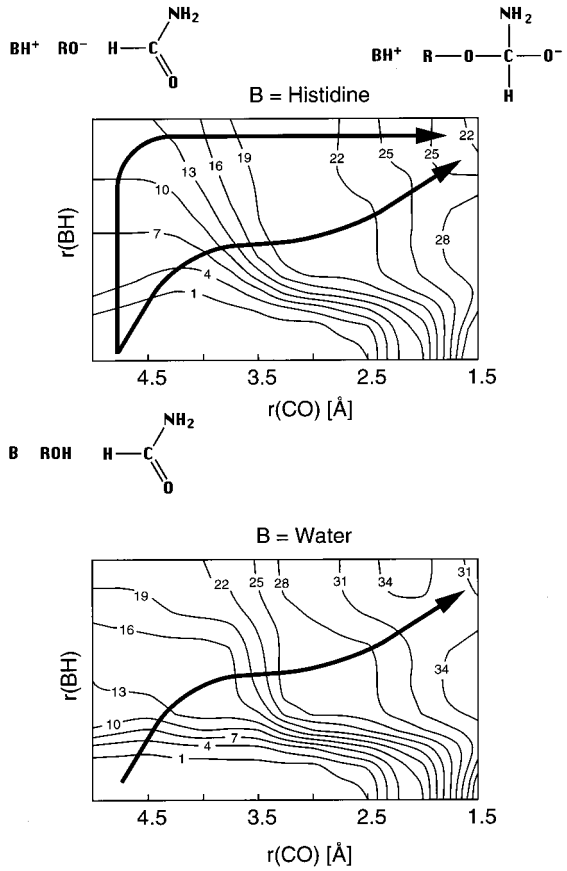


Figure 5. A schematic representation of the pK_a corrected free energy surfaces for the general base-catalyzed formation of the tetrahedral intermediate in the attack of methanol on formamide in aqueous solution (I → III in eq 1, Table 1) calculated using B3LYP$^+$+LD method. The isoenergetic levels (3 kcal/mol spacing) were plotted using the dgrid3d algorithm in the program GNUPLOT. Top, histidine as a base; bottom, water as a base. The arrows represent the least energy path(s) on the given surface.

histidine serves as a general base. Here, the reaction surface is flat enough to support both concerted and stepwise paths. This result is quite different than results obtained by Daggett et al.,[16] who concluded (based on gas-phase calculations) that the reaction is fully concerted.

The Unassisted Attack by R″OH. The presence of a base as an acceptor or donor of a proton is not, in principle, necessary for the nucleophilic attack to proceed. Therefore, in order to cover all the mechanistic alternatives for the solution reaction, we evaluated a mechanism in which the proton is transferred directly from the methanol to the nitrogen or oxygen of the substrate. The calculated free-energy barriers corresponding to

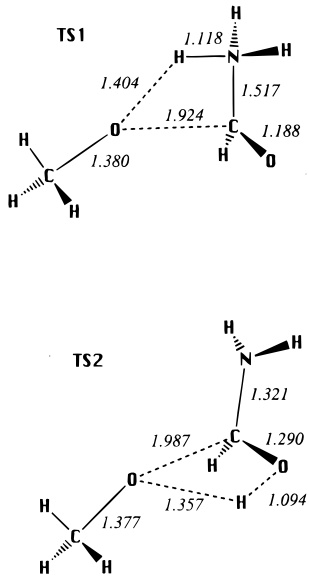


Figure 6. Gas-phase (HF/6-31G*) state structures of the unassisted nucleophilic attack of a CH_3OH molecule on formamide. Numbers in italics denote interatomic distances (Å).

the transition state structures TS1 and TS2 (Figure 6) amount to 39 and 41 kcal/mol, respectively. The geometries of these transition states correspond to the HF/6-31G* optimized structures. We are aware that a more rigorous approach would involve exploration of the two-dimensional free energy surface. However, considering the fact that the calculated barriers are about 7−9 kcal/mol higher than the barrier of the water-mediated proton transfer, it is clear that the direct proton-transfer mechanism is rather improbable. A similar conclusion was reached by Antonczak et al.,[38] who compared activation barriers for unassisted and H_2O-assisted neutral hydrolysis of formamide.

3.3. The Basicity of the Tetrahedral Intermediate. The basicity of the nitrogen of the leaving group, measured by the pK_a of the corresponding conjugated acid, varies significantly upon going from the reactants to the products of the acylation reaction. The extent of this variation and the relative basicities of a general acid and the $-NH_2$ and $-O^-$ groups of the tetrahedral intermediate (III, eq 1) represent important factors that determine the stability of this intermediate and the reaction mechanism. Unfortunately, despite its fundamental importance, the relevant pK_a cannot be obtained experimentally. Therefore, we calculated the absolute pK_a for the protonation of III (eq 1) and formamide on the N and O atoms using G2 model chemistry[54,55] and the LD solvation model. The results of these calculations (Table 2) indicate that the protonation of the oxygen atoms in the tetrahedral intermediate and formamide is favored over the protonation of the corresponding amino groups. Quantitatively, the calculated pK_a values of the nitrogen atoms of III and of formamide are in remarkable agreement with the estimates of Fersht[6] and Komiyama and Bender.[91] The calculated magnitude of the pK_a of the nitrogen atom of III can be further refined by noting that the G2+LD method overestimates the pK_a of ammonia by about 3 pK_a units. Thus, it seems reasonable to expect that the calculated pK_a for III would be overestimated by a similar amount. We invoke this comparison because the experimental pK_a of NH_4^+ is known and the same type of chemical bond (N−H) is formed by the N-protonation of the tetrahedral intermediate. Thus, assuming that the evaluation of the atomization energy of the N−H group by the

(91) Komiyama, M.; Bender, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 557−560.

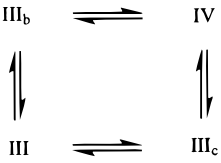
Table 2. A Comparison of the Calculated and Experimental pK_a of Key Substrates and Intermediates of the Serine Protease Reaction

molecule	G_{gas}[a]	ΔG_{gas}[b]	$\Delta\Delta G_{hydr}$[c]	pK_a (G2+LD)	pK_a (exp)
CH_3OH	−115.557 612	377.1	−94.7	16.8	15.6[d]
formamide(OH^+)	−169.984 616	192.5	63.6	−2.4	0 to −4[d]
formamide(NH_3^+)	−169.959 327	176.4	69.9	−9.6	−9.0[e]
III(OH)[g]	−285.164 727	353.7	−73.2	15.4 (14)[f]	
III(NH_3)[h]	−285.201 675	330.6	−56.8	10.5 (8)[f]	
NH_4^+	−56.801 017	198.3	78.5	12.7	9.3[d]

[a] Gas-phase free energy (au) of the protonated compound calculated at the G2 level. [b] $\Delta G_{gas} = G_{gas}(B) - G_{gas}(BH)$ (kcal/mol). [c] $\Delta\Delta G_{hydr} = \Delta G_{hydr}(B) - \Delta G_{hydr}(BH)$, where ΔG_{hydr} denotes the hydration free energy (kcal/mol) calculated by the LD solvation model. [d] Experimentally determined pK_a values taken from ref 98. [e] Experimental estimate (from eq 45 of ref 6). [f] The pK_a values given in parenthesis correspond to the calibrated G2+LD results and represent our best pK_a estimates (see text). [g] The tetrahedral intermediate (eq 1) with a protonated oxygen. [h] The tetrahedral intermediate (eq 1) with a protonated nitrogen. The C′N′ distance was kept fixed (1.51 Å) during the G2 calculation since this tetrahedral structure is not stable in gas phase. We chose a distance constraint of 1.51 Å since the CN bond has this length in the MP2/6-31G(d) fully optimized structure of the negatively charged TI and also in the N-protonated methylamine.

G2+LD level involves the same systematic error for both compounds and taking into account the estimate of Komiyama and Bender, we conclude that the pK_a of the nitrogen and oxygen atoms of III are around 8 and 14, respectively. Because a pK_a of 8 for the nitrogen atom of III is slightly higher than the pK_a of histidine, one can expect only a very shallow minimum on the free energy surface near the N-protonated intermediate. This minimum may completely disappear in the enzyme active site, where the effective pK_a of the histidine residue increases to about 9.[11] Under such circumstances, the reaction coordinate would become fully concerted, as suggested by Komiyama and Bender.[91]

The existence of a tetrahedral intermediate, which is supported by ^{18}O isotopic exchange in the substrate C=O group,[90,92−95] is further supported by the present calculations. To realize this we should consider the kinetic scheme of eq 12.



(12)

The four states in eq 12 include structures III and IV of eq 1 and two other structures that correspond to the O-protonated TI and to the N-protonated TI. These structures are referred to here as IIIb and IIIc, respectively. If the TI is a real intermediate, then we have two competing reactions, III ⇌ IIIb and III ⇌ IIIc. The existence of the first reaction (O-protonation) is supported by the above-mentioned ^{18}O isotope effect and the finding that the O of III has a higher pK_a than the N of III. The existence of the III ⇌ IIIc (N-protonation) is supported by the formation of the product IV, which is probably due to the IIIc → IV (the breakdown of the N-protonated TI). This requires the existence of a high barrier for the IIIb → IV reaction, which is consistent with the high pK_a of the protonated oxygen.

Table 3. The Free Energy Surface for the Breakdown of the Tetrahedral Intermediate in the Base-Catalyzed Methanolysis of Formamide in Aqueous Solution[a] (III → IV in eq 1)

	proton configuration[c]				
r(C′N′)[b] (Å)	A	B	C	D	E
	Ammonia as a Base (B = NH_3)				
1.4	21	26	27	27	26
1.5	18 [I2][d]	22	22 [TS2]	23	20 [I3]
1.7	25	28	27	26	20
1.9	38	37	35	33	22
2.0	43	40	38	35	23
2.1	47	43	42	36	23 [TS3]
2.2	50	44	44	36	23
2.5	54	47	47	34	21
3.0			44		18
3.5					8[e] [P]
	Histidine as a Base (B = His)				
1.4	23	27	28	27	26
1.5	20 [I2]	23	23 [TS2]	23	20 [I3]
1.7	27	29	28	26	20
1.9	40	38	36	33	22
2.0	45	41	39	35	23
2.1	49	44	43	36	23 [TS3]
2.2	52	45	45	36	23
2.5	56	48	48	34	21
3.0			45		18
3.5					8[e] [P]
	Water as a Base (B = H_2O)				
1.4	32	33	31	29	26
1.5	29 [I2]	29	26 [TS2]	25	20 [I3]
1.7	36	35	31	28	20
1.9	49	44	39	35	22
2.0	54	47	42	37	23
2.1	58	50	46	38	23 [TS3]
2.2	61	51	48	38	23
2.5	65	54	51	36	21
3.0			48		18
3.5					8[e] [P]

[a] The table gives the pK_a-corrected free energies (kcal/mol). The actual corrections are given in parentheses. The system depicted in Figure 1 was used for the calculations. [b] The distance between the C′ and N′ atoms (see Figure 1). The r(OC′) distance was kept constant (1.5 Å). [c] A: The hydrogen H′ was bonded to the nitrogen N, r(NO) = 3.0 Å, r(NN′) = 3.0 Å. B: r(NH′) = 1.125 Å, r(NO) = 3.0 Å, r(NN′) = 2.5 Å. C: r(NH′) = 1.25 Å, r(NO) = 3.0 Å, r(NN′) = 2.5 Å. D: r(N′H′) = 1.125 Å, r(NO) = 3.0 Å, r(NN′) = 2.5 Å. E: The hydrogen H′ was bonded to the nitrogen N′, r(NO) = 3.0 Å, r(NN′) = 3.0 Å. [d] Energies that were used to construct the free energy profile of Figure 8 are denoted with the corresponding symbols (I1, I3, TS2, TS3, P) in brackets. [e] No constraints on the N−O and N−N′ distances.

The pK_a values obtained at the G2+LD level (note that this is a substantially more rigorous ab initio approach than the B3LYP$^+$+LD method used by us to generate the reaction surface) can be utilized for the assessment of the accuracy of the B3LYP$^+$ reaction surface. Using pK_a values of 9.3 and 8 for ammonia and the N-protonated tetrahedral intermediate, the correct "experimental" free energy for a proton transfer from NH_4^+ to III is 2 kcal/mol. Because the corresponding free energy difference calculated at the B3LYP$^+$+LD level is around 2 kcal/mol, we can be quite certain that the free energy surface for the general base/acid catalysis is represented correctly by our calculations.

3.4. The Breakdown of the Tetrahedral Intermediate. The General Acid Catalysis. The results calculated for the breakdown of the tetrahedral intermediate, i.e., the reaction III → IV of eq 1, are given in Table 3. Here we assumed that the CN bond breaking and the proton transfer from the general acid to the leaving group nitrogen follows the formation of the tetrahedral intermediate III. This is a reasonable assumption, since the existence of this intermediate has been established by the carbonyl/oxygen exchange experiments using ^{18}O-labeled

(92) Bunton, C. A.; Nayak, B.; O'Connor, C. *J. Org. Chem.* **1968**, *33*, 572−575.
(93) Brown, R. S.; Bennet, A. J.; Slebocka-Tilk, H. *Acc. Chem. Res.* **1992**, *25*, 481−488.
(94) Brown, R. S.; Bennet, A. J.; Slebocka-Tilk, H.; Jodhan, A. *J. Am. Chem. Soc.* **1992**, *114*, 3092−3098.
(95) 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.

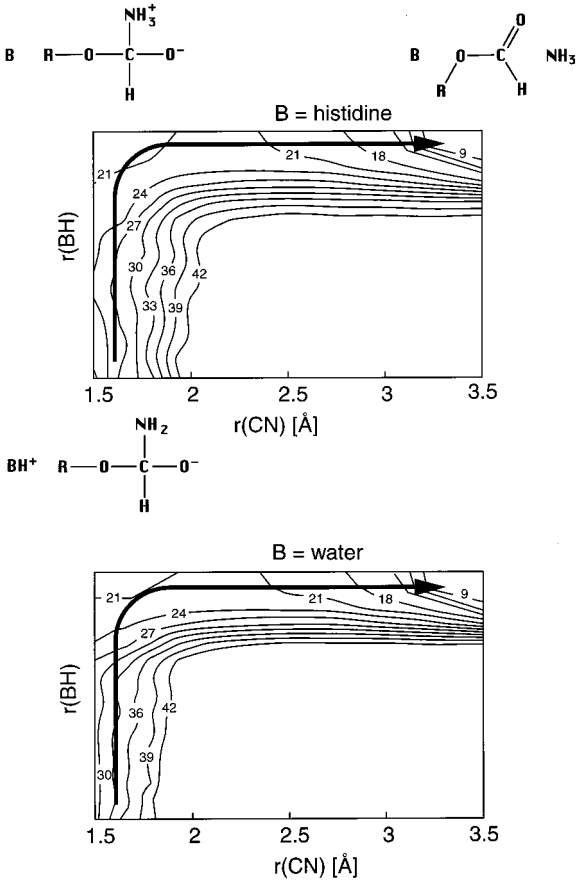


Figure 7. A schematic representation of the pK_a corrected free energy surface for the breakdown of the tetrahedral intermediate in the base-catalyzed methanolysis of formamide in aqueous solution (III → IV in eq 1, Table 3) calculated using B3LYP$^+$+LD method. The isoenergetic levels (3 kcal/mol spacing) were plotted using dgrid3d algorithm in the program GNUPLOT. Top, histidine as a base; bottom, water as a base. The arrows represent the least energy path(s) on the given surface.

compounds.[90,92−95] Furthermore, we focus here on a reaction path that does not involve the oxygen-protonated intermediate, because this mechanism requires the presence of an additional base that would assist in the proton transfer to the leaving group. Such a general base/acid catalyzed proton transfer from the oxygen to the nitrogen atom of the tetrahedral intermediate is necessary, since a direct proton transfer proceeding via a four-membered transition state has a prohibitively large barrier.[21,38] However, such a base is not readily available in protein active sites, which makes this reaction a less suitable candidate for the reference solution reaction for studies of serine proteases.

Because the nitrogen atom of the tetrahedral intermediate is more basic than both general acids considered in this study, the proton transfer between the protonated histidine and the tetrahedral intermediate precedes the CN bond cleavage (Table 3 and Figure 7). As a result, the height of the barrier for the CN cleavage is not affected by the nature of the general acid involved. If the protonated histidine serves as a general acid, this proton transfer is associated only with a small (2 kcal/mol) barrier, whereas for H_3O^+ this barrier entirely disappears. The transition state region for the CN bond cleavage is rather flat and features a maximum at a C−N distance of about 2.2 Å. More importantly, we found that the activation free energy for the breakdown of the tetrahedral intermediate is about 3 kcal/mol lower than the barrier for the histidine-catalyzed nucleophilic attack. This difference further increases when water serves as a general acid.

The Overall Reaction Profile. The free energy diagram for the overall reaction is depicted in Figure 8. Obviously, the rate-determining step is the nucleophilic attack. The activation barrier for the reaction with water as a base is 32 kcal/mol, which is in good agreement with the experimental value (30−32 kcal/mol). The negatively charged tetrahedral structure is not a stable intermediate (though with histidine as a base there might be some shallow local minimum); it converts immediately into the structure with protonated nitrogen. This intermediate is also only quasistable, as only 3 kcal/mol is required to break the CN bond in the next step. The free energy change for the I → IV reaction (see eq 1) is predicted by our calculation to be ∼8 kcal/mol. Naturally, this energy is independent of the nature of the general base/acid. The experimental estimate reported by Guthrie[8] for this reaction free energy is 11 kcal/mol.

To validate our pK_a correction procedure we reevaluated crucial regions on the potential surface of the water-assisted methanolysis of formamide using direct ab initio calculations (rather than obtaining this surface by interpolating the corresponding surface of the reaction with NH_3 as a base). The results, which are presented in Figure 9, were obtained by the B3LYP$^+$+LD calculation in which the water molecule was considered explicitly. These results can be compared with the results of Figure 5 and Tables 1 and 3 that were determined using the pK_a-correction procedure. Here we point out that the predicted overall free-energy barrier is 34 kcal/mol for the calculation with water molecule treated explicitly, while it is 32 kcal/mol when the results for ammonia as a general base are extrapolated to the system involving water as a general base. In addition, both computational methods yield very flat free energy surfaces.

To analyze the results presented in Figure 9 it is useful to consider the geometries and energies of the two transition states that served us as a starting point for the calculations of the reaction coordinate. As seen from the Figure 4b, the gas-phase transition state of this system involves a highly concerted path, where the proton transfer from CH_3OH to H_2O is concerted with the nucleophilic attack and also, at the same time, with the proton transfer from the catalytic water to the amide nitrogen and with the CN bond breaking. This transition state can be described in the two-dimensional diagram of Figure 9, where each axis represents the indicated concerted motion. The gas-phase path does not involve an oxyanion intermediate, but this is not necessarily the situation in solution. Here, the system can move toward the oxyanion region (upper left-hand corner of Figure 9), where the surface is very shallow. Since entropic effects should make the four-bond transition state less likely, it seems to us that in solution the concerted proton transfer and nucleophilic attack (the motion along the (r_1, r_2) direction) and the formation of an oxyanion type transition state occur before the collapse of the system along the (r_3, r_4) coordinate. At any rate, the direct estimate of the activation barrier on the free energy surface in solution is similar to that obtained by our indirect interpolation procedure.

4. Concluding Remarks

The catalytic power of enzymes cannot be understood without relating it to the corresponding reactions in solution. It is quite obvious by now that gas-phase calculations do not provide proper potential surfaces for reactions in solution (and of course in enzymes). Thus, it is important to focus on obtaining reliable results from solution calculations. The present work evaluated the potential surface for the reference solution reaction that is needed for understanding the mechanism of action of serine

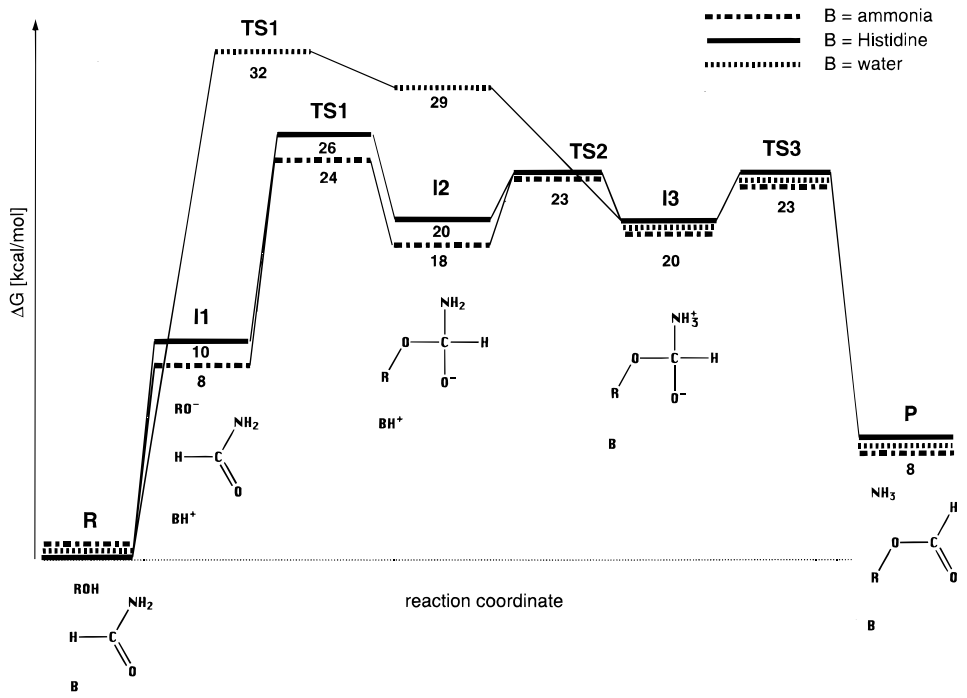


Figure 8. Calculated (B3LYP$^+$+LD) reaction profile of the reference solution reaction for studies of serine protease with three different bases: ammonia, water, and histidine. The relative free energies are given in kcal/mol.

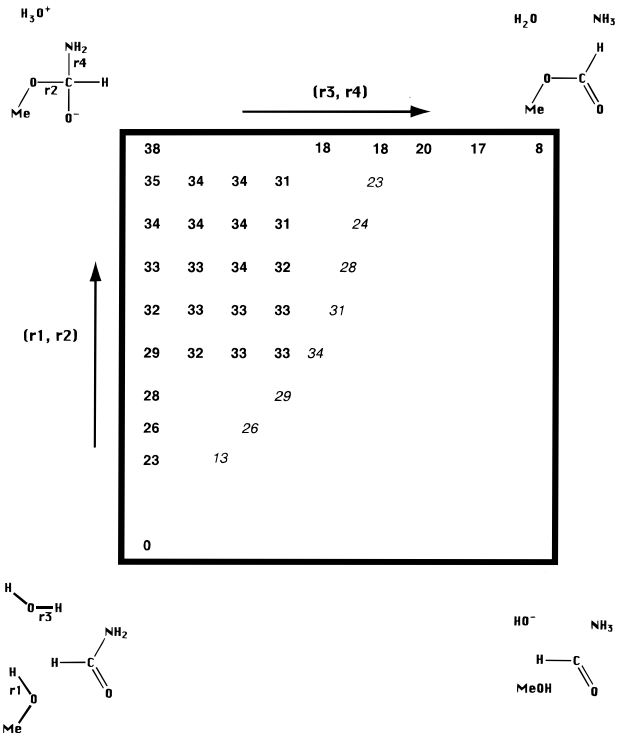


Figure 9. Free energy surface (B3LYP$^+$+LD, kcal/mol) of the general base catalyzed methanolysis of formamide calculated for water as the general base. The water molecule was included explicitly in the quantum mechanics and LD calculations (see text). The geometries for the free energy points given in italics were generated by the gas-phase intrinsic reaction coordinate (IRC) calculation at the HF/6-31G* level. The fully optimized geometry of the corresponding TS was used for this IRC calculation.

proteases. This was done using a judicious combination of high level ab initio theory, the LD solvation model and available experimental data. Thus, we believe that our study has provided a reasonably reliable solution surface. Several mechanistic points can be deduced from our potential surface. First, our study indicates that the methanolysis of formamide with water as the general base is concerted and that the barrier for the formation of the neutral tetrahedral intermediate (see the zwitterion in Figure 8) is significantly higher than the barrier for the cleavage of the CN bond (i.e., 32 and 23 kcal/mol respectively). Second, the surface of Figure 5 indicates that the reaction with histidine as the general base involves a stepwise mechanism with a shallow surface that can allow for some population of both the stepwise and concerted pathways. Thus, it can also support a substantial solvent kinetic isotope effect (KIE). As indicated by early EVB calculations,[27] a similar shallow surface may occur in serine proteases. Therefore, the use of the observed solvent KIE as a proof of a concerted mechanism in serine proteases[25] is not justified. Third, we provide a reliable estimate for the acidity constant (the pK_a) of the amide nitrogen in the tetrahedral intermediate (TI). This constant has been long considered as one of the missing parts in the puzzle of serine proteases.[91] Here, our prediction of the p$K_a \sim 8$ indicates that the protonation of the leaving group may occur prior to the breakdown of the CN bond. This means that the potential surface of the second step of the reaction supports a stepwise mechanism. Finally, our calculations indicate that the formation of the TI is the rate-determining step for our reaction, regardless of the basicity of the assisting general base. That is, the calculated barrier for this step with histidine as the base ($\simeq$26 kcal/mol) is higher than the subsequent barriers ($\sim$23 kcal/mol). However, this result also tells us that all three barriers for the histidine-assisted reaction are of similar height.

The present work focuses on the acylation step (eq 1) of the reference solution reaction for serine proteases. The deacylation step has also been examined in a somewhat less systematic way by the B3LYP$^+$+LD model. Interestingly, it was found that the overall rate-determining step in aqueous solution corresponds to the deacylation reaction. The calculated free energy difference between R of Figure 8 and the highest TS of the deacylation step is around 34 kcal/mol. This preliminary result indicates that the serine proteases catalyze the deacylation step more than

the acylation step (since in the enzyme the acylation step is rate limiting).

Since our potential surfaces for the solution reaction are probably quite accurate, it will be interesting to revisit the reaction in the corresponding proteins. For this purpose we can use the solution study to recalibrate empirical valence bond (EVB) surfaces. These surfaces can be used both in standard EVB simulations or, as a consistent reference potential, in ab initio FEP calculations of enzymatic reactions.[44] Such studies will allow one to explore the catalytic effect of the enzyme with greater certainty about the relevant potential surfaces. It should also allow the exploration of the mechanistic implications of observed kinetic isotope effects studies (see above). This can be done by using the EVB-calibrated surface and a centroid path integral approach (e.g. ref 96). Finally, the use of the EVB approach and our new method of calculating activation entropies[62] should help in determining the role of entropic effects in enzyme catalysis as well as in providing insight about the entropic contributions to solution reactions.

Acknowledgment. This work was supported by the NIH grant GM24492.

Supporting Information Available: Breakdown of gas phase and solvation energies, ZPE and entropy, and numerical justification of eq 2 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992441S

(96) Hwang, J.-K.; Warshel, A. *J. Am. Chem. Soc.* **1996**, *118*, 11745−11751.
(97) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 10282−10288.
(98) Lowry, T. H.; Richardson, K. S. *Mechanism and Theory in Organic Chemistry*; McGraw-Hill: New York, 1987.