Quantum Chemical Studies of a Model for Peptide Bond Formation. 2. Role of Amine Catalyst in Formation of Formamide and Water from Ammonia and Formic Acid

Tetsuro Oie,[†] Gilda H. Loew,*[†] Stanley K. Burt,[‡] and Robert D. MacElroy[§]

Contribution from the Molecular Theory Laboratory, The Rockefeller University, Palo Alto, California 94304, the Molecular Research Institute, Palo Alto, California 94304, and the NASA Ames Research Center, Moffett Field, California 94035. *Received July 26, 1982*

Abstract: The S_N2 reaction between formic acid and two ammonia molecules with the second ammonia as a catalyst has been studied as a model reaction for amine-catalyzed peptide bond formation using an ab initio molecular orbital method. As in previous studies of the uncatalyzed reaction between formic acid and ammonia, two reaction mechanisms have been examined, i.e., a two-step and a concerted reaction. The stationary points of each reaction including intermediate and transition states have been identified and free energies calculated for all geometry-optimized reaction species to determine the thermodynamics and kinetics of each reaction. A comparison with uncatalyzed amide bond formation indicates that the participation of the second ammonia molecule reduces the internal energy of activation but increases the entropy contribution to the free energy of activation, with a net decrease in free energy of activation about 10 kcal/mol for the two-step reaction and only 2 kcal/mol for the concerted reaction. In contrast to the uncatalyzed reaction where the two mechanisms were competitive, the ammonia-catalyzed two-step reaction was found to be more favorable than the concerted one, and for the former the free energy of activation of the second step was slightly higher than that of the first step. The basis set dependence and effect of correlation were also investigated by using a number of basis sets up to 6-31G** and estimates of correlation energy by Møller–Plesset perturbation theory at the second order. Correlation and polarization effects were found to be significant factors in stabilization and destabilization, respectively, of transition states.

Introduction

The molecular events that preceded the formation of life are unknown. For this reason, speculative but plausible models are needed to describe the emergence of life from nonliving precursors. Such models must include descriptions of the origins of the transfer of information from a "genetic" polymer (e.g., DNA or RNA) to a catalytic molecule (protein). Extant processes, involving transfer RNA, messenger RNA, ribosomes, acylases, etc., are ubiquitous and, for this reason, offer very few clues to the origin of either the molecules involved or their functions.

However, the ubiquity of these processes, including the development of the genetic code, suggests either that they were essential to the origin of life or that they were so evolutionarily successful that they rapidly and completely replaced all competing biological processes. In either case, a more simple beginning to the transfer of information from nucleic acids to protein must be sought.

One approach has been to address the problem of amino acid polymers in the absence of enzymes and to use the information collected to postulate the environmental requirements for the formation of prebiotic proteins.

Thus, there have been numerous attempts[1] to synthesize polypeptides from their components amino acids under conditions that resemble those of the primitive earth. Attempts have been made to activate amino acids by a variety of methods including esterification, phosphorylation, and the presence of divalent cations (e.g., Mg^{2+}, Ca^{2+}, Zn^{2+}, Cu^{2+}, Ni^{2+}), since none of the amino acid forms present in solution at any pH (i.e., cationic, zwitterionic, and anionic) are viable candidates for peptide bond formation because of the presence of the protonated amino or anionic carboxylate group.[2] However, the design of effective experiments and the interpretation of results have been hampered by a lack of detailed information of the mechanism of peptide bond formation. Therefore, we have begun a systematic study of this mechanism under a variety of conditions using the techniques of theoretical chemistry.

As a first step in understanding the mechanism of peptide bond formation, we have reported in a previous paper[3] a thorough investigation of the amide bond formation between formic acid and ammonia using MNDO and ab initio molecular orbital methods. The reaction

$$NH_3 + HCOOH \rightarrow NH_2CHO + H_2O$$

should closely correspond to peptide bond formation between neutral amino acids or their alkyl esters. For this amide bond formation, free energies of activation calculated by ab initio method indicated that the two reaction mechanisms examined, i.e., a two-step and a concerted reaction, are fairly competitive in the gas phase or in a hydrophobic environment.

In the study reported here we extend our investigation to the effect of amine catalysis on the mechanism of amide bond formation. Kinetic studies[4] of aminolysis reactions of aliphatic esters (e.g., methyl or ethyl formate) with the primary and secondary amines (e.g., *n*-propylamine, *n*-butylamine, glycylglycine, morpholine) have revealed a rate law containing a term predominantly second order in amine at moderate and high concentrations of amine, indicating that such reactions are susceptible to amine catalysis by a second molecule of the amine. Similarly peptide bond formation could also be amine catalyzed by a second molecule of amino acid.

Despite numerous kinetic studies[4,5] of ester aminolysis, the detailed mechanism of the amine-catalyzed reaction remains uncertain in most instances. Three questions are of primary

[†] The Rockefeller University.
[‡] Molecular Research Institute.
[§] NASA Ames Research Center.

(1) (a) Weber, A. L.; Orgel, L. E. *J. Mol. Evol.* **1981**, *17*, 190. (b) Lahav, N.; White, D. H. *Ibid.* **1980**, *16*, 11. (c) Weber, A. L.; Orgel, L. E. *Ibid.* **1980**, *16*, 1. (d) Mullins, D. W.; Lacey, J. C. *Ibid.* **1980**, *15*, 339. (e) Weber, A. L.; Orgel, L. E. *Ibid.* **1979**, *13*, 185. (f) Weber, A. L.; Orgel, L. E. *Ibid.* **1978**, *11*, 189. (g) Weber, A. L.; Orgel L. E. *Ibid.* **1978**, *11*, 9. (h) Steinman, G.; Cole, M. N. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 735. (i) Steinman, G. *Arch. Biochem. Biophys.* **1967**, *121*, 533. (j) Steinman, G.; Kenyon, D. H.; Calvin, M. *Biochem. Biophys. Acta* **1966**, *124*, 339. (k) Ponnamperuma, C.; Peterson, E. *Science (Washington, D.C.)* **1965**, *147*, 1572. (l) Steinman, G.; Kenyon, D. H.; Calvin, M. *Nature (London)* **1965**, *206*, 707. (m) Steinman, G.; Lemmon, R. M.; Calvin, M. *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *52*, 27. (n) Yamada, S.; Wagatsuma, M.; Takeuchi, Y.; Terashima, S. *Chem. Pharm. Bull.* **1971**, *19*, 2380. (o) Yamada, S.; Terashima, S.; Wagatsuma, M. *Tetrahedron Lett.* **1970**, 1501.

(2) Hay, R. W.; Porter, L. J. *J. Chem. Soc.* **1967**, 1261.

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

(4) (a) Bunnett, J. F.; Davis, G. T. *J. Am. Chem. Soc.* **1960**, *82*, 665. (b) Jencks, W. P.; Carriuolo, J. *Ibid.* **1960**, *82*, 675. (c) Blackburn, G. M.; Jencks, W. P. *Ibid.* **1968**, *90*, 2638.

(5) (a) Jencks, W. P.; Gilchrist, M. *J. Am. Chem. Soc.* **1968**, *90*, 2622. (b) Bruice, T. C.; Hegarty, A. F.; Felton, S. M.; Donzel, A.; Kundu, N. G. *Ibid.* **1970**, *92*, 1370. (c) Jencks, W. P.; Fersht, A. R. *Ibid.* **1970**, *92*, 5442.

0002-7863/83/1505-2221$01.50/0 © 1983 American Chemical Society

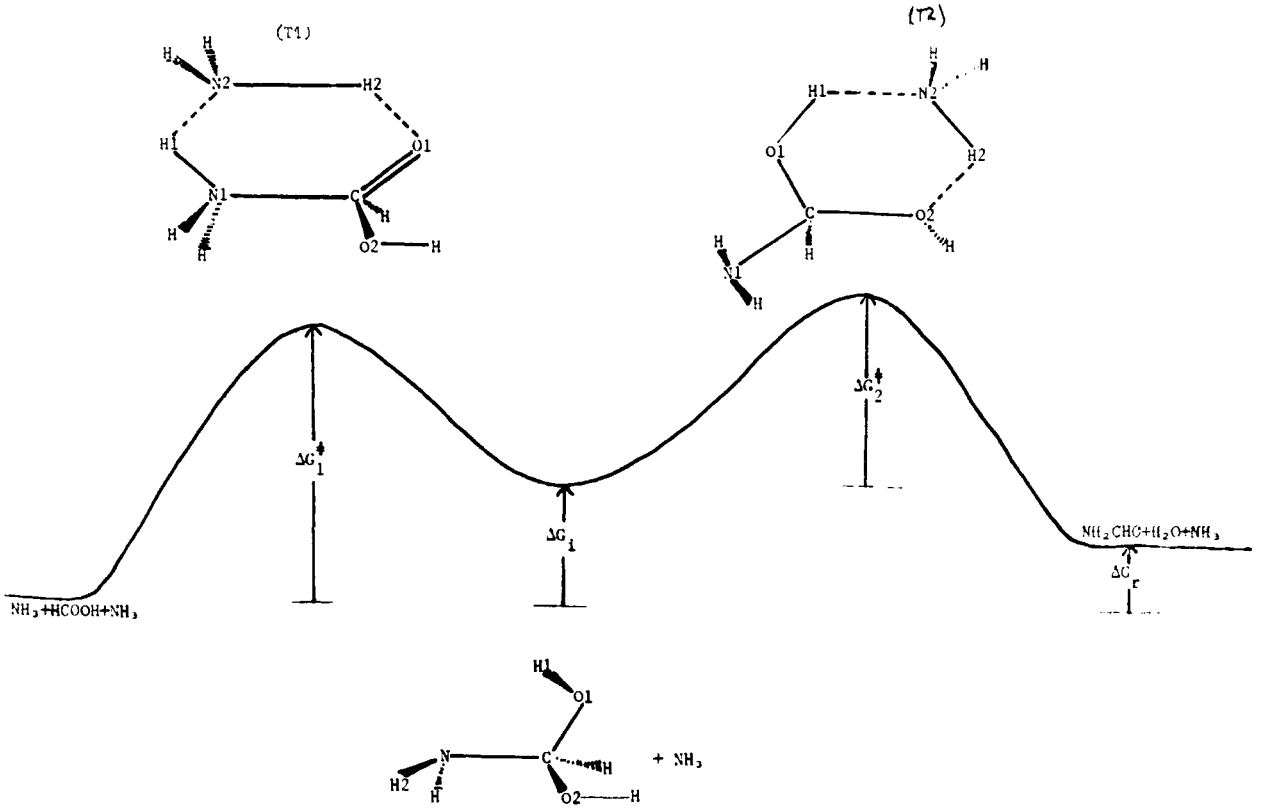

Figure 1. Two-step reaction mechanism for the catalyzed ammonia–formic acid reaction and the free-energy diagram of the reaction.

interest: the role of the second amine molecule in catalyzing amide bond formation, whether such catalysis favors the concerted or two-step reaction via a tetrahedral intermediate, and whether the formation or the breakdown of the intermediate is the rate-determining step, if such an intermediate is formed. Indirect experimental evidence for the existence of an intermediate has been reported for the reverse reaction of ester aminolysis, i.e., hydrolysis of amides in two specific instances: the analysis of hydrolysis product of the methoxymethylenemorpholinium cation,[4c] which presumably generates the same tetrahedral intermediates upon hydrolysis as is formed in the reaction of methyl formate with morpholine (secondary amine); study[6] of carbonyl oxygen exchange with the medium during hydrolysis of benzamide-^{18}O. The results of these experiments support the hypothesis that amide or peptide bond formation proceeds through a tetrahedral intermediate. Nevertheless, such an intermediate has never been directly observed nor isolated in any single case of ester aminolysis or peptide bond formation.

Theoretical studies by quantum mechanical methods can provide valuable detailed information about reaction mechanisms, i.e., the nature of transition states and metastable intermediates, which cannot be readily characterized by experimental techniques.

Therefore, we report here the results of extensive studies of amide bond formation between formic acid and ammonia in the presence of a second ammonia molecule by using ab initio molecular orbital method. For the reaction

$$NH_3 + HCOOH + NH_3 \rightarrow NH_2CHO + H_2O + NH_3$$

the second ammonia is a catalyst.

As in the previous study of the uncatalyzed reaction, we have considered two possible mechanisms for this catalytic reaction: a two-step reaction (Figure 1) and a concerted reaction (Figure 2).

The catalytic role of the second ammonia is defined by the nature of the transition states for each reaction. As seen in these figures, it serves as both acid and base, i.e., simultaneous proton

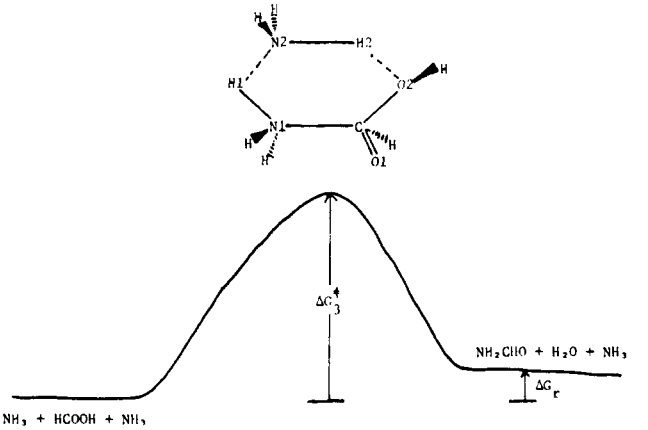

Figure 2. Concerted mechanism for the catalyzed ammonia–formic acid reaction and the free-energy diagram of the reaction.

donor and acceptor in both reactions. In the two-step mechanism, the reaction takes place through a stable of metastable intermediate (INT). The first step, reactant to intermediate, represents C–N bond formation through a nucleophilic attack of the nitrogen N1 of one ammonia on a carbon atom and simultaneous transfer of two hydrogen atoms from one ammonia to the second and from the second to a carbonyl oxygen atom. In the second step, intermediate to product, a water molecule is released by O–H and C–O bond cleavages with the help of ammonia, which again acts as a hydrogen atom relay system. Other molecules with proton-donor and -acceptor capabilities such as water could also catalyze one of these steps, a possibility that is consistent with the observed rate law.

In the second proposed mechanism, the reaction is concerted: both C–N bond formation and release of water by N–H and C–O bond cleavages take place simultaneously, again with the help of the second ammonia acting as a proton-relay catalyst. However, in contrast to the first mechanism, in which a hydrogen atom is transferred to a carbonyl oxygen, the concerted mechanism involves transfer of a hydrogen atom to the hydroxyl oxygen.

(6) Bender, M. L.; Ginger, R. D.; Kemp, J. C. *J. Am. Chem. Soc.* **1954**, *76*, 3350.

Table I. Ab Initio Optimized Structures of Transition States and Intermediate[a]

parameter[b]	STO-3G	3-21G	parameter[b]	STO-3G	3-21G
		Transition State T1			
N1-C	1.785	1.587	∠H2N2H1	84.4	86.6
C=O1	1.287	1.320	∠N2H1N1	153.8	155.4
C-O2	1.416	1.441	∠H1N1C	100.2	100.5
N1-H1	1.107	1.322	∠N1CO2	101.9	99.9
N2-H1	1.462	1.265	τH1N1CO1	26.4	−23.9
N2-H2	1.092	1.096	τN1CO1H2	24.7	−28.4
O1-H2	1.421	1.490	τCO1H2N2	24.6	−16.0
∠N1CO1	108.1	112.5	τO1H2N2H1	−4.3	1.0
∠CO1H2	108.1	107.9	τH2N2H1N1	−1.3	−1.5
∠O1H2N2	157.6	151.3	τN2H1N1C	−10.9	13.6
		Transition State T2			
N1-C	1.468	1.393	∠H1N2H2	85.8	89.4
C-O1	1.283	1.298	∠O1H1N2	153.6	152.5
C-O2	1.736	1.756	∠H1O1C	107.7	111.2
O1-H1	1.404	1.456	∠N1CO1	119.4	117.8
N2-H1	1.099	1.106	τH2O2CO1	28.8	19.8
O2-H2	1.068	1.363	τN2H2O2C	−12.4	−10.6
N2-H2	1.393	1.155	τH1N2H2O2	−0.8	1.1
∠O1CO2	107.1	107.3	τO1H1N2H2	−5.1	−2.6
∠H2O2C	100.4	100.6	τCO1H1N2	27.1	17.5
∠N2H2O2	155.9	154.6	τO2CO1H1	−31.0	−21.3
		Transition State T3			
N1-C	1.640	1.555	∠H1N2H2	85.7	89.6
C=O1	1.249	1.260	∠N1H1N2	155.0	150.9
C-O2	1.563	1.598	∠H1N1C	107.9	108.3
N1-H1	1.323	1.607	∠N1CO1	114.9	115.6
N2-H1	1.177	1.095	τH1N1CO2	9.6	15.6
O2-H2	1.224	1.388	τH2O2CN1	−8.7	−13.6
N2-H2	1.182	1.130	τN2H2O2C	2.6	1.5
∠N1CO2	99.9	101.0	τH1N2H2O2	2.9	5.5
∠H2O2C	112.5	114.9	τN1H1N2H2	1.9	4.4
∠N2H2O2	158.1	152.7	τCN1H1N2	−8.8	−16.7
		Intermediate (INT)			
C-N	1.482	1.428	τH1O1CN	44.1	30.7
C-O1	1.426	1.418	τH2NCO1	−88.3	−124.6
C-O2	1.434	1.421	τH2NCO2	33.7	−4.7
∠NCO1	111.7	112.2	τH3O2CO1	−51.2	−35.1
∠NCO2	108.2	105.3			

[a] See Figures 1-5 for the structures. [b] Bond length (X-Y or X=Y) in Å; bond angles (∠XYZ) and dihedral angle (τWXYZ) in degrees, where W, X, Y, and Z refer to atoms in Figures 1-5. Note that the intermediates formed in the catalyzed and uncatalyzed reactions are the same.

Methods and Results

To characterize these two reaction mechanisms, standard ab initio molecular orbital calculations were carried out by using the GAUSSIAN 80 program.[7] Analytical gradient optimization methods were used to locate the minima, corresponding to reactant, intermediate and product, and saddle points corresponding to transition states, using both STO-3G[8] and 3-21G[9] basis sets (convergence of root mean square of gradients is 0.0003 hartree/bohr or radian). For completely asymmetric structures such as transition states, such optimization involved 33 degrees of freedom. All transition states were identified by finding one negative eigenvalue of the analytic force constant matrix.

For all stationary points found by STO-3G and 3-21G basis sets, the harmonic vibrational frequencies and normal coordinates were determined by solving for eigenvalues and eigenvectors of the mass-weighted Cartesian force constant matrix, $\mathbf{M}^{-1/2}\mathbf{F}\mathbf{M}^{-1/2}$, where $\mathbf{M}$ is a diagonal matrix with atomic masses and F is a Cartesian force constant matrix.[10] For each of the three transition states observed, the eigenvector corresponding to a negative eigenvalue (i.e., the reaction coordinate) of this matrix is displayed in Figures 3–5 as a set of normalized mass-weighted Cartesian displacement vectors, $\mathbf{M}^{1/2}\mathbf{X}$, where $\mathbf{X}$ is the Cartesian displacement vector.

In Figures 3–5, the perspective structure of each of the three transition states obtained by both basis sets is also shown. The optimized geometrical parameters found for the transition states and intermediate by both STO-3G and 3-21G basis sets are given in Table I. Reported values represent convergence to within 0.001 Å for bond length and within 0.1° for bond and dihedral angles. The net charges at the transition states on both ammonia molecules, nucleophilic and catalyst, and on the leaving water molecule are given in Table II.

Table II. Net Positive Charge on Incoming (NH_3) and Leaving (H_2O) Molecules at Transition States

state	molecule[b]	charge[a] STO-3G	charge[a] 3-21G
T1	$(NH_3)_n$	0.22	0.24
	$(NH_3)_c$	0.01	0.10
T2	$(NH_3)_c$	0.08	0.14
	H_2O	0.13	0.04
T3	$(NH_3)_n$	0.19	0.20
	$(NH_3)_c$	0.13	0.20
	H_2O	0.15	0.10

[a] Unit of charge = $+e$. [b] Subscripts n and c for NH_3 refer to nucleophile and catalyst, respectively.

Table III. Effect of Correlation on the Relative Energies[a]

basis set	T1	T2	T3	INT	R
3-21G(MP2-HF)	−6.1	−15.7	−15.6	10.3	2.7
6-31G(MP2-HF)	−8.9	−15.2	−12.5	7.9	2.1
6-31G*(MP2-HF)	−21.9	−19.6	−24.1	−1.0	−0.7
6-31G**(MP2-HF)	−17.0	−19.3	−23.5	−1.0	−0.1

[a] The difference between perturbative and Hartree-Fock energies (in kcal/mol) are given. All the energies are relative to the energy of reactant, except that the energies of T2 are relative to that of INT, which is the sum of the energies of the intermediate and NH_3. The energies of product relative to those of the reactant are given in the last column. All the geometries were obtained with HF/3-21G.

Table IV. Effects of Basis Set Size and Polarization on Relative Energies[a]

basis set	T1	T2	T3	INT
HF/6-31G − HF/3-21G	15.8	10.1	14.4	2.2
HF/6-31G* − HF/6-31G	17.0	12.6	18.3	9.5
HF/6-31G** − HF/6-31G	11.4	12.4	17.6	8.1

[a] The difference in Hartree-Fock energies (in kcal/mol) with different basis sets are given. See also footnote *a* in Table III for the definition of the relative energies. All the geometries were obtained at HF/3-21G.

The STO-3G and 3-21G optimized geometries were then used for single-point energy calculations with more extensive basis sets and including electron correlation effects. Specifically 6-31G,[11] 6-31G* (6-31G + d-type polarization functions on heavy atoms),[12] and 6-31G** (6-31G* + p-type polarization functions on hydrogen atoms)[12] basis sets were used for single-point calculation with 3-21G optimized geometries and 3-21G and 6-31G basis sets with STO-3G optimized geometries.

The effects of valence-electron correlation were incorporated at the level of second-order Møller–Plesset (MP2) perturbation theory[13] with a 6-31G** basis set and 3-21G optimized geometries. Previous work on uncatalyzed amide bond formation between NH_3

(7) Binkley, J. S.; Whiteside, R. A.; Krishnan, R.; Seeger, R.; DeFrees, D. J.; Schlegel, H. B.; Topiol, S.; Kahn, L. R.; Pople, J. A. "Gaussian 80", QCPE 406, Indiana University.
(8) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.
(9) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
(10) Wilson, E. B.; Decius, J. C.; Cross, R. C. "Molecular Vibrations"; McGraw-Hill: New York, 1955.
(11) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
(12) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(13) (a) Møller, C.; Plesset, M. *Phys. Rev.* **1934**, *46*, 618. (b) Binkley, J. S.; Pople, J. A. *Int. J. Quant. Chem.* **1975**, *9*, 229. (c) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quant. Chem. Symp.* **1976**, *10*, 1. (d) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quant. Chem.* **1978**, *14*, 545.

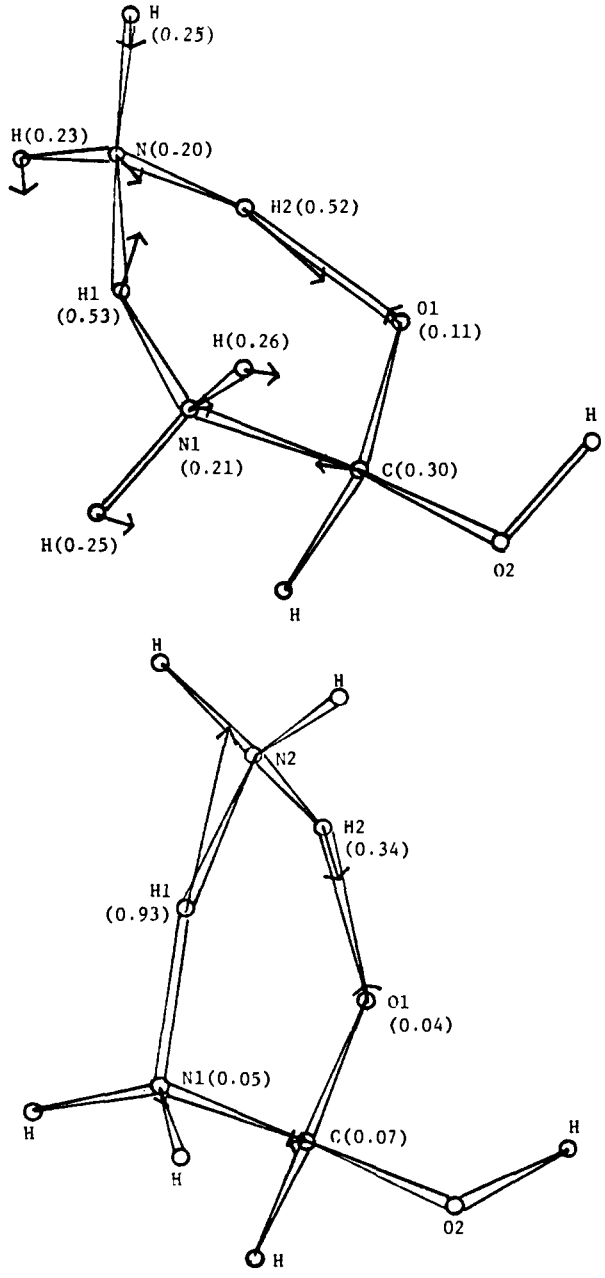

Figure 3. (a, top) Structure of transition state T1 optimized by STO-3G basis set. The components of the normalized reaction coordinate are shown as mass-weighted Cartesian displacement vectors of each atom with the relative amplitude given in parentheses and the direction shown by the arrow. (b, bottom) Structure of transition state T1 optimized by 3-21G basis set. The reaction coordinate is also given.

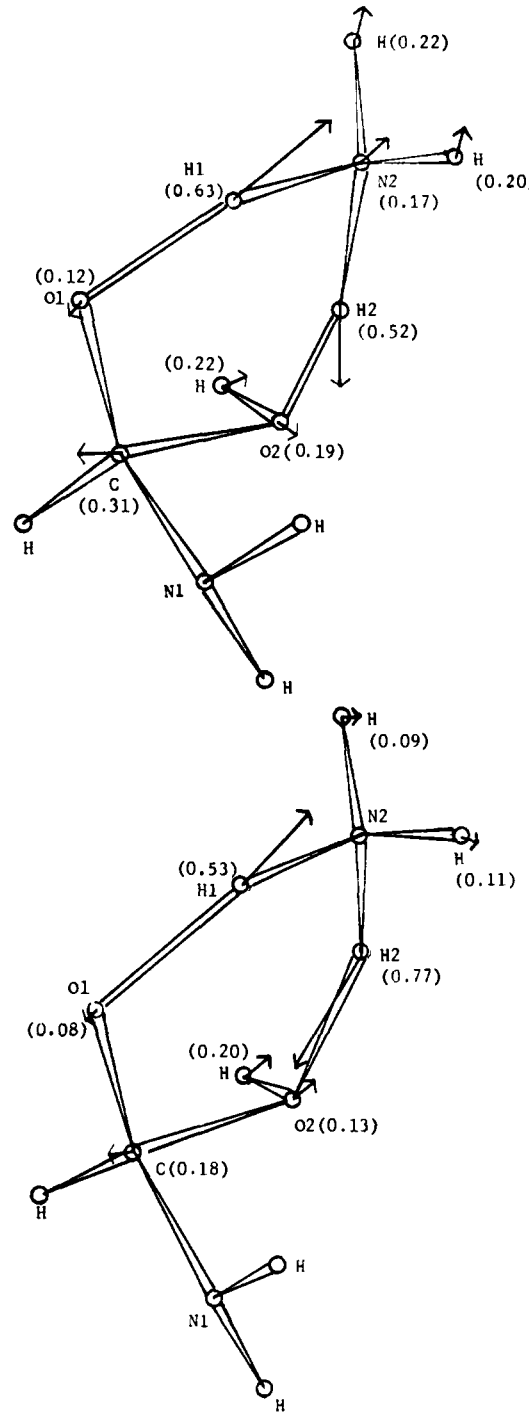

Figure 4. (a, top) Structure of transition state T2 optimized by STO-3G basis set. The reaction coordinate is also given. (b, bottom) Structure of transition state T2 optimized by 3-21G basis set. The reaction coordinate is also given.

and HCOOH showed that the difference in correlation energies obtained at the fourth- and second-order Møller–Plesset perturbation method with 6-31G** basis sets is nearly constant (4–5 kcal/mol) for the three transition states and that the order of the relative stabilities of these transition states remains unchanged. Therefore, the MP2 treatment appears to be a reliable means of taking into account the effect of electron correlation on the relative stabilities of the similar transition states of interest here. With all the smaller basis sets used here, MP2 calculations were also carried out in order to see the effects of correlation on the relative stabilities of the reaction species. These results and the effects of basis set size and polarization are summarized in Tables III and IV.

The validity of using a single-energy calculation with STO-3G and 3-21G optimized geometries to estimate the effect of basis set size and polarization function on the energies of activation and reaction is based on the assumption that the optimized geometries obtained at the lower level will not be significantly different from those obtained with the more extensive basis set or that any differences will not significantly affect the relative energetics. One set of calculations made here have allowed us to test this assumption, namely, comparison of the energetics obtained by single-point 3-21G calculations with STO-3G optimized geometries with those obtained totally from the 3-21G optimized structures. This comparison is shown in Table V, i.e., HF/3-21G//A vs. HF/3-21G//B. As seen in Table V, the comparisons made at this level indicate that single-point calculations reproduce the energetics of the reoptimized structures extremely well. However, such behavior may not be generalized to higher levels. Continued comparisons are necessary to further explain this point.

STO-3G and 3-21G optimized geometries were also used as the basis for calculating entropies and translational, rotational,

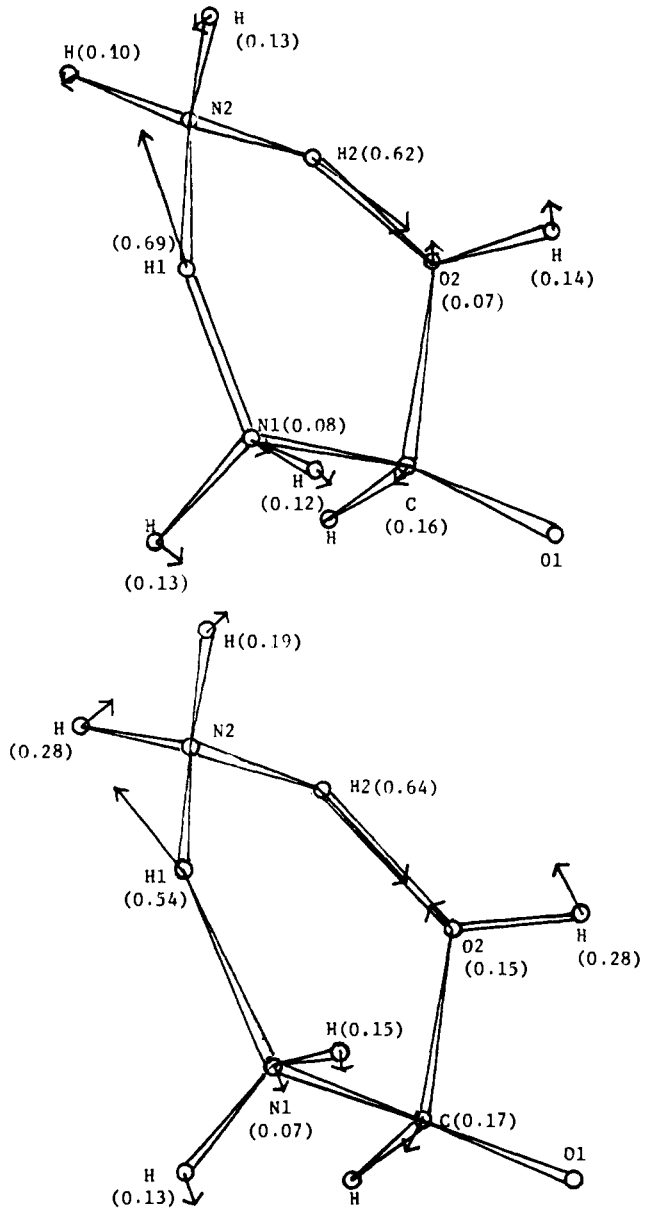

Figure 5. (a, top) Structure of transition state T3 optimized by STO-3G basis set. The reaction coordinate is also given. (b, bottom) Structure of transition state T3 optimized by 3-21G basis set. The reaction coordinate is also given.

and vibrational contributions (i.e., thermal energy correction) to the relative enthalpies. The formulas for these entities were derived from the molecular partition function[14] and used to calculate the relative free energies, as described in the previous work on the uncatalyzed reaction. The relative free energies, entropies, and thermal energy corrections calculated with the STO-3G and 3-21G basis sets (at 298.15 K and 1 atm) are summarized in Table V.

Discussion

Reaction Mechanisms: Nature of Transition States and Intermediates. For both the two-step and concerted mechanisms of amine-catalyzed amide bond formation, the first transition states (see Figures 3 and 5) involve simultaneous nucleophilic attack of one ammonia nitrogen on the carbonyl carbon (i.e., partial C–N bond formation) and the transfer of hydrogen atoms to and from the second ammonia. This catalytic ammonia acts as both a hydrogen atom (H1) acceptor, from the first ammonia leading to partial N–H bond formation, and a hydrogen atom (H2) donor to the carboxylic acid oxygen leading to partial O–H bond formation.

As is seen in Figure 3, a and b, the overall conformational structures of T1 obtained by STO-3G and 3-21G basis sets are similar. Fairly good agreement is seen in the bond and dihedral angles of the six-membered ring in the two optimized geometries. The difference in sign of the dihedral angles of the ring is most likely due to the existing local symmetries and indicates that either direction of the twisting leads to energetically equivalent conformers of the ring. Both optimized structures show partial planarity of the ring around the H1–N2 and H2–N2 bonds, which allows the maximum overlap of the hydrogen atoms (H1 and H2) to be transfered with the lone-pair electrons of the respective acceptor atoms (N2 and O2). The largest difference in the two optimized transition states (T1) is that the one calculated by using a STO-3G basis set resembles the reactants more than that from the 3-21G results, i.e., the partially formed C–N bond is longer (1.785 vs. 1.587 Å) and the partially formed H1–N2 bond is longer than the H1–N1 bond to be broken while the reverse was found in the 3-21G optimized transition state (T1). This difference in extent of progression of the reaction is also reflected in the smaller positive net charge on the second ammonia, i.e., 0.01 (STO-3G) and 0.10 (3-21G) electrons. In both optimized transition-state structures however, a significant amount of electron transfer to the formic acid from both ammonias was observed, i.e., 0.23 (STO-3G) and 0.34 (3-21G). In Figure 3, a and b, the reaction coordinates of the transition states obtained by both the STO-3G and 3-21G basis sets are shown with relative mass-weighted Cartesian displacement vectors of each atom as its components. The Cartesian displacement vectors are obtained by dividing the mass-weighted components by the square roots of the appropriate masses. This reaction coordinate clearly shows the nucleophilic attack of an ammonia on the carboxylic acid and the transfer of hydrogen atoms to and from the second ammonia, where the displacement vectors of H1 and H2 atoms are the predominant components.

The first transition state (T1) of the two-step reaction leads to the release of ammonia and formation of the same intermediate, $H_2N—C(OH)_2H$, formed in the uncatalyzed reaction of formic acid and ammonia previously reported, and a detailed description of this intermediate is given there. In the second step of this reaction, an ammonia molecule can again function as a catalyst by simultaneous donation and acceptance of hydrogen atoms to and from the OH groups of the intermediate leading to the second transition state (T2).

For this second transition state, both STO-3G and 3-21G optimized structures, shown in Figure 4, a and b, respectively, yield a nonplanar six-membered rings with partial planarity around the H1–N2 and H2–N2 bonds. Although excellent agreement in the bond and dihedral angles and most of the bond lengths was obtained for the two optimized geometries, one significant difference is that the partially formed H2–O2 bond of the leaving water molecule is much more developed in the STO-3G case than in the 3-21G case. The reaction coordinate obtained from both basis sets, given in Figure 4, a and b, clearly shows the formation of formamide, water, and ammonia through this transition state.

The transition state (T3) obtained for the concerted mechanism by using both basis sets also has a nonplanar six-membered ring, but it is more planar than the other two transition states. For this transition state, the STO-3G and 3-21G optimized structures indicate comparable timing of transfer of both hydrogens, H1 and H2, to and from the catalytic ammonia: the forming H1–N2 bond is shorter than the breaking H1–N1 bond, and the forming H2–O2 bond is longer than the breaking H2–N2 bond. A significant amount of electron transfer to formic acid from the two ammonia molecules is also seen in the transition state obtained by both basis sets, i.e., 0.32 (STO-3G) and 0.40 (3-21G). In this transition state, the amount of net positive charges remaining on both the nucleophile and catalyst ammonia molecules are similar (Table II), while for transition state T1, the positive charge remained predominantly on the nucleophile ammonia. The reaction coordinate obtained by both basis sets given in Figure 5, a and b, display the

(14) Hertzberg, G. "Molecular Spectra and Molecular Structure: II. Infrared and Raman Spectra of Polyatomic Molecules"; Van Nostrand Company: Toronto, 1968; Chapter 5.

Table V. Ab Initio Calculations on Relative Energies[a]

method[b]	T1	T2	INT	T3	R
	$\Delta G_1^{\ddagger}(\Delta G_1^{\ddagger})$	$\Delta U_2^{\ddagger}(\Delta G_2^{\ddagger})$	$\Delta U_i(\Delta G_i)$	$\Delta U_3^{\ddagger}(\Delta G_3^{\ddagger})$	$\Delta U_r(\Delta G_r)$
HF/STO-3G//A	12.6 (37.3)	43.9 (54.6)	−25.6 (−12.5)	31.7 (57.0)	9.8 (9.4)
HF/3-21G//B	3.6 (29.6)	11.9 (22.6)	−5.7 (7.4)	16.5 (41.8)	1.0 (0.6)
MP2/3-21G//B	−2.5 (22.6)	−3.8 (6.9)	4.6 (17.7)	0.9 (26.2)	3.7 (3.3)
HF/6-31G//B	19.4 (44.5)	22.0 (32.7)	−3.5 (9.6)	30.9 (56.2)	−6.0 (−6.4)
MP2/6-31G//B	10.5 (35.6)	6.8 (17.5)	4.4 (17.5)	18.4 (43.7)	−3.9 (−4.3)
HF/6-31G*//B	36.4 (61.5)	34.6 (45.3)	6.0 (19.1)	49.2 (74.5)	1.8 (1.4)
MP2/6-31G*//B	14.5 (39.6)	15.0 (25.7)	5.0 (18.1)	25.1 (50.4)	1.1 (0.7)
HF/6-31G**//B	30.8 (55.9)	34.4 (45.1)	4.6 (17.7)	48.5 (73.8)	0.1 (−0.3)
MP2/6-31G**//B	13.8 (38.9)	15.1 (25.8)	3.6 (16.7)	25.0 (50.3)	0.0 (−0.4)
HF/3-21G//A	0.1 (25.2)	11.3 (22.0)	−3.7 (9.4)	17.9 (43.0)	5.4 (5.0)
HF/6-31G//A	13.6 (38.7)	21.3 (32.0)	−0.6 (12.5)	33.4 (58.7)	1.2 (−1.6)
	$T\Delta S_1^{\ddagger}$	$T\Delta S_2^{\ddagger}$	$T\Delta S_i$	$T\Delta S_3^{\ddagger}$	$T\Delta S_r$
HF/STO-3G//A	−22.6	−12.2	−10.6	−23.1	0.1
HF/3-21G//B	−22.2	−11.3	−10.7	−22.2	0.0
	$\Delta V_1^{\ddagger}$	$\Delta V_2^{\ddagger}$	ΔV_i	$\Delta V_3^{\ddagger}$	ΔV_r
HF/STO-31G//A	3.4	0.3	2.9	2.6	−1.0
HF/3-21G//B	2.9	−0.6	2.4	3.1	−0.4

[a] All relative energies, ΔU, ΔG, $T\Delta S$, and ΔV are in kcal/mol. G (free energy), S (entropy), and V (sum of translational, rotational, and vibrational enthalpies) are calculated at 298.15 K and 1 atm. U is the Hartree–Fock (HF) or perturbative (MP) energies. All the energies are relative to the energy of reactant, except that the energy at T2 is relative to that at INT, which is the sum of the energies of the intermediate and NH_3. The energies of the product relative to those of the reactant are given in the last column. Note that $G = U - TS + V$. [b] The method and basis sets used for geometry optimization and single point energy calculation are given at the right and left sides of //, respectively, where A = HF/STO-3G and B = HF/3-21G.

Table VI. Comparison of Amine-Catalyzed and Uncatalyzed NH_3–HCOOH Reactions[a]

method[b]	system	T1, $\Delta U_1^{\ddagger}(\Delta G_1^{\ddagger})$	T2, $\Delta U_2^{\ddagger}(\Delta G_2^{\ddagger})$	INT, $\Delta U_i(\Delta G_i)$	T3, $\Delta U_3^{\ddagger}(\Delta G_3^{\ddagger})$	R, $\Delta U_r(\Delta G_r)$
M1	uncat	40.7 (51.8)	44.4 (41.5)	−3.7 (9.4)	43.2 (54.4)	5.4 (5.0)
	cat.	0.1 (25.2)	11.3 (22.0)	−3.7 (9.4)	17.9 (43.0)	5.4 (5.0)
M2	uncat	42.2 (53.3)	47.8 (44.9)	−0.6 (12.5)	54.6 (65.2)	−1.2 (−1.6)
	cat.	13.6 (38.7)	21.3 (32.0)	−0.6 (12.5)	33.4 (58.7)	−1.2 (−1.6)
M3	uncat	38.9 (50.0)	38.5 (35.6)	3.6 (16.7)	42.0 (52.6)	0.0 (−0.4)
	cat.	13.8 (38.9)	15.1 (25.8)	3.6 (16.7)	25.0 (50.3)	0.0 (−0.4)

[a] All relative energies, ΔU and ΔG, in kcal/mol. G (free energy) is calculated at 298.15 K and 1 atm. U is the Hartree–Fock (HF) or perturbative (MP) energy. All energies are relative to the energy of reactant, except that the energies at T2 are relative to those at INT. The energies of the product relative to those of the reactant are given in the last column. Note that the relative energies of INT and R are the same as those for amine-catalyzed reaction studied here. See ref 3 for more details. [b] M1 = HF/3-21G//HF/STO-3G, M2 = HF/6-31G//HF/STO-3G, and M3 = MP2/6-31G**//HF/3-21G. See Table V for the notation. [c] cat. and uncat refer to amine-catalyzed and uncatalyzed reactions, respectively.

separation of the three molecules about to occur leading to direct product formation.

All three transition states are much more polar than the intermediate. The calculated dipole moments from use of using 6-31G** basis set with 3-21G optimized geometries are 4.72, 4.19, 8.47, and 0.87 D for the transition states T1, T2 and T3 and the intermediate, respectively. Despite the close structural similarity between the transition states T1 and T3, the larger dipole moment observed with T3 is understandable considering the total number of bonds cleaved and formed, i.e., five for T1 and six for T3.

Energetics of Stationary Points in the Reaction Pathway. Table V summarizes the Hartree–Fock (HF) and perturbative energies (ΔU) and free energies (ΔG) calculated for all the stationary points of both reaction pathways at all levels of calculations. All the energies given in Tables III–VI are relative to the energy of the reactant (i.e., formic acid and two ammonia molecules), except that the energies of the second transition state (T2) are given relative to the sum of the energies of the intermediate and NH_3. Throughout the following discussion all energy comparisons refer to free energy, unless otherwise noted.

Experimental values of $\Delta G_r = -0.3$ kcal/mol and $T\Delta S_r = -0.4$ kcal/mol were calculated from the experimental heats of formation and entropies of the reactant molecules in the gas phase.[15] These results are in excellent agreement with the results of our highest level calculations, i.e., $\Delta G_r = -0.4$ kcal/mol (MP2/6-31G**) and $T\Delta S_r = 0.0$ kcal/mol.

For ΔG values listed in Table V, entropies (ΔS) and thermal energy corrections (ΔV) calculated with 3-21G basis set were used. As shown in Table V, ΔS and ΔV values calculated with 3-21G basis set are in fairly satisfactory agreement with those calculated with STO-3G basis set.

This agreement indicates that these quantities are insensitive to differences in geometries of transition states, reactants, intermediates, and products obtained by the two different basis sets. The effect of those geometry differences on $\Delta U^{\ddagger}$, ΔU_i and ΔU_r can also be seen in Table V. Comparing results using STO-3G optimized geometries (HF/STO-3G//A) with those using 3-21G optimized geometries (HF/3-21G//B), we see that $\Delta U_1^{\ddagger}$, $\Delta U_2^{\ddagger}$, $\Delta U_3^{\ddagger}$, and ΔU_r are lowered by 9.0, 32.0, 15.2, and 8.8 kcal/mol, respectively, while ΔU_i is increased by 19.9 kcal/mol. The interesting question arises as to how much of this correction one obtained by a single-point energy calculation using the 3-21G basis set with the STO-3G optimized geometries. These results are also shown in Table V (HF/3-21G//A). Comparing HF/3-21G//A with HF/3-21G//B, we see that $\Delta U_1^{\ddagger}$, $\Delta U_2^{\ddagger}$, $\Delta U_3^{\ddagger}$, and ΔU_i agree within 3.5, 0.6, 1.4, and 2.0 kcal/mol, respectively, while value of ΔU_r differs by 4.4 kcal/mol. While incomplete, this comparison does support at least the qualitative validity of investigating the effect of polarization function and electron correlation in this system by single-point calculations.

In Table III, the effect of correlation on the stationary point energies is given as the difference between the perturbative and

(15) (a) Cox, J. D.; Pilcher, G. "Thermochemistry of Organic and Organometallic Compounds"; Academic Press: New York, 1970. (b) Stull, D. R.; Prophet, H. *Natl. Stand. Ref. Data Ser.* **1971**, No. 37. (c) Wagman, D. D.; Evans, W. H.; Parker, V. B.; Halow, I.; Baily, S. M.; Schumm, R. H. *Natl. Bur. Stand. Tech. Note* **1968**, No. 270-3.

Hartree–Fock energies at various basis set levels. For calculations that include polarization, significantly larger correlation effects are observed that tend to lower the relative energies for all three transition states, T1, T2, and T3, while only small effects are seen for the stable reaction species.

Despite the close structural similarity between the transition states T1 and T3, the larger correlation effect observed with T3 is understandable considering the total number of bonds formed and cleaved, i.e., five for T1 and six for T3.

It is interesting to note that even for 3-21G and 6-31G basis sets, the directions of correlation effects on the activation energies are the same as those observed for larger basis sets, though the correlation energies for T1 are fairly underestimated compared with those for T2 and T3 with the smaller basis sets.

Table IV summarizes the effects of the basis set size and polarization on the stationary-point energies. Contrary to correlation effects, which lower the activation energy at every level, increasing basis set size from 3-21G to 6-31G and adding d-polarization functions to 6-31G basis set increase the relative energies for both the transition states and intermediate. However, a decrease in activation energies (a maximum of 5.6 kcal/mol) is obtained for all three transition states by addition of p-polarization functions to the hydrogen atoms in the 6-31G* basis set. This decrease is most likely due to the increased stabilization of multiple bonds formed by the H1 and H2 atoms in the six-membered-ring transition states.

For all transition states, correlation and polarization effects are in opposite directions, but both affect each transition state in the same order, T1 < T2 < T3, at the MP2/6-31G** level. Thus, by cancellation, the net effect on these activation energies is very similar, i.e., −5.6, −6.9, and −5.9 kcal/mol for T1, T2, and T3, respectively. In consequence of the opposite effects of correlation and polarization, HF/6-31G results for the transition states are very similar to the highest level MP2/6-31G** calculations we have performed, as shown in Table V.

The higher level of calculation, which include both correlation and polarization effects, predicts an intermediate less stable than the reactant (i.e., $\Delta G_i > 0$) and the order of activation energies $\Delta G_2^{\ddagger} < \Delta G_1^{\ddagger} < \Delta G_3^{\ddagger}$. All of these results are qualitatively the same as those obtained by single-point Hartree–Fock calculations at the 3-21G or 6-31G level.

For comparison, the activation energies obtained for the previous work on uncatalyzed amide bond formation between NH_3 and HCOOH are included in Table VI along with selected results obtained here. As is seen in Table VI, use of STO-3G optimized geometries is a fairly satisfactory approximation for a higher level of calculation: the order of activation energies obtained at HF/3-21G//HF/STO-3G and HF/6-31G//HF/STO-3G is the same as that obtained at highest level of theory. Thus, for larger systems currently under consideration such as amide bond formation between ammonia and glycine molecules or peptide bond formation between two glycine molecules, the reasonable level of calculation chosen was HF/3-21G//HF/STO-3G or HF/6-31G//HF/STO-3G.

A comparison of the free energies of activation, $\Delta G_1^{\ddagger}$ and $\Delta G_2^{\ddagger}$, for the two-step reaction mechanism clearly shows that the first step (i.e., nucleophilic attack by nitrogen atom on the carbonyl carbon atom) is slower than the second step (i.e., breakdown of the intermediate). Further, since the potential energies $\Delta U_1^{\ddagger}$ and $\Delta U_2^{\ddagger}$ (at MP2/6-31G**) are nearly equal, the entropy contribution to the free energy (i.e., $T\Delta S$ of −22.2 kcal/mol to $\Delta G_1^{\ddagger}$ and −11.3 kcal/mol to $\Delta G_2^{\ddagger}$) is the most critical factor in determining which step has the larger activation energy in this mechanism. The overall rate of the reaction is determined by the free-energy difference between the reactant and the highest point of the energy surface. Relative to the reactant, the highest point of the energy surface is the energy of the second transition state ($\Delta G_2^{\ddagger} + \Delta G_i$), which is greater than $\Delta G_1^{\ddagger}$ at the four highest levels of calculation. However, the small differences between these two quantitites, i.e., 42.5 vs. 38.9 kcal/mol at the MP2/6-31G** level, implies that both steps nearly equally affect the overall rate of the reaction.

Comparing the two possible mechanisms of amide bond formation, the free energies of activation of both steps in the two-step reaction are calculated to be significantly lower than those for the concerted one ($\Delta G_3^{\ddagger}$) at all levels of calculations. At the highest level of theory, MP2/6-31G**, $\Delta G_2^{\ddagger} + \Delta G_i$ and $\Delta G_3^{\ddagger}$ are calculated to be 42.5 and 50.3 kcal/mol, respectively. Thus, for amine-catalyzed amide bond formation we predict that in a hydrophobic environment, the two-step reaction is more favorable than the concerted one. At this same level of theory ΔG_1 and $\Delta G_2^{\ddagger} + \Delta G_i$ for the two-step reaction are calculated to be 38.9 and 42.5 kcal/mol, respectively, indicating that both steps affect the overall rate of reaction but not equally, with the second step dominating.

For the three transition states, a comparison of amine-catalyzed and uncatalyzed reactions (see Table VI) clearly shows that the substantial decreases in the potential energies of activation ($\Delta U^{\ddagger}$), i.e., 25.1, 23.4, and 17.0 kcal/mol (at MP2/6-31G** level) for $\Delta U_1^{\ddagger}$, $\Delta U_2^{\ddagger}$, and $\Delta U_3^{\ddagger}$, respectively, caused by the catalytic ammonia are significantly offset by the increase in the entropy (about 10 kcal/mol in $T\Delta S$). The smallest decrease in potential energy of activation for transition state T3 is associated with the greatest increase in the polarity. Dipole moments for the transition states T1, T2, and T3 increase by 0.72, 0.62, and 4.49 D, respectively, in going from the uncatalyzed reaction to the catalyzed reaction.

Conclusions

Several conclusions can be drawn from this study. These conclusions are based on the assumption that the relative energetics of the concerted and two-step reactions and of the catalyzed and uncatalyzed reactions would not be qualitatively altered by changes in geometries obtained at higher than the 3-21G level of calculation.

(1) Two-step and concerted reaction mechanisms of amine-catalyzed amide bond formation have been characterized that involve simultaneous nucleophilic attack of the amine on the carbonyl carbon atom and indirect transfer of a hydrogen atom from the nucleophilic amine to a carboxylic acid oxygen atom via the second amine, which acts as both an acid and base. In both gas-phase and hydrophobic environments, the two-step reaction mechanism is predicted to be more favorable than the concerted one. For the two-step mechanism the first step, reactant to intermediate, is slower than the second step, intermediate to product, due mainly to the entropy contribution to the free energy of activation. However, the activation energy of the second step relative to the reactant is predicted to be slightly higher than that of the first step, indicating the overall rate of amide formation by this mechanism is determined by both steps. The intermediate state is found to be less stable than the reactant.

(2) The amine catalyst substantially decreases the potential energies of all three transition states but increases the entropy changes, resulting in a net decrease of free energy of activation by about 10 kcal/mol for the two-step reaction and by about 2 kcal/mol for the concerted one from those in the uncatalyzed reaction. Entropy contributions to the free energies of activation are much greater than thermal energy corrections for all three transition states.

(3) Correlation effects and p-polarization functions lower the activation energies which are raised by d-polarization functions. The net effect of including both correlation and polarization is to lower the activation energies by 5.6–6.9 kcal/mol with nearly equal lowering for each of the three transition states of interest. Therefore, the relative order of free energies of activation predicted at the highest level of theory used here is the same as that predicted by a HF/3-21G or HF/6-31G level of calculation using a STO-3G optimized geometry, i.e., $\Delta G_2^{\ddagger} < \Delta G_1^{\ddagger} < \Delta G_3^{\ddagger}$.

Acknowledgment. Support from NASA Consortium Agreement NCA2-OR630-001 with The Rockefeller University is acknowledged. We are indebted to Drs. D. Spangler, D. DeFrees, and S. Binkley for helpful discussions.

Registry No. Formic acid, 64-18-6; ammonia, 7664-41-7; formamide, 75-12-7.