Understanding the Hydrogen Bond Using Quantum Chemistry

MARK S. GORDON* AND JAN H. JENSEN

Department of Chemistry, Iowa State University, Ames, Iowa 50011-3111

Received April 22, 1996

Introduction

Understanding the origins and the fundamental nature of the chemical bond has been a central focus of chemists for decades. A chemical bond that has received considerable attention due to its importance in biological phenomena and in the relations between gas and condensed phases is the *hydrogen bond*, typically a relatively weak interaction involving an electronegative donor X, a hydrogen, and the electronegative acceptor Y:

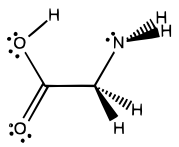
$$R{-}X{-}H\cdots{:}Y{-}R'$$

R−X−H and R′−Y are commonly (although not necessarily) separately identifiable molecules bound together by the weak hydrogen-bonding interaction. One may qualitatively understand the formation of the hydrogen bond in terms of the interaction between the electron deficient H and "available" lone pairs on Y. If R−X−H and R′−Y are indeed independent molecules, the stabilization due to this *inter*molecular hydrogen bond formation is simply

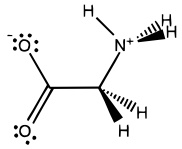
$$\Delta E_b = E[R{-}X{-}H] + E[R'{-}Y] - E[R{-}X{-}H\cdots Y{-}R'] \quad (1)$$

The most ubiquitous example of an intermolecular hydrogen bond is the interaction between two water molecules to form the water dimer (R = R′ = H, X = Y = O). The binding energy for the water dimer is about 5 kcal/mol (0.22 eV). Of course, the condensation of water vapor to liquid water occurs in large part as a result of the formation of hydrogen bonds between the individual water molecules. The process of freezing amounts to the formation of networks of hydrogen bonds in highly regular patterns. Hydrogen bonding is important in biological processes, because of the common occurrence of electronegative atoms, especially N and O, in biomolecules, such as amino acids and nucleotide bases. For example, the two strands of DNA are held together solely by hydrogen bonds. Since water is the common solvent in living systems, hydrogen bonding between biomolecules and water plays an important role in determining the energetics and dynamics of bioprocesses.

While it is sensible to ascribe the origin of hydrogen bonding to interactions between H and the lone pairs on Y, if we are to understand the fundamental nature of these interactions, hydrogen-bonded complexes (e.g., water dimer) must be considered carefully using a reliable theoretical model. To facilitate our discussion of how quantum mechanics can reveal the inner workings of hydrogen bonding, we will consider the fundamental nature of hydrogen bonding involving water. Following an analysis of the origin of the stability of the water dimer, we will turn our attention to the more complex problem of why solvation by water stabilizes the zwitterionic form of amino acids (specifically glycine), a structure that does not even exist in the gas phase!



Neutral Glycine



Glycine Zwitterion

Molecular Orbital Interpretations

The more sophisticated the wave functions used in the calculations, the more important it becomes to formulate interpretations that are easily understood. Two points must be addressed: (1) What are the best terms in which to formulate the interpretation? (2) How can the usual quantum chemical equations be rewritten to conform to these terms? If possible, we want terms that are already familiar to chemists. Soon after the electron was discovered, and before the advent of quantum chemistry, Lewis[1] and Langmuir,[2] among others, discovered that much of chemistry can be explained by assigning a pair of electrons to each bond, with the remaining electrons distributed in pairs as atomic inner shells and "lone pairs" localized on the atoms. For example, in the Lewis notation, boron hydride would be written as

$$:B{:}H \quad \text{or} \quad :B{-}H$$

Each dot represents a valence electron (the two core electrons are implicitly included in the "B"). This localized electron model also proved able to qualitatively explain the structure of molecules through the valence shell electron pair repulsion theory.[3] The concepts of inner shell, bond, and lone pair electrons

Mark S. Gordon was born in New York, NY, in 1942. He received his B.S. degree from Rensselaer Polytechnic Institute in 1963 and his Ph.D. with John Pople at Carnegie Mellon University in 1968. Following postdoctoral research with Klaus Ruedenberg at Iowa State University, he joined the faculty at North Dakota State University, where he achieved the rank of Distinguished Professor. In 1992, he joined the chemistry faculty at Iowa State University. His research spans a broad range of interests in quantum chemistry and dynamics.

Jan H. Jensen was born in Ærøskøbing, Denmark, in 1969. He received his B.A. degree from Concordia College, Minnesota, in 1989 and his Ph.D. with Mark Gordon at Iowa State University in 1995. He is currently a postdoctoral researcher with Professor Gordon. His research interests include interpretation of quantum mechanical calculations in general and intermolecular interactions in particular.

(1) Lewis, G. N. *J. Am. Chem. Soc.* **1916**, *38*, 762−785.
(2) Langmuir, I. *J. Am. Chem. Soc.* **1919**, *41*, 868−934.
(3) Gillespie, R. J.; Nyholm, R. S. *Q. Rev.* **1957**, *11*, 339−381.

S0001-4842(96)00059-3 CCC: $12.00

has become so ingrained in chemists' minds that it is scarcely thought of as a model. So, a quantum chemical interpretation of results in terms of these three types of electron pairs is sensible.

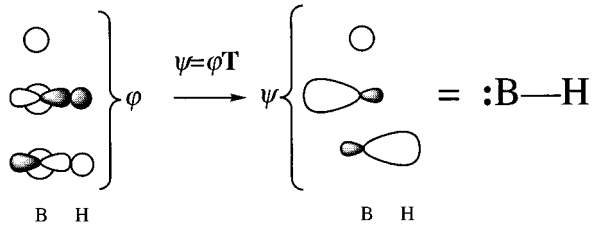
Localized Molecular Orbitals. The molecular orbitals obtained by solving the Hartree−Fock equations are delocalized over the entire molecule and do not conform to the pairs outlined above. Indeed, such delocalization is required by the molecular symmetry. However, these *canonical* MOs are not unique, since any unitary transformation among the set of doubly occupied orbitals yields the same total electron probability density and therefore the same properties, including the molecular energy. The same is true for the singly occupied subspace. While there is an infinite number of such unitary transformations, a particularly useful and well-defined transformation is that which converts the canonical orbitals into orbitals that are maximally localized. The basic concept for transforming to such *localized molecular orbitals* was suggested by Lennard-Jones and Pople,[4] and a practical method for obtaining the most localized orbitals was proposed by Edmiston and Ruedenberg,[5] who noted that the total energy within the Hartree−Fock scheme may be written as a sum of one- and two-electron terms,

$$\langle E\rangle = \langle E_1\rangle + \langle E_2\rangle \tag{2}$$

$\langle E_1\rangle$ and $\langle E_2\rangle$ are separately invariant to any unitary transformation of the canonical orbitals. $\langle E_2\rangle$ can be further subdivided as

$$\langle E_2\rangle = C - X = C' - X' + D \tag{3}$$

where C and X, the total Coulomb and exchange terms, respectively, are each also invariant to any unitary transformation of the canonical orbitals. Defining D as the sum of all intraorbital repulsions, C' and X' are the net interorbital Coulombic and exchange repulsions, respectively. Since C', X', and D *do* change under a unitary transformation of the canonical orbitals, one can seek that transformation **T** which maximizes D (or equivalently minimizes $C' - X'$). The resulting orbitals $\psi = \varphi\mathbf{T}$ are referred to as the energy-localized molecular orbitals (LMO). These LMOs do indeed look like bond, lone pair, and inner shell orbitals. So, they conform more closely to a chemist's view of molecular electronic structure than do the delocalized canonical orbitals. The BH molecule has six electrons in three doubly occupied MOs. The three canonical orbitals can be energy localized to give a core, lone pair, and bond MO:



Similar transformations can be applied to the integrals over canonical MOs to transform them to integrals over LMOs.

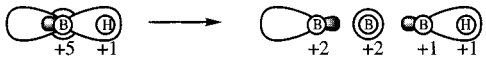
Local Nuclear Charges. Using LMOs, one can analyze the electron−electron repulsion, exchange, and electronic kinetic energy in terms of inner shells, bonds, and lone pairs. However, the remaining two terms in the total energy—the electron−nuclear attraction and nuclear−nuclear repulsion—are not easily assigned to these LMOs. To accomplish such a partitioning, the nuclear charge distribution must be partitioned in a way that is complementary to the partitioning of the electronic wave function. This is accomplished by assigning a local nuclear charge distribution[6,7] [$Z_i(A)$ for all atoms A] to LMO i such that

$$\begin{aligned} Z_i(A) &= 2 \text{ if } \psi_i \text{ is an inner shell or a} \\ &\qquad \text{lone pair LMO localized on atom } A \\ &= 1 \text{ if } \psi_i \text{ is a bond LMO localized on} \\ &\qquad \text{atom } A \text{ and its bonded partner} \\ &= 0 \text{ otherwise} \end{aligned} \tag{4}$$

In this way the total nuclear charge of an atom is preserved,

$$\sum_{i=1}^{N/2} Z_i(A) = Z_A \tag{5}$$

Consider the BH molecule. The three LMOs obtained above can be used to define their corresponding localized nuclear charge distributions: inner shell and lone pair LMOs are assigned +2 charges positioned at the *one* atom on which they are localized, whereas the bond LMOs are assigned +1 charges on each of the *two* atoms on which they are localized:



These three types of localized charge distributions can be used to describe most, but not all, charge distributions. A charge distribution with a formal net charge, for example, requires special attention and is addressed in later sections. This "local" nuclear charge (or charges) and the LMO, respectively, constitute the nuclear and electronic parts of an electrically neutral *localized charge distribution* (LCD).

Using the expression for Z_A in eq 4, the total energy of the system can be written as

$$\begin{aligned} E &= \sum_{i}^{N/2}[T_i + \sum_{j}^{N/2}(V_{ij} + G_{ij} + g_{ij}) + \sum_{j}^{N/2} e_{ij}^{(2)}] \\ &= \sum_{j=1}^{N/2}[T_i + \sum_{j=1}^{N/2} v_{ij} + \sum_{j}^{N/2} e_{ij}^{(2)}] \end{aligned} \tag{6}$$

where T, V, G, and g in eq 6 represent the kinetic, electron−nuclear attraction, electron−electron repulsion, and nuclear−nuclear repulsion energies, respectively, and $e^{(2)}$ gives the correlation correction from second-order perturbation theory.

Computational Details

Level of Theory. Molecular structures can be predicted by minimization of the energy, within the Born−Oppenheimer approximation. For the water

(4) Lennard-Jones, J.; Pople, J. A. *Proc. R. Soc. London* **1950**, *A202*, 166−180.
(5) Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457−465.
(6) England, W.; Gordon, M. S. *J. Am. Chem. Soc.* **1971**, *93*, 4649.
(7) Jensen, J. H.; Gordon, M. S. *J. Phys. Chem.* **1995**, *99*, 8091.

dimer we use geometries obtained with correlated wave functions. For the larger glycine−water study we use Hartree−Fock[8,9] geometries that serve as input for correlated energy calculations at the MP2 level of theory.[10,11] The levels of theory used in both studies have been shown to give results that are essentially identical to results obtained with higher levels of theory. So, our results are well converged with respect to the MO basis and electron correlation corrections.

Parallel GAMESS. The calculations described in the following two sections have been facilitated by dramatic advances in computational hardware and software, notably the development of algorithms that take advantage of parallel architecture. The suite of programs used to carry out the calculations discussed below is the General Atomic and Molecular Electronic Structure System (GAMESS).[12] The most time-consuming features of GAMESS have been rewritten for parallel computers.[13−15] The scalability of these codes is excellent at the Hartree−Fock level of theory and is also good for correlated wave functions.

Water Dimer[7]

The water dimer has become *the* paradigm for the hydrogen bond. The molecular structure of the water dimer and a sketch of the valence LMOs are shown in Figure 1a. One water (hydrogen donor, **D**) donates a hydrogen to the hydrogen bond. The other water (hydrogen acceptor, **A**) accepts this hydrogen. The water dimer energy is about 5 kcal/mol below the energy of two separated waters. If we freeze all coordinates but the oxygen−oxygen distance (R) and calculate the binding energy for various values of R, we obtain the curve shown in Figure 1a. A satisfactory understanding of the water dimer hydrogen bond would be an explanation of why, as we decrease R, the energy drops and then rises. To proceed, we consider the following: the water dimer has 10 doubly occupied LMOs resulting in well over 100 energy terms, even considering the three potential energy components as one term! One solution is to divide the energy of the dimer into a contribution from the two water molecules plus an interaction energy,

$$E = E(\mathbf{A}) + E(\mathbf{D}) + E(\mathbf{A}|\mathbf{D}) \qquad (7)$$

This subdivision of the energy is accomplished simply by restricting the sums in eq 6 to run over a subset of the LCDs:

$$E(\mathbf{A}) = \sum_{i\in\mathbf{A}}[T_i + \sum_{j\in\mathbf{A}} v_{ij} + \sum_{j\in\mathbf{A}} e_{ij}^{(2)}] \quad \text{and} \quad E(\mathbf{A}|\mathbf{D}) = \sum_{i\in\mathbf{A}}[\sum_{j\in\mathbf{D}} 2v_{ij} + \sum_{j\in\mathbf{D}} 2e_{ij}^{(2)}] \qquad (8)$$

Once the energy is obtained in the form of eq 6 the energy analysis can be done on a spreadsheet. A plot of the three energy components at various values of R is shown in Figure 1b. It is apparent that the hydrogen-bonding curve arises from two competing forces: an attractive interaction energy and two repulsive internal energies. The former dominates for large values of R, and the latter dominates at short R. Within this energy partitioning scheme, the repulsion part of the water dimer potential energy curve as a function of R stems from the internal energies, *not* the interaction energy. This is unexpected, since we are accustomed to thinking in terms of intermolecular repulsion energies that depend explicitly on the separation.

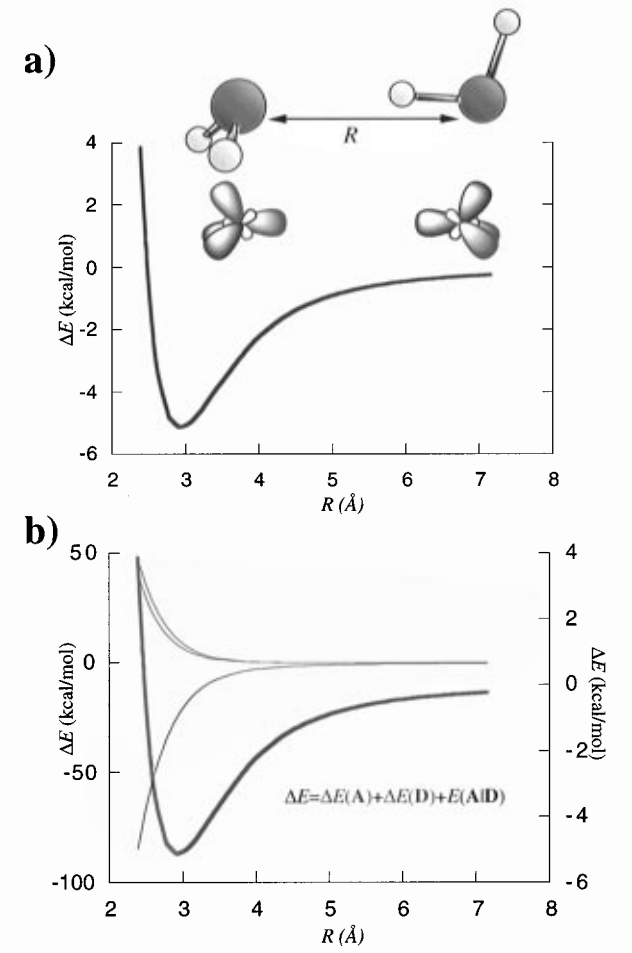


Figure 1. (a) Schematic representation of the water dimer geometry and lone-pair (red) and bond (gray) LMOs, and a plot of the change in the water dimer energy relative to that of free water and as a function of R. (b) Internal (red) and interaction energy (blue) components (left y-axis) of the total energy (composite curve, right y-axis).

Interaction Energy. A comparison of the SCF and correlation energy components of the interaction energy reveals that the latter is at most 9% of the total. We therefore defer the discussion of the correlation energy to a later section. The interaction potential energy [$PE(\mathbf{A}|\mathbf{D})$] has 25 LCD terms from the 5 LCDs on **A** interacting with the 5 LCDs on **D** (16 of these terms are symmetry-unique). However, we now show that only *three* are really responsible for the rapid decrease of $E(\mathbf{A}|\mathbf{D})$ as R is decreased. Figure 2a shows a breakdown of the interaction potential energy in terms of each LCD on **A** interacting with

(8) Levine, I. R. *Quantum Chemistry*; Allyn and Bacon: Newton, 1983.
(9) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; McGraw-Hill: New York, 1989.
(10) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, 618−622.
(11) Pople, J. A.; Binkley, S.; Seeger, R. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1−19.
(12) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(13) Windus, T. L.; Schmidt, M. W.; Gordon, M. S. *Chem. Phys. Lett.* **1993**, *216*, 375−379.
(14) Windus, T. L.; Schmidt, M. W.; Gordon, M. S. *Theor. Chim. Acta* **1994**, *89*, 77−88.
(15) Windus, T. L.; Schmidt, M. W.; Gordon, M. S. *ACS Symposium Series on Parallel Computing in Chemistry*, No. 592; Mattson, T., Ed.; American Chemical Society: Washington, DC, Chapter 2.

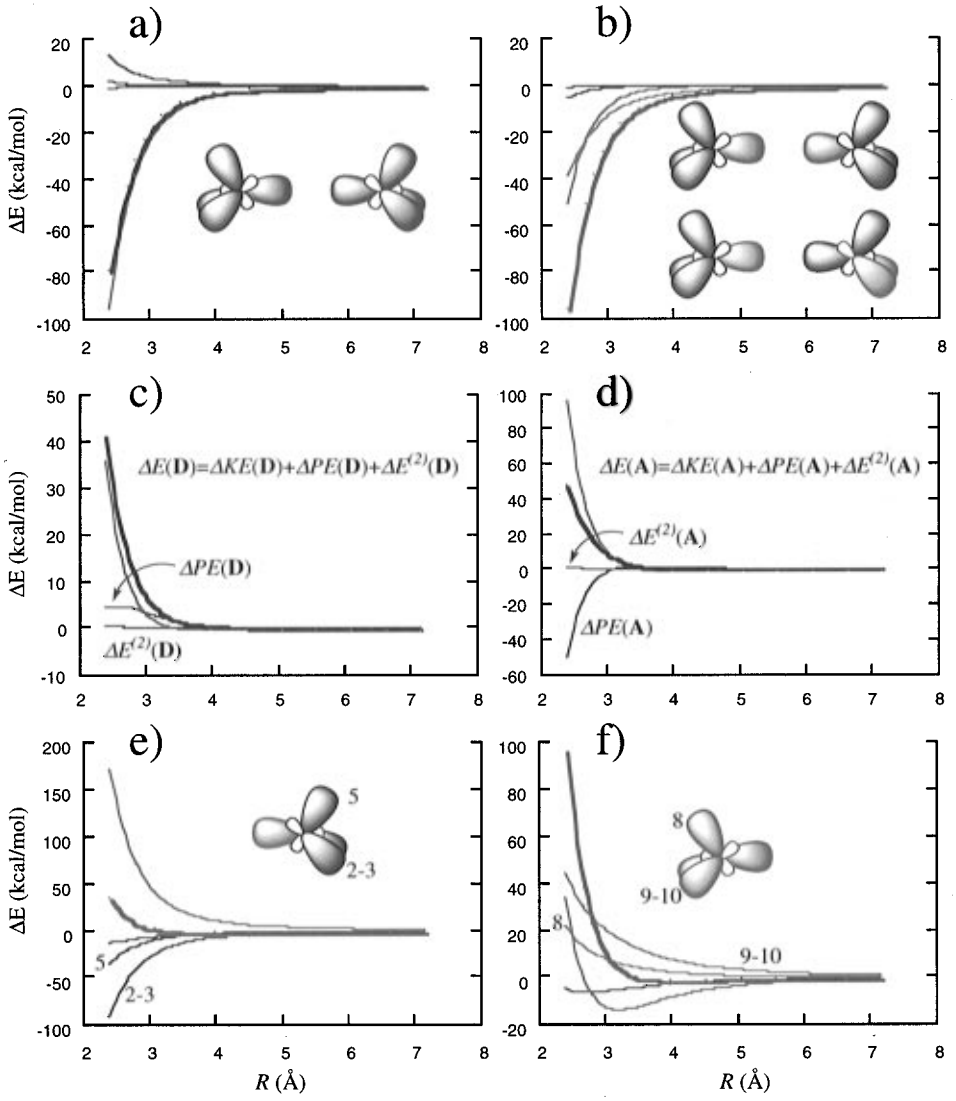
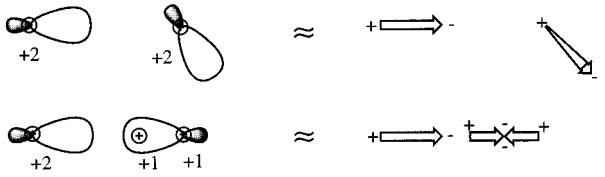


Figure 2. (a) Breakdown of the potential interaction energy in terms of each LCD on **A** interacting with all of **D**. (b) Breakdown of the lone pair−**D** interaction energy into individual LCD components. (c, d) Total energy change, and kinetic, potential, and correlation energy components of **D** and **A**, respectively. (e, f) LCD components of the electronic kinetic energy of **D** and **A**, respectively. The potential interaction energy is dominated by two LCD interaction terms; the rise in internal energy is due to the rise in the *kinetic energy* of certain LCDs.

all of **D**. Clearly the interaction term due to the hydrogen-bonded lone pair dominates, so we decompose this term into its LCD components on **D** in Figure 2b. Two terms stand out (the green curve represents two identical terms). These two terms correspond to the interaction of a lone pair LCD on **A** with an O−H bond and with two identical lone pairs on **D**. These two contributions constitute 51% and 66%, respectively, of the interaction potential energy. Only at this point is it necessary to consider the three potential energy components (eq 6) separately to obtain a deeper understanding.

As R is decreased, the components of the interaction potential energy will necessarily increase in *magnitude*: the electron−electron and nuclear−nuclear repulsions will increase, while the electron−nuclear attraction will decrease. The total potential interaction energy decrease indicates that the electron−nuclear attraction dominates. The question is why. To answer this, consider a generic lone pair and bond LCD. A lone pair LCD has a positive (nuclear) region and an equally negative (electronic) region, while a bond LCD has two positive regions separated by a negative region that contains twice the charge of either positive region. This corresponds to a dipole and quadrupole for the lone pair and bond LCD, respectively:

The attractive lone pair−lone pair and lone pair−quadrupole interactions are then easily explained as due to the negative part of the **A** lone pair being attracted to the positive regions of the two lone pair LCDs and one of the OH bond LCDs on **D**, respectively:



So the latter interaction represents, in essence, the conventional view of hydrogen bonding presented in

the Introduction. However, this is only half the picture. The other half comes from an interaction between the acceptor lone pair and the partly deshielded oxygen nucleus on the donor water.

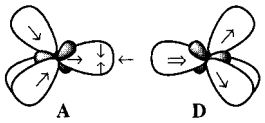
Internal Energies. Now, consider the source of the internal energy increase in the donor and acceptor waters as they are brought together. First we separate these two energies into their kinetic, potential, and correlation energy components of eq 6, as shown in Figure 2c,d. Clearly, the electronic kinetic energy dominates the other components and is thus responsible for the increase in the total energy at small R. The individual LCD components to the kinetic energy, shown in Figure 2e,f, show that (1) the kinetic energy increase of the hydrogen donor is clearly due to an increase in kinetic energy of the hydrogen-bonded OH bond LCD and (2) the origin of the sharp increase in the electronic kinetic energy of the hydrogen acceptor at short R is due to the hydrogen-bonded lone pair LCD, although the remaining valence LCDs contribute to the magnitude.

What causes these changes in electronic kinetic energy? A physical interpretation of the electronic kinetic energy can be obtained by writing the kinetic energy integral in terms of the gradient:

$$T_i = \int \psi_i(\mathbf{r}_1)[-{}^1/_2\nabla_1^2]\psi_i(\mathbf{r}_1)\, d\mathbf{r}_1 = {}^1/_2\int [\nabla_1\psi_i(\mathbf{r}_1)]^2\, d\mathbf{r}_1 \tag{9}$$

It is apparent that a localized function will have a higher kinetic energy than a more delocalized one:

Comparing the LMOs of water dimer to their counterparts in free water, one finds that some have contracted and some have expanded. As the two waters are brought close together, the electron clouds of the lone pair on **A** and the OH bond on **D** involved in the hydrogen bond are pushed apart.



The OH bond electrons accumulate around the oxygen of **D**. This pulls the electrons closer to the nucleus, the LMO contracts, and the kinetic energy rises. The remaining valence LMOs on **D** are pushed away from the oxygen, so they expand slightly. The hydrogen-bonded lone pairs on **A**, while repelled by the OH bond LMO, are attracted by the H in the OH bond on **D** and thus concentrate nearby: The LMO contracts and the kinetic energy rises. This concentration causes the lone pair electrons to be pulled away from the **A** oxygen, and the remaining valence LMOs contract. Thus, their kinetic energies also goes up.

Of course, this electronic rearrangement will also affect the internal potential energy. However, these changes are dominated by the kinetic energy (Figure 2c−f) and will not be discussed further.

Electron Correlation. The large magnitude of the various Hartree−Fock energy components relative to

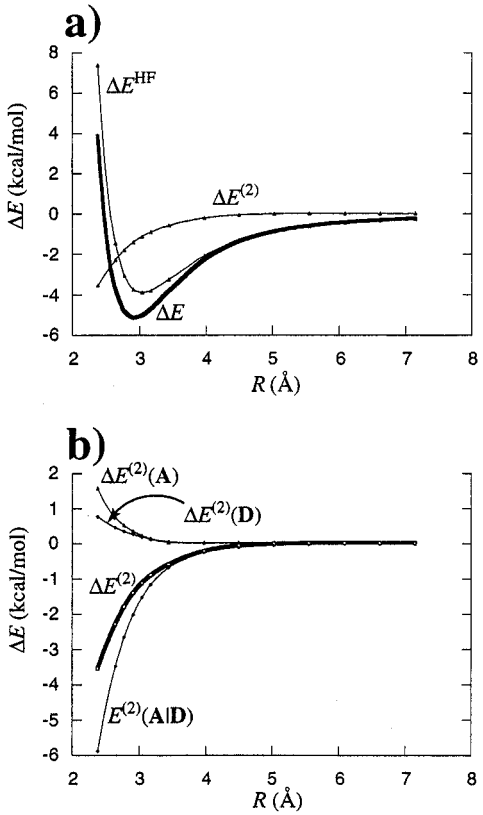


Figure 3. (a) Breakdown of the total energy in terms of a Hartree−Fock and electron correlation component. (b) Breakdown of the electron correlation component into internal and interaction components. The main error introduced by the Hartree−Fock approximation is an underestimation of the interaction energy.

that of the respective correlation energy components makes it difficult to analyze both simultaneously. If we separate the Hartree−Fock and correlation energy components of the total energy (Figure 3), we find that Hartree−Fock theory gives a hydrogen-bonding curve similar to that of the total energy. As the two water molecules are brought closer together, electron correlation, which lowers the relative energy, becomes more important. The addition of electron correlation lowers the binding energy by 1.4 kcal/mol and decreases the minimum energy value of R by 0.1 Å. A partition of $E^{(2)}$ into internal and interaction components (Figure 3) reveals that this energy increase comes exclusively from the interaction energy component. Thus, the main error introduced by the Hartree−Fock approximation for this particular system is an overestimation of the interaction energy. For example, MP2 lowers the binding energy by 2.0 kcal/mol at the equilibrium value of R.

Glycine and Water[16]

One reason that hydrogen bonding is worthy of detailed examination is its importance in solvation and biomolecular chemistry. To study the effect of solvation by water on biomolecular chemistry, it is useful to choose a biomolecule whose chemistry is relatively well understood when it is solvated and when it is in the gas phase. The difference in the chemistry in the two environments can then be attributed to solvent effects. The amino acid glycine provides a good test case for studying solvent effects. Depending on the

(16) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1996**, *117*, 8159.

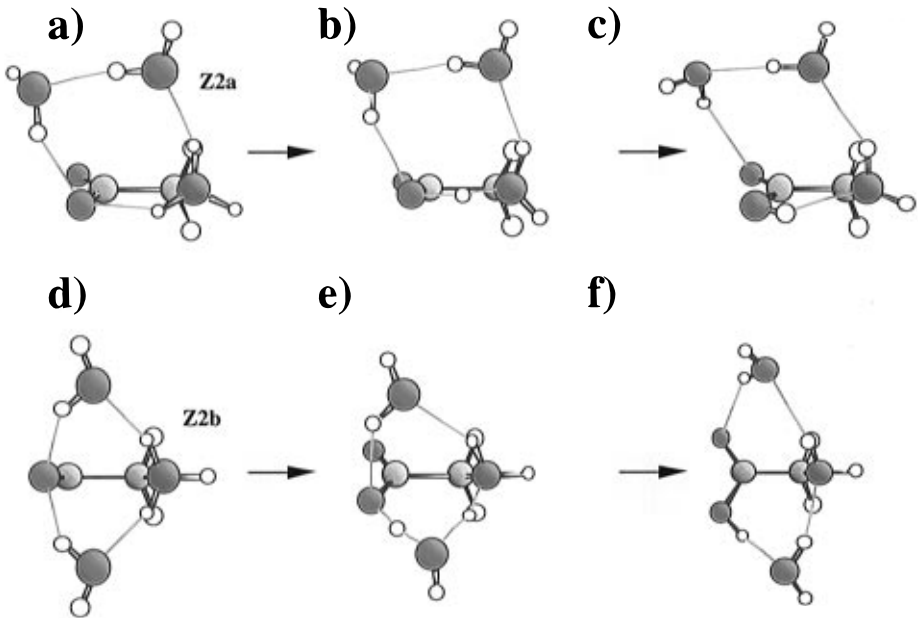


Figure 4. (a, d) Two dihydrated glycine zwitterion structures and their respective reaction paths leading to neutral structures.

environment, glycine can exist in one of two forms: zwitterionic or neutral. In the gas phase the neutral structure is preferred by about 20 kcal/mol (based on theoretical calculations on a hypothetical zwitterion structure), while the zwitterionic form is preferred by about 10 kcal/mol when glycine is dissolved in water. There is no barrier separating the zwitterion and neutral structures in the gas phase, so the zwitterion does not exist. The barrier height in the aqueous phase is unknown. In principle it should be possible to observe the transition from the gas phase behavior to that of the aqueous phase by successively increasing the number of water molecules that surround glycine. The first key point in this transition would be the minimum number of water molecules necessary to induce a barrier to proton transfer and thus stabilize the zwitterion. The research described in the last part of this Account concerns itself with this point.

Dihydrated Glycine. In Figure 4a,d we present two dihydrated zwitterionic glycine structures (**Z2a** and **Z2b**) that have essentially identical energies. Structure **Z2a** has an intramolecular hydrogen bond in addition to the intermolecular hydrogen bonds to the water molecules. Structure **Z2b** has more intermolecular hydrogen bonds but lacks an intramolecular hydrogen bond. Therefore, intramolecular transfer of the proton directly from the NH_3^+ group to the COO^- group is possible only for **Z2a**. Proton transfer in **Z2b** must be assisted by one of the water molecules. The two proton transfer mechanisms are illustrated in Figure 4. The energy associated with each proton transfer mechanism is shown in Figure 5a. The intramolecular proton transfer occurs without a barrier, just as in the gas phase, while the water-assisted proton transfer must overcome a barrier of ~6 kcal/mol. The main conclusion is that just two water molecules can suffice to stabilize the zwitterionic form of glycine. The stabilization is accomplished by breaking the intramolecular hydrogen bond that provides a barrierless path to proton transfer, by forming intermolecular hydrogen bonds to the solvent. Why does the water-assisted proton transfer mechanism require more energy than the intramolecular mechanism? A complete answer would entail an analysis

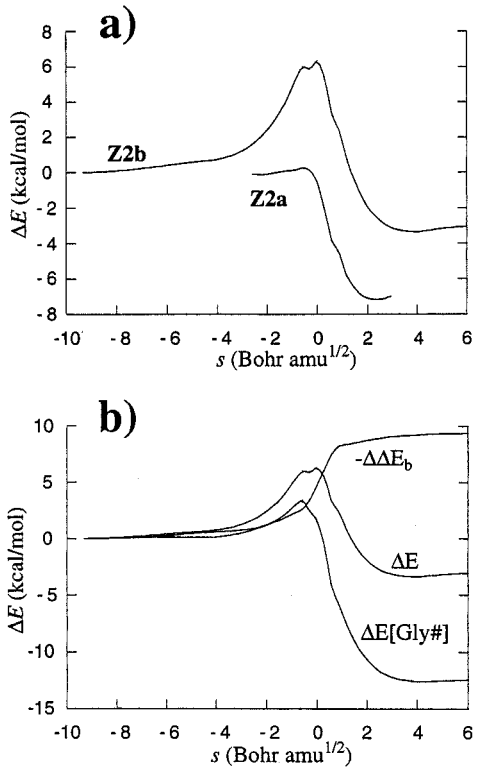


Figure 5. (a) Proton transfer reaction path for **Z2a** and **Z2b**: there is a barrier to water-assisted proton transfer but not to intramolecular proton transfer. (b) Decomposition of the **Z2b** reaction path into intrinsic and binding energy components: both contribute to the barrier.

of both mechanisms. Here we concentrate on the origin of the barrier for water-assisted proton transfer.

The analysis starts by defining the binding energy as

$$\Delta E_b = E[\mathrm{Gly}(\mathrm{H_2O})^{\#}] + E[\mathrm{H_2O}] - E[\mathrm{complex}] \quad (10)$$

$E[H_2O]$ is the (constant) energy of a water molecule and $E[Gly(H_2O)^{\#}]$ is the intrinsic energy. The latter is the energy of the glycine molecule plus that of the water molecule assisting the proton transfer. It is obtained by removing the other ("spectator") water

molecule in a given structure and recalculating the energy. The *relative* binding energy (calculated along the reaction path relative to the zwitterion complex) is calculated as

$$\Delta\Delta E_b = \Delta E[\mathrm{Gly(H_2O)^{\#}}] - \Delta E \qquad (11)$$

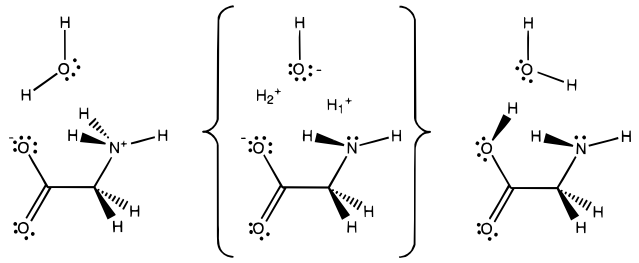
This expression can be rearranged so that the total relative energy is expressed in terms of relative intrinsic ($\Delta E[\mathrm{Gly(H_2O)^{\#}}]$) and binding energy ($\Delta\Delta E_b$) components:

$$\Delta E = \Delta E[\mathrm{Gly(H_2O)^{\#}}] - \Delta\Delta E_b \qquad (12)$$

Figure 5b shows these components evaluated as the water-assisted proton transfer proceeds. Both energy components contribute to the barrier: the water-assisted proton transfer mechanism intrinsically gives rise to a barrier, which is increased by the presence of the solvent molecule. This barrier increase is due to a decrease in the binding energy (i.e., *the intermolecular hydrogen bonds are weakening*) as the proton transfer proceeds.

Intrinsic Energy. Figure 5b suggests that proton transfer via the water molecule (water-assisted proton transfer) has an intrinsic barrier. To investigate this further, the water-assisted proton transfer reaction path was recalculated without the spectator water. This "gas phase" reaction path is virtually identical to the reaction path calculated in the presence of a spectator water. So we can conclude that the barrier to water-assisted proton transfer in dihydrated glycine is not caused by the presence of the spectator water molecule.

The energy localization of the SCF wave function of monohydrated glycine yields 25 localized molecular orbitals (LMOs). To obtain a continuous description of the structures along the reaction path, reactants, products, and all intermediate structures must be described by identical sets of LCDs. One solution is to describe all structures as the composite $[NH_2CH_2COO^- + OH^- + H_1^+ + H_2^+]$:



To facilitate this, two additional "LCDs", each consisting only of a +1 charge at the position of the proton being transferred are defined (i.e., no LMO is associated with the LCD). These 2 additional LCDs result in a total of 27 LCDs, hence over 400 LCD interaction terms. So these LCDs must be grouped into larger groups, as in the analysis of the water dimer.

We rely on chemical intuition to make sensible choices. Glycine contains three functional groups: COO^-, CH_2, and NH_2. Thus, the 27 LCDs are divided into 6 functional groups as shown in Figure 6a: (1, 2) the proton LCDs, (3) the NH_2 group contains the N lone pair, the two NH bonds, the N core, and the CN bond LCD, (4) the CH_2 group contains the C core and the two CH bond LCDs, (5) the OH^- group contains

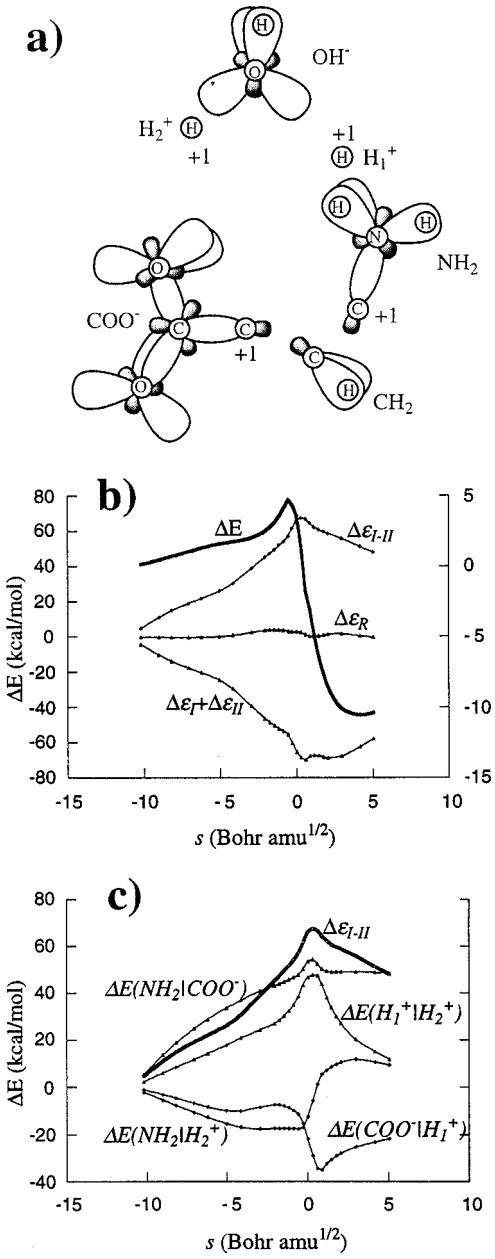


Figure 6. (a) Definition of the six functional groups of monohydrated glycine. (b) Energy components of the total energy along the reaction path of the gas phase water-assisted proton transfer: the barrier is due to the interaction term between the two simultaneous proton transfer processes. (c) Decomposition of the interaction energy in terms of its components: the energy of interaction of the NH_2 and COO^- groups and of the two protons dominates.

the O core, the OH bond, and the three O lone pairs, (6) the COO^- group consists of the remaining LCDs. The formal −1 charges on the latter two groups are assigned for convenience to the O atoms involved in the proton transfer by assigning +5/3 to each of their three lone pair LMOs.

The energy of each functional group and the energy of interaction with other groups are calculated in a manner similar to that for the water dimer (eq 7). We can simplify the energy analysis even further by decomposing the proton transfer into two separable but simultaneous processes: (I) proton transfer from the NH_2 group to the OH^- group and (II) proton transfer from the OH^- group to the COO^- group. Both processes have an associated proton transfer energy (ϵ_I and ϵ_{II}, respectively) which, together with an

interaction energy and a "remainder" energy ϵ_R due to the methylene group, sum to the total energy:

$$E = \epsilon_I + \epsilon_{II} + \epsilon_{I-II} + \epsilon_R \quad (13)$$

The energy terms are given by

$$\epsilon_I = E(NH_2) + E(NH_2|H_1^+) + {}^1/_2E(OH^-) + E(OH^-|H_1^+) + E(NH_2|OH^-)$$

$$\epsilon_{II} = E(COO^-) + E(COO^-|H_2^+) + {}^1/_2E(OH^-) + E(OH^-|H_2^+) + E(COO^-|OH^-)$$

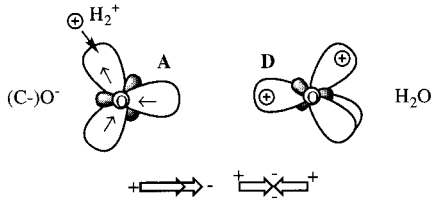
$$\epsilon_{I-II} = E(H_1^+|H_2^+) + E(NH_2|COO^-) + E(NH_2|H_2^+) + E(COO^-|H_1^+) \quad (14)$$

$$\epsilon_R = E(CH_2) + \sum_{Y \neq CH_2} E(CH_2|Y)$$

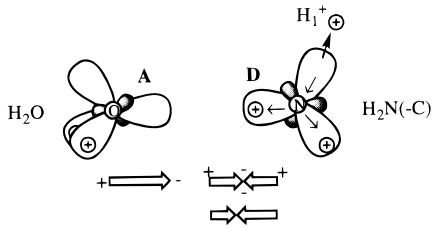
Only half of the OH^- energy is included in ϵ_I and ϵ_{II} to avoid double counting.

Figure 6b shows a plot of $\Delta\epsilon_I + \Delta\epsilon_{II}$, $\Delta\epsilon_{I-II}$, and $\Delta\epsilon_R$ evaluated relative to their values in the zwitterion, along the water-assisted proton transfer reaction path. Within this energy partitioning scheme the energy barrier is due to $\Delta\epsilon_{I-II}$. The combined energy from the two simultaneous proton transfer reactions, $\Delta\epsilon_I + \Delta\epsilon_{II}$, is dominated by $\Delta\epsilon_I$ since H_1^+ is being transferred between the two groups with the largest difference in proton affinity (NH_2 and OH^-). When the proton is transferred from NH_2 to OH^-, $\Delta\epsilon_I$ is always decreasing. Figure 6d shows a breakdown of $\Delta\epsilon_{I-II}$ into its four components (eq 14). The two most important contributions to $\Delta\epsilon_{I-II}$ are the interactions between the two protons and between the NH_2 and COO^- groups. The former energy is simply a reflection of the change in the proton−proton distance during the reaction. The source of the latter energy increase can be traced to the increased repulsion between the N lone pair and the lone pairs on the negatively charged O. As the protons are transferred the negative end of the N lone pair LCD dipole is brought closer to the negatively charged O, leading to a strong energy increase.

Binding Energy. From the analysis of the water dimer we have a good understanding of what affects the binding energy. This allows us to offer a qualitative description of the effect that the double proton transfer has on the water−glycine interaction energy. We qualitatively depict the hydrogen bonds to the negatively charged oxygen in the COO^- group and the NH bond in the NH_2 group (where the water molecule is the hydrogen donor and acceptor, respectively) as dipole−quadrupole interactions. First consider the effect of protonating the COO^- group on the former hydrogen bond. The approaching proton (H_2) will attract the electrons in the receiving lone pair and thus pull them away from the oxygen. Therefore, the electrons in the other two lone pairs are pulled toward the oxygen, as depicted in the schematic. This will shorten the lone pair LCD dipole and thus weaken the hydrogen bond.



The hydrogen bond to the NH_2 group may be explained in a similar manner. The departing proton (H_1) exerts a weakening pull on the lone pair electrons which are therefore pulled toward the nitrogen. This electron pair then repels the electron pairs in the two adjacent bonds which consequently move closer to the hydrogen end of the bond. The shift in the quadrupole leads to a weaker hydrogen bond.



From the water dimer analysis we know that lone pair−lone pair interactions contribute to the binding energy as well, and these interactions will be affected by the proton transfer. However, the qualitative picture outlined above captures the essence of the effect of proton transfer on the two hydrogen bonds.

Summary and Conclusions

Reformulations of the Hartree−Fock and MP2 equations facilitate the interpretation of quantum mechanical results in terms of familiar chemical concepts, such as inner shell, bond, and lone pair electrons. Quantum mechanical calculations show that the hydrogen bond in the water dimer can be understood as the competition between the attractive interaction potential energy and the repulsive internal electronic kinetic energy of the water molecules. The former term can be approximated by classical electrostatic interactions. An understanding of this hydrogen-bonding model facilitates the analysis of the effect of solvation on amino acids, especially the role played by water molecules on the stabilization of the zwitterion form.

This work was supported by grants from the National Science Foundation (CHE-9311911) and the National Institute for Standards and Technology.

AR9600594