J. Am. Chem. Soc. **1996,** *118,* 11237−11243

What Happens to Salt-Bridges in Nonaqueous Environments: Insights from Quantum Mechanics Calculations

Ya-Jun Zheng[†] **and Rick L. Ornstein***

Contribution from the Environmental Molecular Sciences Laboratory,
Pacific Northwest National Laboratory, Richland, Washington 99352

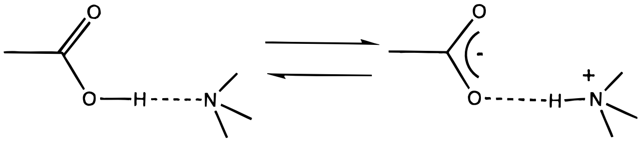
Received January 4, 1996[⊗]

Abstract: To examine the effect of solvent environment on protein salt-bridges, we performed high-level *ab initio* molecular orbital calculations in the gas phase and in three different solvents on a salt-bridge as modeled by formate and guanidinium ions. The energy difference between the neutral hydrogen-bonded complex and the zwitterionic form and the interconversion barrier between them are investigated in detail at RHF/6-31G*, RHF/6-311+G**, MP2/6-31G*, and MP2/6-311+G** levels. In the gas phase, the neutral conventional hydrogen-bonded complex is predicted to be favored at all four levels of theory and there is a small barrier for the interconversion. In a nonpolar, hydrophobic solvent like CCl_4, the energy difference between these two forms is small and the barrier that separates them is also low, but the neutral hydrogen-bonded complex still seems to be slightly favored. However, in polar solvents like DMSO and water, the zwitterionic form becomes much more favored. In polar solvents, the barrier for conversion of the neutral hydrogen-bonded form to the zwitterionic form is small at the Hartree−Fock level, but it disappears at the correlated level (MP2). The implication of these findings toward stabilizing an enzyme in nonaqueous solvents is briefly discussed.

Introduction

Recently there has been increasing interest in nonaqueous enzymology.[1] Many advantages of carrying out enzymatic reactions in nonaqueous environments have been identified experimentally. To fine tune enzyme activity and stability in organic solvents, several "design" rules were suggested, such as introducing internal cross-links, maximizing intra-protein hydrogen bonds, and removing surface charges and surface hydrogen-bonding sites.[2] Most of these are yet to be tested, but the effect of removal of surface charges has been investigated using site-directed mutagenesis on α-lytic protease[3a] and subtilisin 8397.[3b] Although removal of surface charges in α-lytic protease[3a] and in two mutants of subtilisin 8397 (Lys43→Asn (K43N) and Lys256→Tyr (K256Y)) did improve the stability of subtilisin in high concentrations of hydrophilic organic solvents, the D181N (Asp181→Asn) mutant of subtilisin 8397 was actually less stable than subtilisin 8397.[3b] Clearly, to obtain a fundamental understanding of what causes these changes in enzyme stability and function in organic solvents and to predict enzyme function in new solvents, structural and thermodynamic information concerning enzymes in nonaqueous solvent environments are a prerequisite.

Scheme 1



Recently, both experimental (X-ray crystallographic and solution NMR)[4−7] and theoretical (molecular dynamics simulation)[8−11] techniques have been used to probe enzyme structure and dynamics in organic solvents. Crystallographic studies show that protein structures in organic solvents are very similar to their aqueous structures except for some differences in solvent exposed side chain conformation. NMR studies are controversial and so far no three-dimensional solution structure has been solved in organic solvent. Molecular dynamics simulations have demonstrated that additional hydrogen bonds and salt-bridges are formed when proteins are put into organic solvents. The increase in total number of intra-protein hydrogen bonds is probably responsible for some of the unusual properties of enzymes in nonaqueous environments such as increased thermostability.

In principle, a hydrogen-bonded complex between Asp (or Glu) and Lys (or Arg) can be either a neutral conventional hydrogen-bonded complex or a charged zwitterionic form (Scheme 1), depending on the environment. For example, in the gas phase, glycine exists in a neutral form, but in the solid

---

* Author to whom correspondence should be addressed: (509) 375-6904, Fax; rl_ornstein@pnl.gov, Internet.

[†] Present address: Department of Chemistry, University of California, Santa Barbara, CA 93106.

[⊗] Abstract published in *Advance ACS Abstracts*, November 1, 1996.

(1) (a) Klibanov, A. M. *Trends Biochem. Sci.* **1989**, *14*, 141. (b) Chen, C.-S.; Sih, C. J. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 695. (c) Klibanov, A. M. *Acc. Chem. Res.* **1990**, *23*, 114. (d) Dordick, J. S. *Enzyme Microb. Technol.* **1989**, *11*, 194. (e) Gupta, M. N. *Eur. J. Biochem.* **1992**, *203*, 25. (f) Faber, K.; Riva, S. *Synthesis* **1992**, 895. (g) Wescott, C. R.; Klibanov, A. M. *Biochim. Biophys. Acta* **1994**, *1206*, 1. (h) Halling, P. J. *Enzyme Microb. Technol.* **1994**, *16*, 178.

(2) Hwang, J. Y.; Arnold, F. H. In *Applied Biocatalysis*; Blanch, H. W., Clark, D. S., Eds.; M. Dekker: New York, 1991; pp 53−86.

(3) (a) Martinez, P.; Arnold, F. H. *J. Am. Chem. Soc.* **1991**, *113*, 6336. (b) Sears, P.; Schuster, M.; Wang, P.; Witte, K.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 6521.

(4) Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653. Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675.

(5) Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326. Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *J. Am. Chem. Soc.* **1995**, *117*, 577.

(6) (a) Desai, U. R.; Osterhout, J. J.; Klibanov, A. M. *J. Am. Chem. Soc.* **1994**, *116*, 9240. (b) Desai, U. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1995**, *117*, 3940.

(7) Wu, J.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 6843.

(8) (a) Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 10113. Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 6529. Toba, S.; Hartsough, D. S.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1996**, *118*, 6490. (b) Norin, M.; Edholm, O.; Haeffner, F.; Hult, K. *Biophys. J.* **1994**, *67*, 548.

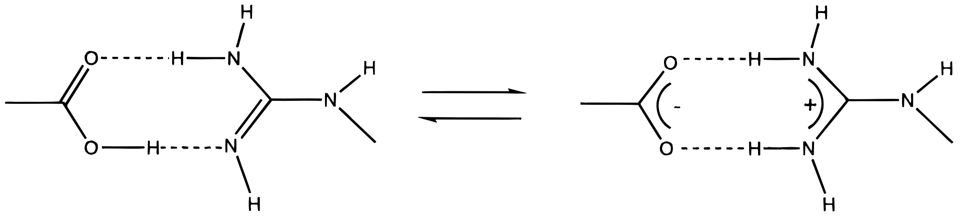
(9) Zheng, Y.-J.; Ornstein, R. L. *Biopolymers* **1996**, *38*, 791.

(10) Zheng, Y.-J.; Ornstein, R. L. *Protein Eng.* **1996**, *9*, 485.

(11) Zheng, Y.-J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 4175.

S0002-7863(96)00041-8 CCC: $12.00

Scheme 2



Neutral H-Bond Complex **Zwitterionic H-Bond Complex**

phase and aqueous solution it exits as the zwitterionic form.[12] The interconversion between these two forms of the hydrogen-bonded complex can be easily accomplished by proton tunneling, which also provides a simple mechanism for charge annihilations in nonaqueous environments. Which form the "salt-bridge" takes could have great implications in enzyme catalysis, in the structure of membrane proteins, and in the stability and function of enzymes in nonaqueous environments; it could also be a potential problem for modeling enzymes in different solvent environments. Although salt-bridges in proteins have been investigated intensively,[13−27] they are routinely assumed to be in the zwitterionic form. The energy difference between the neutral and zwitterionic form was considered in the work by Warshel and co-workers[26] and Honig and Hubbell.[27] The interconversion between the two forms of a salt-bridge involves a proton transfer process; proton transfer in aqueous solution and in enzymatic reactions has been investigated extensively by Warshel and co-workers using the empirical valence bond method.[28−31] Honig and Hubbell[27] also examined the stability of salt-bridges in membrane proteins based on electrostatic considerations; they concluded that the dielectric constant of the medium would have to be less than about 4 before the neutral form becomes favored. However, this latter approach has been criticized for assuming unit formal charges on oxygen and nitrogen atoms.[32] Furthermore, such an electrostatic approach is unable to give the barrier for this interconversion. An alternative approach is to use *ab initio* quantum mechanics in combination with reaction field theory.[33−35] Here we report an investigation using *ab initio* quantum mechanics in combination with reaction field theory to study the interaction between the carboxylic group and the guanidinium group of an arginine as modeled by formate and guanidinium ion (Scheme 2).

Theoretical Methods

All *ab initio* molecular orbital calculations were carried out using Gaussian 92.[36] Geometries of the species involved were fully optimized at the RHF/6-31G*, RHF/6-311+G**, and MP2/6-31G* levels. First, we performed calculations in the gas phase and the potential energy surface was searched in detail. The transition state for the interconversion between the neutral hydrogen-bonded complex and the zwitterionic hydrogen-bonded complex was located at all three levels. Each stationary point was then characterized by harmonic frequency analysis at both RHF/6-31G* and RHF/6-311+G** levels. For calculations in solutions, the Onsager solvent reaction field model was used,[37] which has been successfully implemented in Gaussian 92 and applied to sulfamic acid (in both neutral form (H_2NSO_2OH) and zwitterionic form ($^+H_3NSO_3^-$)), formamide, and 2-pyridone.[37−41] Three solvents were considered and the dielectric constants are 78.3, 46.45, and 2.238 for water, dimethyl sulfoxide (DMSO), and carbon tetrachloride (CCl_4), respectively.[42] The compounds we used in this study are shown in Chart 1. The neutral hydrogen-bonded complex, the zwitterionic form, and the transition state for the interconversion between them are designated as **1**, **3**, and **2** in the gas phase; in solutions, they are designated as **1a**, **3a**, and **2a** in CCl_4, as **1b**, **3b**, and **2b** in DMSO, and **1c**, **3c**, and **2c** in water. Since geometry optimization using the self-consistent reaction field (SCRF) model at the MP2 level is not available, only single point calculations were performed at MP2/6-31G* and MP2/

---

(12) Zheng, Y.; Neville, J. J.; Brion, C. E. *Science* **1995**, *270*, 786 and the references cited therein.

(13) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1983**, *168*, 867.

(14) Perutz, M.; Raidt, H. *Nature* **1975**, *255*, 256. Perutz, M. *Science* **1978**, *201*, 1187.

(15) Ruegg, C.; Ammer, D.; Lerch, K. *J. Biol. Chem.* **1982**, *257*, 6420.

(16) Fersht, A. R. *J. Mol. Biol.* **1972**, *64*, 497.

(17) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. *Biochemistry* **1990**, *29*, 2403.

(18) Horovitz, A.; Serrano, L.; Avron, B.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1990**, *216*, 1031.

(19) Erwin, C. R.; Barnett, B. L.; Ovliver, J. D.; Sullivan, J. F. *Protein Eng.* **1990**, *4*, 87.

(20) Sali, D.; Bycroft, M.; Fersht, A. R. *J. Mol. Biol.* **1991**, *220*, 779.

(21) Dao-Pin, S.; Sauer, U.; Nicholson, H.; Matthews, B. W. *Biochemistry* **1991**, *30*, 7142.

(22) Lumb, K. J.; Kim, P. S. *Science* **1995**, *268*, 436. Lavigne, P.; Sonnichsen, F. D.; Kay, C. M.; Hodges, R. S. *Science* **1996**, *271*, 1136. Lumb, K. J.; Kim, P. S. *Science* **1996**, *271*, 1137.

(23) Waldburger, C. D.; Schildbach, J. F.; Sauer, R. T. *Nature Struct. Biol.* **1995**, *2*, 122.

(24) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211.

(25) Hwang, J.-K.; Warshel, A. *Nature* **1988**, *334*, 270.

(26) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283. Warshel, A.; Aqvist, J. *Annu. Rev. Biophys. Chem.* **1991**, *20*, 267 and the references cited therein.

(27) Honig, B. H.; Hubbell, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5412.

(28) (a) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640. (b) Warshel, A. *Biochemistry* **1981**, *20*, 3167.

(29) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629. Warshel, A.; Russell, S. *J. Am. Chem. Soc.* **1986**, *108*, 6569.

(30) Warshel, A. In *Computer Modeling of Chemical Reactions in Enzymes and in Solutions*; Wiley: New York, 1991.

(31) Aqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523.

(32) Kyte, J. In *Structure in Protein Chemistry*; Garland Publishing, Inc.: New York, 1995; pp 165−166.

(33) Rinaldi, D.; Rivail, J.-L. *Theor. Chim. Acta* **1973**, *32*, 57. Yomosa, S. *J. Phys. Soc. Jpn.* **1973**, *35*, 1738. Tapia, O.; Goscinski, O. *Mol. Phys.* **1975**, *29*, 1653.

(34) Miertus, S.; Scrocco, E.; Tomasi, *J. Chem. Phys.* **1981**, *55*, 117. Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Valle, F. O. *J. Mol. Struct.* **1991**, *234*, 401. Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(35) Rinaldi, D.; Ruiz-Lopez, M. F.; Rivail, J.-L. *J. Chem. Phys.* **1983**, *78*, 834.

(36) Gaussian 92; Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Steward, J. J. P.; Pople, J. A.; GAUSSIAN, Inc.: Pittsburgh, PA, 1992.

(37) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.

(38) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776.

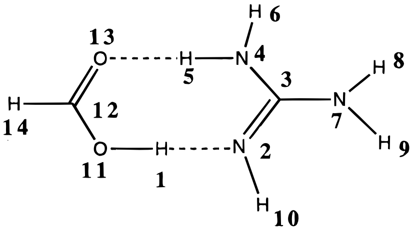
(39) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Am. Chem. Soc.* **1992**, *114*, 523.

(40) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Chem. Phys.* **1991**, *95*, 8991.

(41) Wong, M. W.; Wiberg, K. B.; Frisch, M. J. *J. Am. Chem. Soc.* **1992**, *114*, 1645.

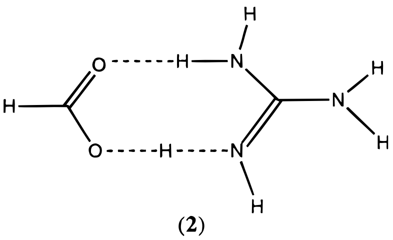
(42) *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, 1994.

Chart 1



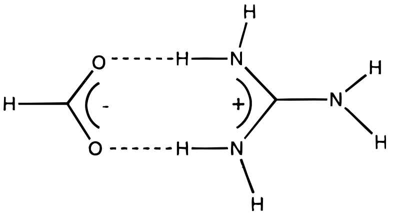
(**1**)

Neutral H-bonded complex



(**2**)

Transition State



(**3**)

Zwitterionic form

6-311+G** levels of theory. For single-point SCRF calculations at the MP2/6-31G* level, the gas-phase MP2/6-31G* geometries were used; for the single-point SCRF calculations at the MP2/6-311+G** level, the corresponding geometries in each solvent at the RHF/6-311+G** level were used. It has been demonstrated by Wiberg and co-workers that the reaction field model using the MP2/6-311+G** level of theory is able to reproduce the experimental solvent effects.[43] Nevertheless, we also tried the MP2/6-311++G** level of theory and indeed it gave the same results as the MP2/6-311+G**, which is consistent with previous work by Wiberg and co-workers.[43]

For the SCRF calculations, the radius of the molecule is needed. This is straightforward for **1**, **2**, and **3**. It has been shown that both **1** and **3** are minima on the potential energy hypersurface at the RHF/6-31G* level in the gas phase.[44] If we want to examine the interaction between the Lys-Asp (or Glu) type of salt-bridges, we would similarly choose formate and methylammonium ions as models; since it has been demonstrated (at the RHF/6-31G* level) that a minimum does not exist on the potential energy hypersurface for the zwitterionic form ($HCOO^-$- - -$^+HNH_2CH_3$),[44] no additional calculations were here pursued.

Results and Discussions

The calculated total electronic energy for each species involved is given in Table 1 and the calculated geometrical parameters are summarized in Tables 2, 3, 4, and 5 for species **1**−**3**, **1a**−**3a**, **1b**−**3b**, and **1c**−**3c**, respectively. For simplicity, the results in the gas phase and different solutions will first be discussed individually.

Table 1. The Calculated Total Electronic Energy (au) for All of the Species Involved

compd[a]	6-31G*	6-311+G**	MP2/6-31G*	MP2/6-311+G**
1	−392.9083333	−393.0425273	−394.0048819	−394.2576221
2	−392.8958128	−393.0305763	−393.9986453	−394.2520633
3	−392.9000250	−393.0338204	−393.9993708	−394.2499742
1a	−392.9097485	−393.0439357	−394.0064194	−394.2591592
2a	−392.9011394	−393.0359810	−394.0039693	−394.2570900
3a	−392.9099480	−393.0436070	−394.0066935	−394.2566755
1b	−392.9119367	−393.0461642	−394.0085188	−394.2616996
2b	−392.9078537	−393.0429110	−394.0112181	−394.2631152
3b	−392.9260870	−393.0605131	−394.0166463	−394.2714496
1c	−392.9120027	−393.0462325	−394.0085759	−394.2617822
2c	−392.9080312	−393.0430956	−394.0114180	−394.2632688
3c	−392.9265873	−393.0610670	−394.0169205	−394.2719452

[a] The neutral hydrogen-bonded complex, the zwitterionic form, and the transition state for the interconversion between them are designated as **1**, **3**, and **2** in gas phase; in solutions, they are designated as **1a**, **3a**, and **2a** in CCl_4, as **1b**, **3b**, and **2b** in DMSO, and **1c**, **3c**, and **2c** in water.

Gas Phase. Geometries for **1** and **3** were fully optimized and the interconversion transition state between them was located at RHF/6-31G*, RHF/6-311+G**, and MP2/6-31G* levels. Table 2 summarizes the calculated geometrical parameters for compounds **1**−**3** at each level of theory. Table 6 gives the calculated energy difference between **1** and **3** and the interconversion barrier between them. At the MP2/6-311+G** level of theory, only a single-point calculation was carried out using the RHF/6-311+G** optimized geometry. It is clear from Table 6 that the neutral hydrogen-bonded complex (**1**) is favored at all levels of theory. As the level of theory increases, the barrier between **1** and **3** decreases. The reverse barrier disappears at the correlated level (MP2) and the neutral complex seems to be the only minimum on the potential energy surface. Calculations at the MP2/6-311++G** level give the same results as those at the MP2/6-311+G** level of theory, indicating that the addition of diffuse functions on hydrogens has little effect on the calculated results.

In general, both RHF/6-31G* and RHF/6-311+G** give very similar geometries; the bond lengths and angles differ by less than 0.02 Å and 2°, respectively. There are two N- - -O distances for each compound. In **1** and **2**, these two N- - -O distances are different, while in **3**, they are the same. The N- - -O distances are longer in **1** than in **3** at each level of theory. The calculated energy difference and interconversion barrier are also very similar at both RHF/6-31G* and RHF/6-311+G** levels. At the correlated level (MP2), the hydrogen-bonding distances become shorter. This is expected since it has been shown that hydrogen bonding is stronger at the MP2/6-31G* level than at the RHF/6-31G* level.[44] It is also interesting to note that compared to the RHF/6-31G* and RHF/6-311+G** level results, the transition state at the MP2/6-31G* level becomes more like the zwitterionic form. This is consistent with the fact that the reverse barrier becomes smaller as the level of theory increases. The calculated geometries for **1**−**3** at the MP2/6-31G* level of theory are shown in Figure 1.

In CCl_4. The calculated results for **1a**−**3a** in CCl_4 are summarized in Tables 3 and 6. Calculations at RHF/6-31G*, RHF/6-311+G**, and MP2/6-31G* levels predict that compounds **1a** and **3a** are similar in energy; calculations at the MP2/6-311+G** level predict compound **1a** to be favored by about 1 kcal/mol. Again the interconversion barrier between them becomes smaller as the level of theory increases. The reverse barrier is only about 0.3 kcal/mol at the MP2/6-311+G** level; thus the potential energy surface is very flat. The trend for the geometry is also similar to that in the gas phase. The calculated geometries at RHF/6-31G* and RHF/6-311+G** levels are very similar (see Table 3). The differences in calculated bond length

---

(43) Wiberg, K. B.; Keith, T. A.; Frisch, M. J.; Murcko, M. *J. Phys. Chem.* **1995**, *99*, 9072.

(44) Zheng, Y.-J.; Merz, K. M., Jr. *J. Comput. Chem.* **1992**, *13*, 1151.

Table 2. The Calculated Geometrical Parameters (Bond Length in Å and Angle in deg) for Compounds **1**−**3** in the Gas Phase

	1[a]			**2**[a]			**3**[a]		
parameters	6-31G*	6-311+G**	MP2/6-31G*	6-31G*	6-311+G**	MP2/6-31G*	6-31G*	6-311+G**	MP2/6-31G*
r_{1-2}	1.8043	1.7960	1.6878	1.2127	1.2033	1.1995	1.0412	1.0401	1.0859
r_{1-11}	0.9814	0.9761	1.0261	1.2576	1.2615	1.3133	1.6415	1.6299	1.5429
r_{11-2}	2.7852	2.7716	2.7122	2.4690	2.4637	2.5112	2.6822	2.6691	2.6288
r_{4-5}	1.0023	1.0004	1.0227	1.0155	1.0136	1.0444	1.0407	1.0399	1.0715
r_{13-5}	2.0599	2.0799	1.9580	1.8140	1.8167	1.7244	1.6439	1.6303	1.5942
r_{13-4}	3.0521	3.0661	2.9746	2.8191	2.8177	2.7650	2.6841	2.6694	2.6654
r_{12-3}	4.1023	4.1043	4.0446	3.8313	3.8341	3.8406	3.8768	3.8703	3.8552
θ_{13-5-4}	170.1	168.2	172.3	169.8	168.7	173.8	177.6	177.0	178.3
θ_{11-1-2}	177.6	177.8	175.8	176.2	176.6	175.9	177.7	177.0	179.7

[a] Structures **1**, **2**, and **3** are the neutral hydrogen-bonded complex, the transition state, and the zwitterionic form, respectively.

Table 3. The Calculated Geometrical Parameters (Bond Length in Å and Angle in deg) for Compounds **1a**−**3a** in CCl_4

	1a[a]		**2a**[a]		**3a**[a]	
parameters	6-31G*	6-311+G**	6-31G*	6-311+G**	6-31G*	6-311+G**
r_{1-2}	1.7750	1.7632	1.2587	1.2518	1.0290	1.0262
r_{1-11}	0.9855	0.9807	1.2099	1.2099	1.7230	1.7257
r_{1-2}	2.7597	2.7431	2.4667	2.4599	2.7515	2.7511
r_{4-5}	1.0014	0.9995	1.0090	1.0064	1.0291	1.0263
r_{13-5}	2.0884	2.1117	1.8974	1.9166	1.7220	1.7252
r_{13-4}	3.0787	3.0955	2.8936	2.9067	2.7507	2.7508
r_{12-3}	4.1015	4.1037	3.8643	3.8727	3.9446	3.9507
θ_{13-5-4}	169.6	167.7	168.7	167.3	177.9	177.2
θ_{11-1-2}	177.1	177.1	175.5	175.7	177.8	177.2

[a] Structures **1a**, **2a**, and **3a** are the neutral hydrogen-bonded complex, the transition state, and the zwitterionic form, respectively.

Table 4. The Calculated Geometrical Parameters (Bond Length in Å and Angle in deg) for Compounds **1b**−**3b** in DMSO

	1b[a]		**2b**[a]		**3b**[a]	
parameters	6-31G*	6-311+G**	6-31G*	6-311+G**	6-31G*	6-311+G**
r_{1-2}	1.7286	1.7101	1.3260	1.3243	1.0157	1.0104
r_{1-11}	0.9935	0.9898	1.1545	1.1486	1.8765	1.9429
r_{11-2}	2.7207	2.6984	2.4777	2.4700	2.8919	2.9528
r_{4-5}	1.0002	0.9983	1.0030	1.0002	1.0156	1.0104
r_{13-5}	2.1413	2.1702	2.0264	2.0720	1.8788	1.9422
r_{13-4}	3.1284	3.1497	3.0131	3.0504	2.8941	2.9522
r_{12-3}	4.1041	4.1062	3.9255	3.9441	4.0823	4.1468
θ_{13-5-4}	168.8	166.6	167.4	165.4	178.3	177.9
θ_{11-1-2}	176.1	176.0	174.6	174.4	178.2	177.9

[a] Structures **1b**, **2b**, and **3b** are the neutral hydrogen-bonded complex, the transition state, and the zwitterionic form, respectively.

Table 5. The Calculated Geometrical Parameters (Bond Length in Å and Angle in deg) for Compounds **1c**−**3c** in Water

	1c[a]		**2c**[a]		**3c**[a]	
parameters	6-31G*	6-311+G**	6-31G*	6-311+G**	6-31G*	6-311+G**
r_{1-2}	1.7269	1.7084	1.3281	1.3266	1.0154	1.0099
r_{1-11}	0.9938	0.9901	1.1530	1.1469	1.8823	1.9529
r_{11-2}	2.7192	2.6970	2.4783	2.4706	2.8974	2.9623
r_{4-5}	1.0002	0.9982	1.0028	1.0001	1.0153	1.0099
r_{13-5}	2.1429	2.1721	2.0306	2.0759	1.8838	1.9540
r_{13-4}	3.1299	3.1515	3.0170	3.0542	2.8989	2.9635
r_{12-3}	4.1041	4.1063	3.9275	3.9461	4.0872	4.1570
θ_{13-5-4}	168.7	166.6	167.3	165.4	178.3	178.1
θ_{11-1-2}	176.1	176.0	174.6	174.4	178.2	177.7

[a] Structures **1c**, **2c**, and **3c** are the neutral hydrogen-bonded complex, the transition state, and the zwitterionic form, respectively.

and angle are within about 0.02 Å and 2°. The two N- - -O distances in **1** and **3** are about the same.

Carbon tetrachloride is a nonpolar, hydrophobic solvent with a very small dielectric constant (2.238) and the effect of CCl_4 on the structure and energy of a solute molecule is generally expected to be small. However, because the difference in dipole moment between the neutral hydrogen-bonded complex (3.84 Debye at the RHF/6-31G* level) and the zwitterionic form (10.02 D at the RHF/6-31G* level) is so large, even CCl_4 is able to have a relatively large effect on the energy difference between these two complexes. This is consistent with the $(\epsilon - 1)/(2\epsilon + 1)$ dependence in the Kirkwood−Onsager expression for the free energy of solvation of a dipole[28b,45] and also consistent with the report by Warshel that the calculated solvation energy for the $NH3^{+}$- - -COO^{-} ion pair in a hydrocarbon solvent is about 50% of that in water at a distance of 2.8 Å.[45a] The calculated energy difference between these two forms of hydrogen-bonded complexes is much smaller in CCl_4 compared to the gas-phase values. It has been shown experimentally that the hydrogen bond formed between benzoic acid and triethylamine, two compounds whose pK_a values differ by

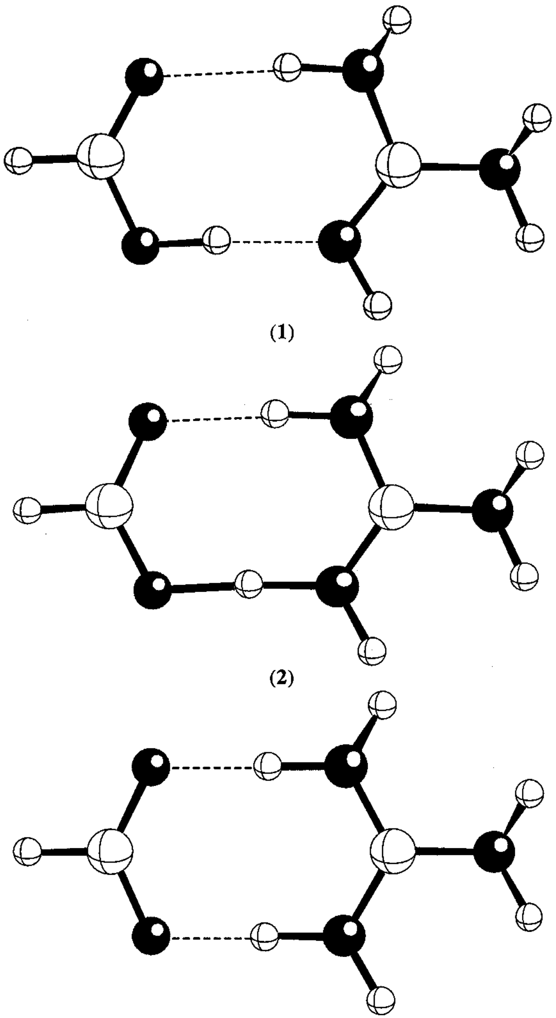
---

(45) (a) Warshel, A. *Isr. J. Chem.* **1981**, *21*, 341. (b) Jorgensen, W. L.; McDonald, N. A.; Selmi, M.; Rablen, P. *J. Am. Chem. Soc.* **1995**, *117*, 11809.

Table 6. The Calculated Energy Difference between the Two Forms of Hydrogen-Bonded Complex and the Interconversion Barrier (kcal/mol)

theory	energy diff[a]	forward barrier[b]
gas phase		
6-31G*	5.4	7.9
6-311+G**	5.5	7.5
MP2/6-31G*	3.4	3.9
MP2/6-311+G**	4.8	3.5
MP2/6-311++G**	4.8	3.5
CCl_4		
6-$_1$G*	−0.1	5.4
6-311+G**	0.2	5.0
MP2/6-31G*	−0.2	1.5
MP2/6-311+G**	1.0	1.3
DMSO		
6-31G*	−8.8	2.6
6-311+G**	−9.0	2.0
MP2/6-31G*	−5.1	0
MP2/6-311+G**	−6.1	0
water		
6-31G*	−9.1	2.5
6-311+G**	−9.3	2.0
MP2/6-31G*	−5.3	0
MP2/6-311+G**	−6.3	0

[a] Negative value means that the zwitterionic form is favored while a positive value means that the neutral form is favored. [b] This refers to energy difference between **2** and **3**.


(**1**)

(**2**)

(**3**)

Figure 1. The calculated geometries for the neutral hydrogen-bonded complex (**1**), the transition state (**2**), and the zwitterionic form (**3**) at MP2/6-31G* level.

as much as that between glutamic acid and lysine, shows little or no ionization in low dielectric, hydrophobic solvents such as CCl_4, cyclohexane ($\epsilon = 2.0$), and carbon disulfide ($\epsilon = 2.6$).[46]

Although the Onsager reaction field model has been successfully applied to several systems, there are still problems that need to be addressed. One problem is associated with defining the cavity radius.[47] The use of an elliptical cavity and including higher order electric moments may improve the situation,[35,48] but it still may not be adequate in some cases. To alleviate this problem, Wiberg and coworkers recently developed a new self-consistent reaction field model,[43,49] which is based on Tomasi's polarizable continuum model.[34] It has been applied to several systems and the results are very encouraging. As a further test, we repeated some of the above calculations in CCl_4 using the newly developed self-consistent reaction field method, which does not have the problem of the spheric cavity. The calculated relative energies and geometrical parameters at the RHF/6-311+G** level (data not shown) are very similar to the results at the same level using the Onsager reaction field model, suggesting that in this system, the Onsager model is quite reasonable. For example, the calculated barrier at the RHF/6-311+G** level using the new self-consistent model is 4.9 kcal/mol and the corresponding value using the Onsager model is 5.0 kcal/mol. As a result, no further calculations using this new self-consistent reaction field model were carried out. It should be pointed out that the new method is time consuming and further refinement is probably required to accelerate the convergence. We also examined the energy difference between the total energy using the Onsager reaction field model and the total energy from the new self-consistent reaction field model for the zwitterion in CCl_4 at two separation distances (3.5 and 4.0 Å); the calculated energy differences are 5.6 and 9.8 kcal/mol, respectively. As the separation distance gets larger, the difference between the Onsager model and the new self-consistent reaction field model increases, where as the molecule deviates more from the spherical model, the Onsager model becomes less reliable. Another problem with the reaction field model is related to the energy of cavity formation. So far, this term has been neglected in the *ab initio* implementation of the reaction field model. If the molecules are of similar size, the energy of cavity formation will cancel. In the present study, this poses no problem since the size of these three compounds (**1**−**3**) is essentially the same. However, if one wants to use the reaction field model to study a process such as a chemical reaction involving the association of two molecules, the energy of cavity formation has to be taken into consideration; this term can then be estimated using scaled-particle theory.[50]

In DMSO. As shown in Table 6, in a polar solvent like DMSO, the zwitterionic form becomes favored. At RHF/6-31G* and RHF/6-311+G** levels, both **1b** and **3b** are minima, but at the correlated levels the minimum for **1b** seems to disappear. The calculated geometries at both RHF/6-31G* and RHF/6-311+G** levels are again similar, but the difference is larger in DMSO than in the gas phase and in CCl_4. The difference in angles is about 2° and the difference in bond lengths can be as large as 0.06 Å (Table 4). At the RHF/6-

---

(46) DeTar, D. F.; Noval, R. W. *J. Am. Chem. Soc.* **1970**, *92*, 1361.

(47) Warshel, A.; Luzhkov, V. *J. Comput. Chem.* **1992**, *13*, 199.

(48) Rinaldi, D.; Rivail, J.-L.; Rguini, N. *J. Comput. Chem.* **1992**, *13*, 675.

(49) Foresman, J.; Keith, T. A.; Wiberg, K. B.; Frisch, M. J. To be published. For applications of this model, see ref 43 and the following: Wiberg, K. B.; Rablen, P. R.; Rush, D. J.; Keith, T. A. *J. Am. Chem. Soc.* **1995**, *117*, 4261. Wiberg, K. B.; Castejon, H.; Keith, T. A. *J. Comput. Chem.* **1996**, *17*, 185.

(50) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717. Prevost, M.; Oliveira, I. T.; Kocher, J.-P.; Wodak, S. J. *J. Phys. Chem.* **1996**, *100*, 2738.

311+G* level, the N- - -O distance in the zwitterionic form **3** is about 2.95 Å, which is in agreement with the X-ray crystallographic data for salt-bridges in proteins (2.93 Å for Arg-Asp and 2.94 Å for Arg-Glu).[51] As shown in Table 4, the transition state shifts toward the neutral hydrogen-bonded complex, which is in agreement with the Hammond postulate.[52] According to the Hammond postulate, as the reaction becomes more exothermic the transition state resembles the reactant (the neutral form) more than the product (the zwitterionic form).

In Water. The calculated geometrical parameters, energy differences, and interconversion barriers in water are given in Tables 5 and 6. The calculated results in water are very similar to those in DMSO. Although water has a much larger dielectric constant than DMSO, it seems that further increases in the dielectric constant have little effect on the calculated results. As expected, the zwitterionic form is much more favored in aqueous solution than in nonpolar solvent. Again the interconversion transition state resembles the neutral form more than the zwitterionic form. Recently, a statistical analysis of salt-bridges in 94 proteins has been carried out.[51] The average N- - -O distances are found to be 2.93 and 2.94 Å for Arg-Asp and Arg-Glu, respectively. Our calculated value is 2.96 Å at the RHF/6-311+G** level (Table 5), which is in excellent agreement with the experimental values, once again indicating that these salt-bridges are indeed in the zwitterionic form in the crystal structure grown from aqueous solutions.

Early studies by Warshel showed that ion pairs are not stable in nonpolar regions of a membrane.[26,45a] Based on electrostatic considerations, Honig and Hubbell concluded that the zwitterionic form of a salt-bridge is comparable in energy to the neutral form in the range of dielectric constants $\epsilon = 2-4$.[27] They considered salt-bridges of the Lys-Asp (or Glu) type. Although we did not investigate the interactions in Lys-Asp (or Glu) directly, the results from Arg-Asp (or Glu) type of salt-bridges are probably applicable to the Lys-Asp (or Glu) type of salt-bridges, since both types of salt-bridges have basically the same type of interactions. The results of the studies by Warshel and by Honig and Hubbell are consistent with our high-level *ab initio* molecular orbital calculations on the Arg-Asp (or Glu) type of salt-bridges. Warshel and co-workers have also demonstrated that the Asp-His pair is zwitterionic in the active site of serine proteases.[29]

Conclusions

To examine the effect of solvent environment on protein salt-bridges, we performed high-level *ab initio* molecular orbital calculations in the gas phase and in three different solvents using a model system. In the gas phase, the neutral conventional hydrogen-bonded complex is predicted to be favored at all four levels of theory. In a nonpolar, hydrophobic solvent like CCl_4, the energy difference between these two forms is small and the barrier that separates them is also low, but the neutral conventional hydrogen-bonded complex still seems to be slightly favored. However, in polar solvents like DMSO and water, the zwitterionic form becomes much more favored.

Ionized hydrogen bonds are often designated as salt bridges (or ion pairs) in crystallographic structures of proteins. In most cases, it is not known on which atom the proton resides since X-ray crystallography cannot locate hydrogen atoms directly. In low dielectric environments such as in nonaqueous solvents, membranes, and the interior of a protein, it is possible that the zwitterionic form will convert to the neutral hydrogen-bonded form via a proton shift; calculations suggest that this readily occurs in a very low dielectric environment. Since a protein is not a homogeneous system with the same dielectric constant, the "microenvironment" for each salt-bridge could be different. Normally, salt-bridges are also hydrogen bonded to other polar groups or nearby water,[25,26,51] so the "microenvironment" could be rather polar and the "effective" dielectric constant could be larger than normally expected in the interior of a protein, which is assumed to have a dielectric constant of 2−4 in most electrostatic models. Therefore, many interior salt-bridges of proteins can be expected to be zwitterionic.[25,26] Recently, Sauer and co-workers have demonstrated that the buried salt-bridge triad Arg31-Glu36-Arg40 in Arc repressor can be replaced with Met31-Tyr36-Leu40 (among other nonpolar combinations).[23] Therefore, the interconversion between the neutral and zwitterionic forms could be possible for surface salt-bridges of proteins in nonpolar organic solvents.

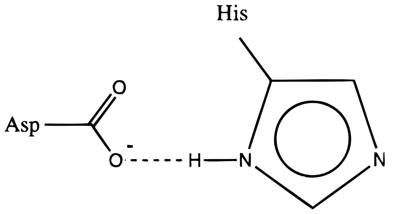
Molecular dynamics simulations on proteins in nonaqueous solvents have demonstrated that additional intra-protein hydrogen bonds and salt-bridges are formed when a protein is transferred from aqueous to nonaqueous solution.[8−11] An unanswered question is whether any surface salt-bridges are in the zwitterionic form in nonaqueous solution. Based on the present calculations, a salt-bridge can be expected to be in the zwitterionic form in a relatively polar "microenvironment". If it is in a nonpolar environment with very small dielectric constant, it will probably be in the neutral form. From the perspective of optimizing the stability of enzymes in nonaqueous solvents, an isolated charged group on a protein's surface, with no oppositely charged group nearby, may need to be removed (by site-directed mutagenesis), to prevent a reduction in thermostability in a nonaqueous environment, as demonstrated experimentally.[3] However if an isolated charged group which is stabilized by complementary protein dipoles (or electrostatic field of the protein)[25−26] is removed, it may actually decrease protein stability in a nonaqueous environment. Likewise, replacement of an isolated charge by an opposite charge group (such as Arg→Glu charge reversal) can increase protein stability provided this replacement makes favorable electrostatic interactions. If a charged group has an oppositely charged group nearby to which it could form a salt-bridge in a nonpolar, hydrophobic nonaqueous solution, the interconversion between the neutral and zwitterionic forms, via a proton shift, may be an effective way to annihilate charge separations. Removal of the latter type charged groups may not necessarily provide additional stability in hydrophobic, nonaqueous solution.

As noted above, the D181N mutant of subtilisin 8397 is less stable than subtilisin 8397 in high concentration of hydrophilic organic solvents, such as dimethylformamide. In the aqueous crystal structure of subtilisin 8397, Arg186 is very close to Asp181, but the orientation of the guanidinium part of Arg186 is such that a salt-bridge cannot form.[53] However, in a nonaqueous or mixed-solvent environment, the guanidinium part of Arg186 of subtilisin 8397 could undergo a rotation to form a salt-bridge with Asp181. Such a salt-bridge is not possible in the D181N mutant and could explain why this mutant is destabilizing relative to subtilisin 8397. Interestingly, a crystallographic study by Farber and co-workers on α-chymotrypsin in hexane[5] does indicate that nonaqueous solvents may induce significant rotation of the side chains of surface residues. This proposal can be verified easily by solving the crystal structure of the D181N mutant of subtilisin 8397 in nonpolar solvent such as hexane. Techniques such as neutron diffraction and IR

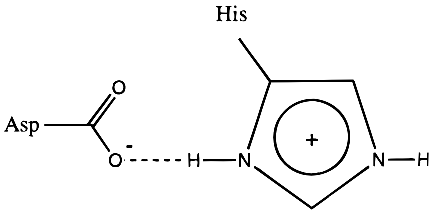
---

(51) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol*. **1995**, *254*, 761.
(52) Hammond, G. S. *J. Am. Chem. Soc*. **1955**, *77*, 334.
(53) Kidd, R. D.; Yennawar, H. P.; Sears, P.; Wong, C.-H.; Farber, G. K. *J. Am. Chem. Soc*. **1996**, *118*, 1645.

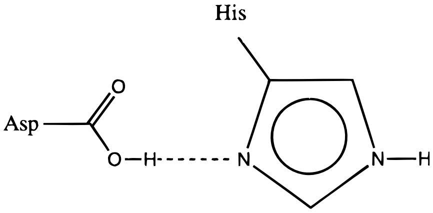
Chart 2



(**4**)



(**5a**)



(**5b**)

spectroscopy can also be used to examine the surface salt-bridges of proteins in nonaqueous solvents.

Are the present results in conflict with the notion of pH memory (or pH dependence) of enzymes in organic solvents?[1g,54] According to the studies by Klibanov and co-workers, the activity of lyophilized enzyme (such as subtilisin and α-chymotrypsin) powder in organic solvents depends on the pH of the aqueous solution from which the enzyme was lyophilized; thus it appears that titratable active site groups of a protein retain their ionization state upon lyophilization and that the same state of ionization is required of active site groups for catalytic activity in aqueous and nonaqueous solutions. For the enzymes studied by Klibanov and co-workers, the question is what is the form of the Asp-His pair of the catalytic triad (Asp102-His57-Ser195 in chymotrypsin and Asp32-His64-Ser221 in subtilisin) (**4** or **5**, see Chart 2); the Asp-His pair is more stable as a zwitterion (**5a**).[29] Our prediction of surface charge annihilation does not contradict the notion of pH memory, since charge annihilation is generally restricted to surface exposed or very nonpolar interior regions, while active site residues are often significantly buried and usually possess polar groups. Charge annihilation, however, could affect the pK_a's of catalytic residues in the active site, which may be partly responsible for some loss of activity and the observed shift in the activity versus pH profile.[55]

Acknowledgment. We thank the referees for their insightful comments and suggestions. This work was supported by the Laboratory Directed Research and Development program at Pacific Northwest National Laboratory as part of the Microbial Biotechnology Initiative (R.L.O). Pacific Northwest National Laboratory is a multiprogram national laboratory operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC06-76RLO 1830. Y.-J. Zheng acknowledges the Associated Western Universities for administering his postdoctoral fellowship.

JA960041O

---

(54) Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194.

(55) Schmitke, J. L.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3360.