JCTC Journal of Chemical Theory and Computation

# The Low-Barrier Double-Well Potential of the $O^{\delta 1}-H-O^{\delta 1}$ Hydrogen Bond in Unbound HIV Protease: A QM/MM Characterization

Melissa A. Porter and Pablo A. Molina*

*Department of Chemistry, Murray State University, Murray, Kentucky 42071*

Received June 14, 2006

**Abstract:** The presence of a low-barrier hydrogen bond (LBHB) in aspartyl proteases and its implications in drug design have been the subject of intense study. Here, we present a combined quantum mechanical/molecular mechanical (QM/MM)−Numerov procedure and use it to characterize the $O^{\delta 1}-H-O^{\delta 1}$ hydrogen bond (HB) in unbound HIV protease. The QM/MM scheme fully traces the shape of the HB's potential energy curve. The potential is used to obtain numerical solutions to the wave functions and vibrational energies of hydrogen, deuterium, and tritium. The vibrational eigenfunctions are used to compute expectation values for interatomic distances and vibrationally and thermally averaged spectroscopic properties of the $O^{\delta 1}-H-O^{\delta 1}$ HB. Our work corroborates previous results by Piana and Carloni who found a LBHB via an ab initio molecular dynamics simulation (Piana, S.; Carloni, P. *Proteins* **2000**, *39*, 26−36). Our predictions of isotope effects on the chemical shift of unbound HIV protease are consistent with experimental measurements in similar HBs. These results support the predictive power of this method and its potential use in screening inhibitors of aspartyl proteases.

## Introduction

Twelve years ago, Frey and others postulated that low-barrier hydrogen bonds (LBHBs) are key components of enzymatic acceleration.[1,2] They proposed that the transformation of a weak hydrogen bond (HB) in the ground state into a LBHB in the transition state or in a transient intermediate could significantly contribute to catalytic efficiency. The structural characteristics of a LBHB are a short heteroatom separation and a highly deshielded proton occupying a central position in a double well where the zero-point vibrational energy is close to the energy barrier.[3] Yet, its physicochemical nature, requirements for existence, and actual importance in catalysis have been subject to an intense debate.[4−6] The $N^{\delta 1}-H-O^{\delta 1}$ HB of serine proteases was the first protein system in which the existence of this type of HB was postulated. The LBHB was then defined in terms of measured spectroscopic properties: an extreme downfield hydrogen chemical shift (16.0 ppm $\leq \delta_H \leq$ 21.0 ppm), a low fractionation factor ($\theta < 1$), and large differences in the chemical shifts upon isotopic substitution. However, as noted early by Perrin and Nielson, the ambiguity in the range of values given for these properties creates some difficulty in predicting the exact circumstances under which a LBHB will be observed.[3] In particular, $\delta_H$ and $\theta$ measurements cannot be used as the sole evidence for the presence of a LBHB: both a $\delta_H >$ 16.0 ppm and a $\theta < 1$ can correspond to LBHBs, but they can also correspond to "strong" (asymmetric double-well) or "very strong" (symmetric single-well) HBs.[7,8] $\Delta\delta_{H-D}$ measurements are not accurate because of the large errors arising from $\delta_D$'s broad band in the NMR spectrum. For instance, the $\Delta\delta_{H-D}$ of the chymotrypsin−AcL−$CF_3$ complex is reported as 1.1 ± 0.5 ppm.[9] In contrast, $\Delta\delta_{H-T}$—because of $\delta_T$'s narrow line width in the spectrum—is a much more precise measure of HB strength and provides reliable experimental evidence for LBHBs. Indeed, $\Delta\delta_{H-T}$ measurements suggest that, while the $N^{\delta 1}-H-O^{\delta 1}$ HB in the low-pH form of unbound serine proteases is "strong", it becomes a "LBHB" in transition-state analogues. Among chymotrypsin−inhibitor complexes, $\Delta\delta_{H-T}$ varies from 0.63 to 0.68.[10]

* Corresponding author fax: (270) 809-6474; e-mail: pablo.molina@murraystate.edu.

10.1021/ct600200s CCC: $33.50 
Published on Web 09/22/2006

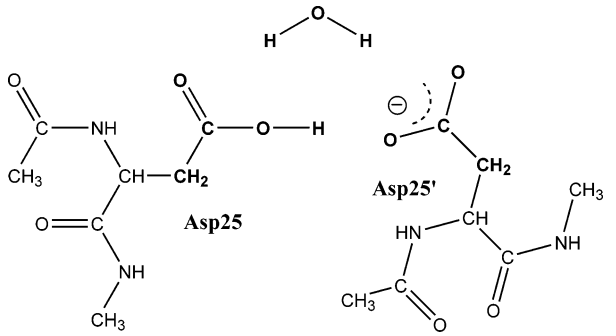


***Figure 1.*** Small model used for the initial optimization. Protomer Asp25-H is shown. The "QM" atoms relaxed during the optimization appear in bold.

LBHBs have also been reported in many other systems, for instance, aspartyl proteases. This family of enzymes is widely distributed in nature, and a number of aspartyl proteases have been identified as useful targets for chemotherapeutic intervention in human diseases.[11−13] Two notable members are memapsin 2—an enzyme associated with the pathogenesis of Alzheimer's disease—and HIV protease (Hp), a target for AIDS. Hp is a homodimer that cleaves the polyprotein products of the HIV-1 genome, producing several proteins necessary for viral growth and cellular infection. The active site of Hp consists of a monoprotonated aspartate dyad with a water molecule bridging the two Asp residues (Figure 1). The Hp−substrate complex exhibits $C_2$ symmetry, and numerous structure-based $C_2$-symmetric inhibitors have been designed to exploit this feature.[14] The symmetric nature of the $O^{\delta1}-H-O^{\delta1}$ HB is suggestive of the presence of a LBHB. Indeed, a computational study by Piana and Carloni on unbound Hp—via ab initio molecular dynamics simulations—found a LBHB between the $O^{\delta1}$ atoms that compensates for their negative charge repulsion and holds them within a distance of 2.5 (±0.1 Å).[15] Following this study, LBHB-based reaction mechanisms for Hp have been postulated and their implications in drug design have been discussed.[16] Northrop proposed that "looking to new designs, one might expect to achieve tighter binding by developing inhibitors that do not break the short, strong LBHB".[17] Given the incidence of aspartyl proteases as targets of major human diseases, efforts taken in Northrop's direction (i.e., relating "LBHB character" to inhibitory efficacy) could prove to be tremendously beneficial. As discussed above, experimental measurements of $\delta_H$, and more particularly $\Delta\delta_{H-T}$, can serve as indicators of HB strength. Yet, in aspartyl proteases, the absence of experimental data of these NMR properties for the $O^{\delta1}-H-O^{\delta1}$ HB hinders this line of study. In general, $\delta_H$ measurements of carboxylic acids are complicated because of rapid proton exchange with water; moreover, in the case of Hp, autolysis might create an additional problem.[15,18,19] In this paper, we introduce a computational model that can correct this experimental inadequacy. Our model can be used to investigate low-barrier double-well hydrogen bonds of biological importance such as the $O^{\delta1}-H-O^{\delta1}$ HB in aspartyl proteases but should also be transferable to other HB systems of a similar nature. The model is used here to characterize the $O^{\delta1}-H-O^{\delta1}$ HB of unbound Hp. Results are consistent with Piana and Carloni's benchmark work and offer the first predictions of NMR isotope effects in aspartyl proteases. To our knowledge, this is the first quantum mechanical/molecular mechanical (QM/MM) study that traces the entire shape of a double-well "LBHB" potential energy curve. The paper is organized as follows. First, we present the computational methodology in a stepwise manner. Next, we discuss our computed structural and spectroscopic properties in light of available experimental results and theoretical considerations. We finish by discussing the implications of our work and the usefulness of the methodology.

## Method

The model builds upon the method used by Molina and Jensen to characterize strong hydrogen bonds in serine proteases.[8] We trace the potential energy curve for proton transfer at the B3LYP/6-31G+(d,p) level of theory using the effective fragment potential (EFP) methodology, a QM/MM scheme in which the QM and MM regions are separated by a "buffer region" consisting of localized molecular orbitals.[20] We then obtain numerical solutions of the 1-D Schrödinger equation via the Numerov method.[21] The vibrational eigenfunctions are used to compute expectation values for interatomic distances and vibrationally and thermally averaged spectroscopic properties of the $O^{\delta1}-H-O^{\delta1}$ HB.

**A. Geometry Optimization.** The model for the $O^{\delta1}-H-O^{\delta1}$ HB is based on the same crystal structure utilized by Piana and Carloni (PDB code: 3PHV, 2.7 Å resolution).[22] A small 52-atom model of the active site (Figure 1) is built from the crystal structure coordinates; hydrogens are added with the online utility WhatIf Web Interface.[23] A water molecule is added manually and preoptimized semiempirically at the PM3 level. The small model is optimized at the B3LYP/6-31G(d) level of theory. When two different conformations with opposite protonation states are started with, two minima which we will call protomers Asp25-H and Asp25′-H are obtained. During the optimization, the Asp side chains and water molecule are relaxed while coordinates for the rest of the atoms remain frozen. This geometry setup leaves the carboxylates flexible and free to rotate, whereas their surrounding protein frame is kept fixed, thus mimicking the actual structural conditions of that protein region.[15] For practical purposes, the minima are equivalent with regard to energy and interatomic distances, as one could expect from the symmetry of the unbound Hp homodimer. The protomers differ in energy by just $\Delta E = 2.16 \times 10^{-2}$ kcal/mol, explicable within computational error; $\Delta r_{OH}$ is only $7.40 \times 10^{-4}$ Å, and differences in other internal coordinates are also very small. The protomers present a short $r_{OO}$ of 2.50 Å and an elongated $r_{OH}$ of 1.05 Å. The carboxylates exhibit a small departure from coplanarity ($\Omega = 32.6°$). The external oxygens are separated by 4.59 Å. The dihedral angle between the protonated oxygen, the oxygens of the facing carboxylate, and the water oxygen is 16.2°.

**B. Least Linear Motion Path.** Next, a 19-point least linear motion path (LLMP) is traced along the proton-transfer coordinate that connects the two minima. Single-point energies are obtained using a large QM/MM model of the dimer at the B3LYP/6-31+G(d,p)//B3LYP/6-31G(d) level of theory. The LLMP corresponds to the potential energy

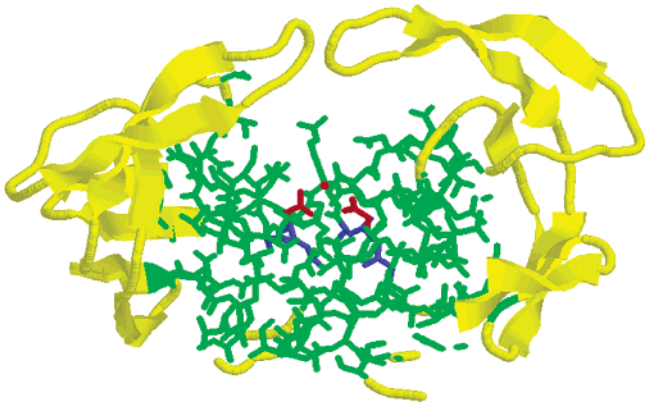

***Figure 2.*** The three computational regions used in the QM/MM model. QM/buffer/EFP regions (red, blue, and green). Note: hydrogen atoms not shown for clarity; water added manually.

surface (PES) of the $O^{\delta 1}-H-O^{\delta 1}$ hydrogen bond; coordinates changed are those of the 16 atoms that were allowed to relax during the optimization of the small model. The LLMP is constructed as follows: *Z* matrices with identical connectivity are built at the coordinates of the two minima and are used to calculate the changes in bond lengths, bond angles, and dihedral angles on going from Asp25-H to Asp25′-H. Then, nine intermediate points between the minima are generated by adding successive increments representing each internal coordinate change. Two additional points are created by scaling the coordinates by the factors necessary to decrease the $O^{\delta 1}-H$ bond length in both Asp residues to the O−H bond length (0.979 64 Å) of propionic acid optimized at the same level of theory. Last, using the *Z* matrix difference scheme, six additional points (three on each side) are placed between the minima and the points adjusted to the $r_{OH}$ in propionic acid.

**C. QM/MM Methodology.** The QM/MM model has a total of 814 atoms. It consists of three computational regions: QM/buffer/EFP (Figure 2). The QM region contains the 16 atoms that were allowed to relax in the optimization of the small model. The coordinates of these QM atoms are modified using the LLMP scheme and single points computed to construct the potential energy surface (see previous subsection). Note that, unlike HBs in nonhomogeneous environments, the unbound Hp homodimer affords an uncomplicated case of study of double-well potentials; the QM/MM−LLMP approach provides a sensible description of the energy barrier along proton transfer in this simple and symmetric system; structural relaxation might improve the accuracy of the computed barrier height but should not differ markedly from our estimate. The surrounding protein frame of the QM atoms is described by the buffer and EFP regions. The EFP region describes the electrostatic potential of the enzyme within roughly 12 Å of the active site by a multipole expansion (charges through octupoles at all atoms and bond midpoints and polarizable points for each valence-localized molecular orbital).[24] To create the EFP region, we use a "divide-and-conquer" approach.[25] The EFP has a total of six fragments that comprise residues 7−12, 81−91, and 21−33 (except QM and buffer atoms) in both monomers of unbound Hp. Each fragment is obtained via an RHF/6-31G(d) ab initio calculation, and then, all fragments are assembled together. The EFP and QM regions are separated by a buffer region of localized molecular orbitals obtained at the RHF/6-31G(d) level of theory.[20] There are a total of 28 buffer atoms comprising the backbone atoms around each Asp side chain (ending at the α carbons of neighboring residues). The geometries of the EFP and buffer regions are those of the crystal structure coordinates.

**D. Symmetrization of the Potential.** Polynomial regression of the PES traced by the computed data points yields the quartic function $f(x) = 985.86x^4 - 5066.7x^3 + 9656.4x^2 - 8084.9x + 2510$. Our potential is a little asymmetric to the left of the second minimum, most likely because of the sensitivity of the EFP points to small geometry changes along the LLMP. Note that point "12" had to be dropped because of an unreasonable energy. Thus, we symmetrize the PES (Figure 3) to improve the accuracy of the eigenvalues and

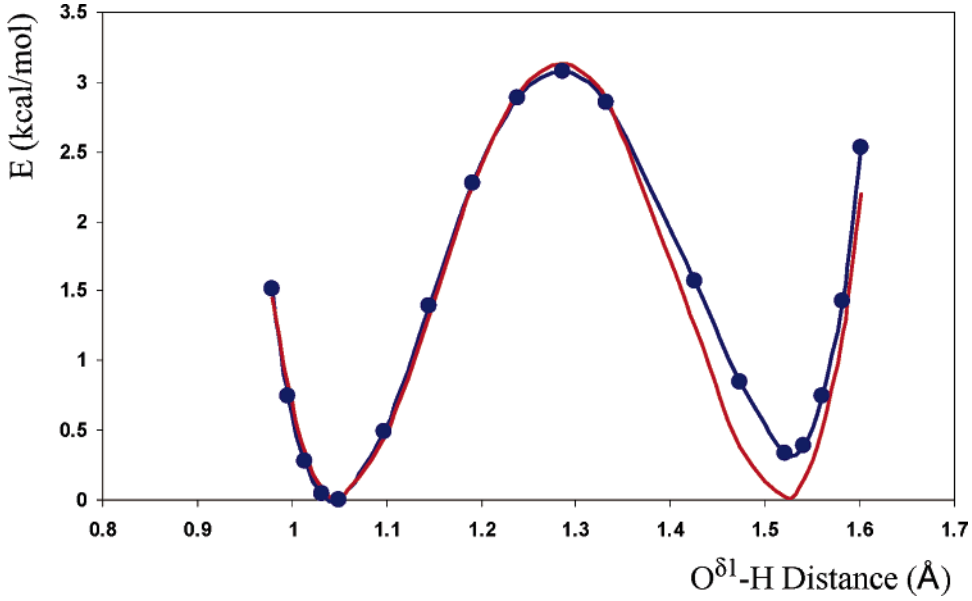


***Figure 3.*** Original curve (blue) obtained with the LLMP points. The symmetrized potential (red) is traced with the same points.

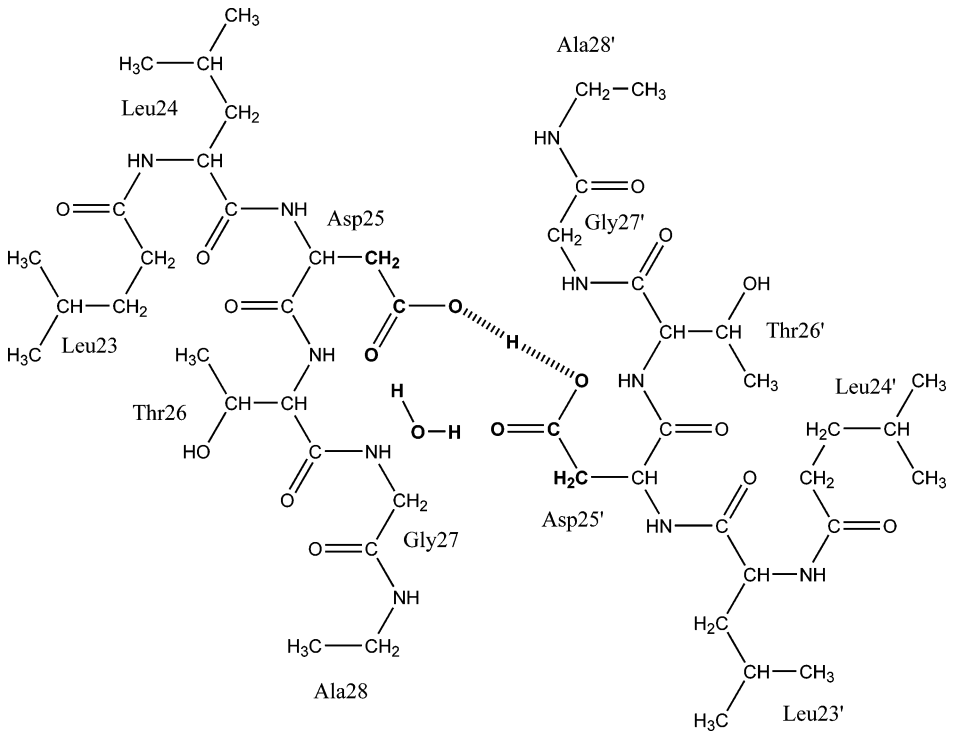


***Figure 4.*** The 162-atom model used to calculate the NMR chemical shifts. "QM" atoms are shown in bold.

the quality of the wave functions that are needed to compute structural and spectroscopic properties. Notice that unbound Hp is actually a symmetric homodimer, so no artificial correction is introduced. We construct a new symmetric function using the critical numbers of the quartic function: the $r_{OH}$ for the two minima ($x = 1.048\,963\,326$ and $x = 1.518\,033\,637$) and the $r_{OH}$ for the barrier maximum ($x = 1.283\,498\,482$). These are set as absolute minima and the relative maximum, respectively.

Hence,

$$f'(x) = a(x - 1.048\,963\,326)(x - 1.283\,498\,482)(x - 1.518\,033\,637)$$

This implies that

$$f(x) = a(-2.043\,793\,713x + 0.250\,000\,000x^4 - 1.283\,498\,482x^3 + 2.443\,549\,159x^2) + c$$

Now we have

$$f(1.048\,963\,326) = -0.633\,903\,878a + c \text{ and}$$
$$f(1.283\,498\,482) = -0.633\,147\,443a + c$$

Solving for "$a$" and "$c$" with this system of linear equations, we obtain a new symmetric function:

$$V(x) = 1031.647\,795x^4 - 5296.473\,517x^3 + 10\,083.528\,41x^2 - 8433.901\,12x + 2615.878\,315$$

**E. Numerov Method.** We now proceed to obtain eigenvalues and eigenfunctions for the vibrational energy levels of H, D, and T. The one-dimensional Schrödinger equation

$$-\frac{\hbar^2}{2m}\hat{d}^2\,\Psi(x) + \hat{V}(x)\,\Psi(x) = E\Psi(x) \qquad (1)$$

is solved numerically using 10 000 grid points along the above potential via the Numerov method.[6] For H, the kinetic energy operator $\widehat{E_K}$ in units of kilocalories per mole is

$$\frac{\hbar^2}{2m} = 0.0478\frac{\text{kcal}}{\text{mole}}\text{Å}^2 \qquad (2)$$

and $V(x)$ is given by our symmetrized potential. After correcting for the isotope mass, $\widehat{E_K}$ is 0.0239 for D and 0.0159 for T. The Numerov procedure uses a recursive relationship to compute the value of the wave function along the PES.[21] In practice, we obtain eigenvalues and eigenfunctions using the solver applet created by Schmidt.[26]

**F. Structural Properties.** The expectation value for the $r_{OH}$ bond length is calculated with

$$\langle r_{OH}\rangle = \sum_{i=1}^{100}[\psi(i)]^2\, r_{OH}(i)\,\Delta x \qquad (3)$$

The square of the normalized ground-state vibrational wave function ($\psi_0{}^2$) corresponds to the probability distribution function. The first point ($x_{min}$) starts at an $r_{OH}$ distance of 0.5 Å, and the remaining points are constructed via incremental steps ($\Delta x = 0.0\widehat{15}$ Å) up to $r_{OH} = 2.0$ Å ($x_{max}$).

$$\langle \pi\rangle = \sum_{i=1}^{23}[\psi_0(i)]^2\,\pi(i)\,\Delta x \qquad (4)$$

For hydrogen, the Boltzmann's population fraction of the first excited state is only slightly bigger than 0.1 (see the next subsection) and $\psi_1$ has two symmetric peaks around the PES midpoint. Hence, when the expectation value is averaged over the ground and first excited states, $<r_{OH}>$ is imperceptibly affected. Other structural properties ($\pi$) are obtained via expectation values, $<\pi>$, averaged over 23

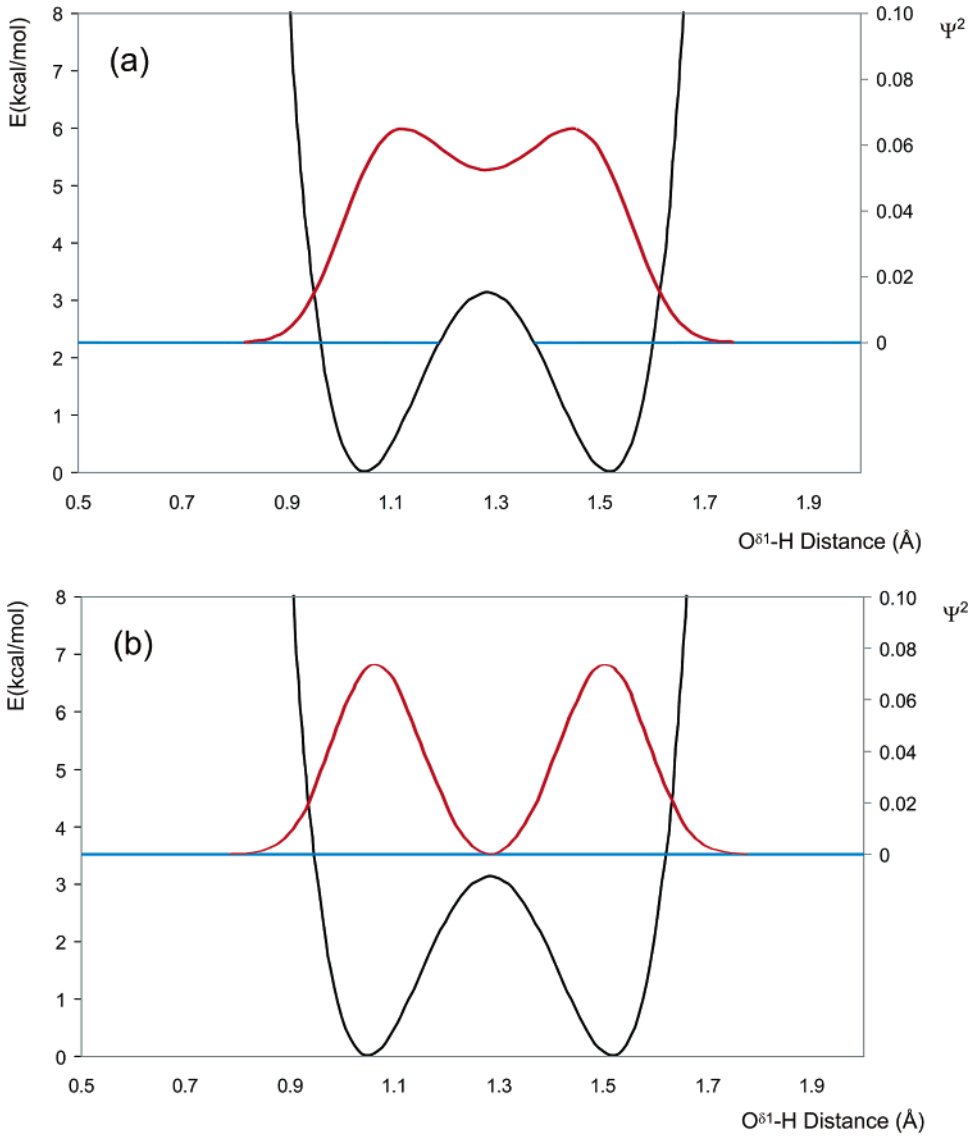


***Figure 5.*** Hydrogen. The energy values (horizontal lines) and the probability distribution functions (in red) of the (a) ground state and (b) first excited state.

points. The geometries of these points are built using the *Z* matrix difference scheme described in section B. In this case, $x_{min} = 0.768\ 38$ Å, $x_{max} = 1.802\ 54$ Å, and $\Delta x = 0.047$ Å.

For comparison, $<r_{OH}>$ averaged over 23 points is only $3.06 \times 10^{-4}$ Å larger than $<r_{OH}>$ averaged over 100 points, which shows that the use of 23 points is amply sufficient.

**G. Spectroscopic Properties.** The hydrogen NMR chemical shift is calculated using the linear scaling technique of Rablen et al.[27]

$$\delta_H = [30.60 - 0.957<\sigma_H>] + 0.33 \quad (5)$$

To their original equation (shown in brackets), we add a correction factor of 0.33 to obtain the aqueous-phase value of the chemical shift.[28]

$$\langle\sigma_H\rangle = \sum_{i=1}^{23}[\psi(i)]^2\,\sigma_H(i)\,\Delta x \quad (6)$$

In eq 5, $<\sigma_H>$ is the expectation value for the chemical shielding and $\sigma_H$ is computed using a 162-atom model (Figure 4) that comprises roughly all atoms within 7 Å of the active site. The B3LYP/6-311++G(d,p) level of theory is used for the atoms in the QM region, and STO-3G is used for the rest of the atoms.[8] The first excited state for hydrogen is thermally accessible at 298.15 K, yielding a Boltzmann's population ratio $p_1/p_0$ slightly bigger than 0.1. It has a small but non-negligible contribution to the chemical shielding. Hence, we thermally average the chemical shielding over the ground and first excited states with eq 7

$$\langle\sigma_H\rangle = f_0\sum_{i=1}^{23}[\psi_0(i)]^2\,\sigma_H(i)\,\Delta x + f_1\sum_{i=1}^{23}[\psi_1(i)]^2\,\sigma_H(i)\,\Delta x \quad (7)$$

where $f_0$ and $f_1$ are the population fractions in the ground and first excited states, respectively.

The isotope effects on the NMR chemical shifts, $\Delta\delta_{H-D}$ and $\Delta\delta_{H-T}$, are computed by

$$\langle\sigma_H\rangle - \langle\sigma_X\rangle \quad (8)$$

In eq 8, X = D or T. Once again, $<\sigma_X>$ is thermally averaged as in eq 7.

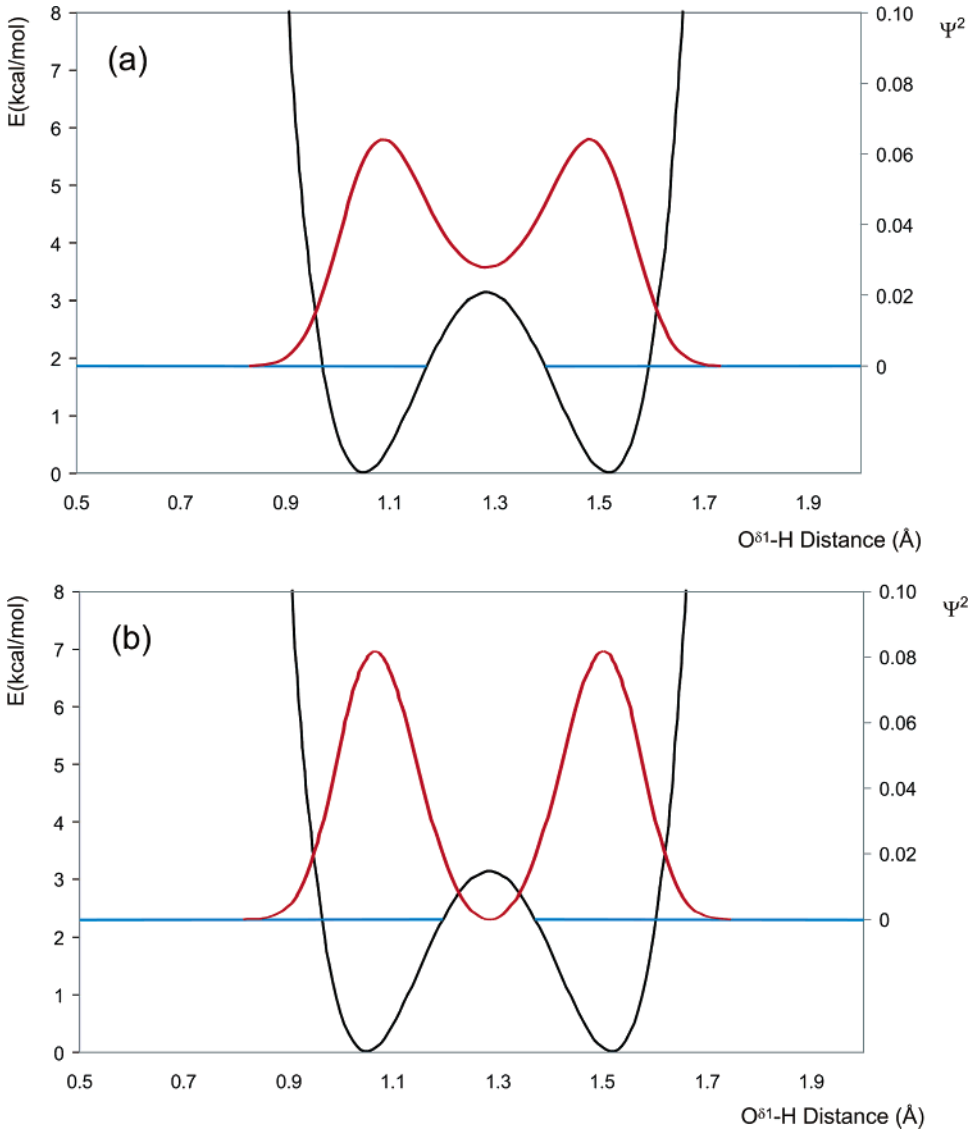


***Figure 6.*** Deuterium. The energy values (horizontal lines) and the probability distribution functions (in red) of the (a) ground state and (b) first excited state.

**H. Miscellaneous.** The GAMESS[29] program was used for all calculations, except the NMR calculations, which were performed using Gaussian.[30]

## Results and Discussion

The original quartic function $f(x)$ yields a proton distribution function that presents irregularities, in particular, a lower "hump" above the second minimum (Figure 3). The uneven quality of the wave function somewhat affects the accuracy of our expectation values. Our symmetrized potential is used to correct this problem. The new potential $V(x)$ is in very good agreement with both the quartic function and the curve obtained with the computed data points. The correlation coefficients are 0.995 and 0.987, respectively. Differences in values of calculated properties with each potential are small but not trivial. For comparison, the eigenvalue of the proton ground-state vibrational level is 2.27 kcal/mol using the symmetrized potential versus 2.33 kcal/mol with the unsymmetrized quartic function. Likewise, $<r_{OH}>$ is 1.26 Å with the quartic function versus $<r_{OH}> = 1.28$ Å with the symmetrized potential. The symmetrized potentials and the probability distribution functions for the ground-state and excited-state vibrational energies of hydrogen, deuterium, and tritium are shown in Figures 5, 6, and 7, respectively.

The symmetric double-well potential energy curve along proton transfer is consistent with the homodimeric nature of Hp. The structural properties of the HB (Figure 8) are also symmetrical. The expectation values (Table 1) reveal an equidistant water, a short $O^{\delta 1}-O^{\delta 1}$ distance of 2.46 Å, and an $O^{\delta 1}-H$ distance of 1.28 Å. The HB is quite nonlinear ($O\hat{H}O = 151°$) as also occurs in intramolecular O···H···O HBs. The angles of HBs commonly adjust to the local environment; notice also that, in aspartyl proteases, the proton is "sandwiched" between four lone pairs.[31,32]

In LBHBs, the closeness of the ground-state vibrational energy ($E_0$) of the hydrogen to the barrier has been alternatively described as "at or slightly above", "similar to", and "below" the barrier.[6,33,34] Our potential energy curve places the energy barrier at 3.14 kcal/mol, 0.87 kcal/mol above the ground-state vibrational energy of the hydrogen.

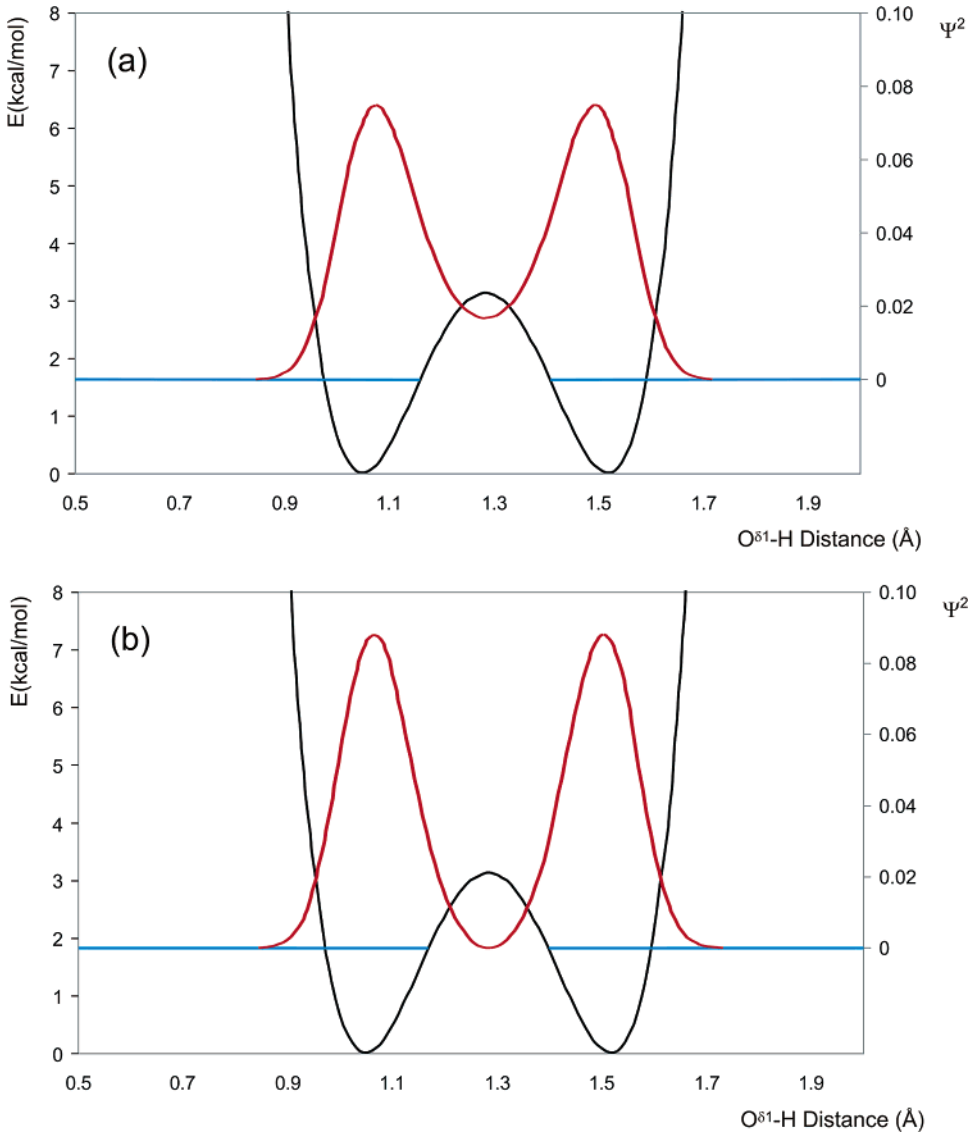


***Figure 7.*** Tritium. The energy values (horizontal lines) and the probability distribution functions (in red) of the (a) ground state and (b) first excited state.

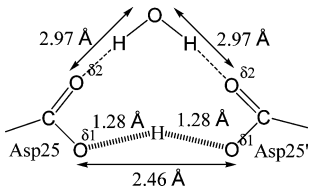


***Figure 8.*** Computed expectation values for interatomic distances in the active site of unbound HIV protease.

The proton is delocalized across a wide portion of the central region. Its diffuse position is revealed by the shape of the proton distribution function (Figure 5a). Notice the small dip in the probability that occurs at the top of the energy barrier.[34,35] The perturbation caused by the barrier has a larger effect on the probability distribution of the heavy isotopes, resulting in increasingly pronounced dips for deuterium and tritium (Figures 6 and 7). It also raises their $E_0$ values relative to the $E_0$ values of a symmetric, no-barrier potential; the ground states are 1.86 and 1.64 kcal/mol for D and T, respectively (Figure 9). In general, even states—having a large amplitude at the center of the potential (i.e., in the region of perturbation)—move up in energy, whereas odd states—with a node in the center—are much less affected (Table 2).[36]

***Table 1.*** Expectation Values Obtained in This Study[a]

| property | expectation value |
|---|---|
| $O^{\delta1}-H$ distance | 1.28 Å |
| $O^{\delta1}-O^{\delta1}$ distance | 2.46 Å |
| O(water)$-O^{\delta2}$ (Asp25) distance | 2.97 Å |
| O(water)$-O^{\delta2}$ (Asp25′) distance | 2.97 Å |
| $O^{\delta1}-H-O^{\delta1}$ angle | 151° |
| $\sigma_H$ | 14.9505 ppm |
| $\sigma_D$ | 15.4446 ppm |
| $\sigma_T$ | 15.6733 ppm |

[a] The chemical shieldings are thermally averaged over the ground state and first excited state.

This behavior is of particular importance for the ground and first excited states and needs to be considered in calculating the expectation values of the NMR chemical shift of the proton ($\delta_H$) and the NMR shifts upon isotopic substitution ($\Delta\delta_{H-D}$ and $\Delta\delta_{H-T}$). Thus, the computed chemi-

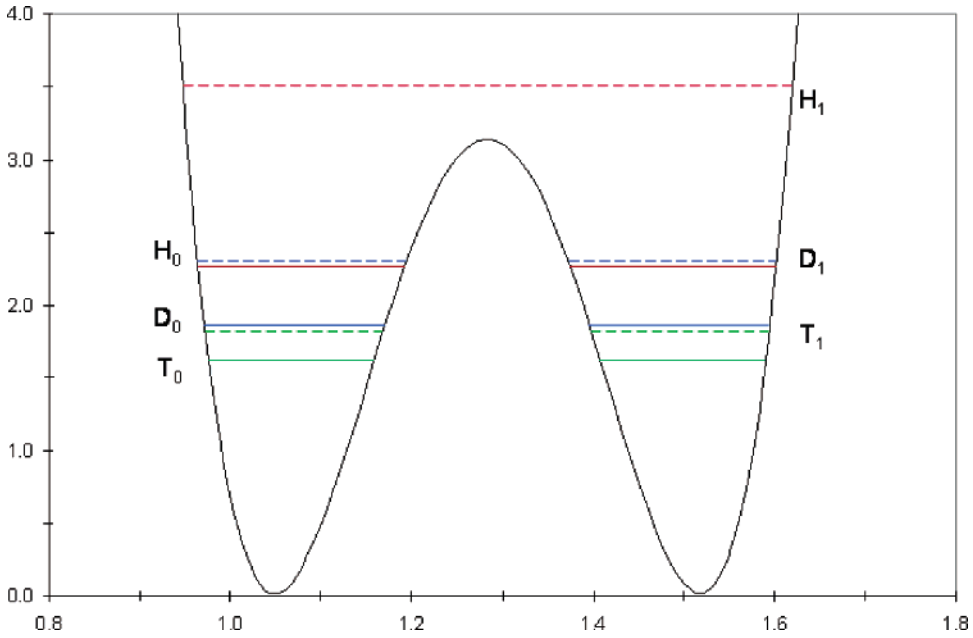


***Figure 9.*** Eigenvalues of the ground state (solid lines) and excited state (dashed lines) of each isotope.

***Table 2.*** The First Six Eigenvalues (in Units of kcal/mol) for Each Isotope

| isotope | $E_0$ | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ |
|---|---|---|---|---|---|---|
| H | 2.27 | 3.52 | 7.47 | 11.67 | 16.66 | 22.21 |
| D | 1.86 | 2.30 | 4.93 | 7.24 | 10.21 | 13.52 |
| T | 1.64 | 1.83 | 4.02 | 5.55 | 7.73 | 10.15 |

cal shifts are averaged over the ground and first vibrational excited states (which is thermally accessible at room temperature). The first excited state for the proton is at 3.51 kcal/mol, 0.37 kcal/mol above the energy barrier. The population fraction in $E_1$ at 298.15 K is 0.11. Hence, the first excited state has a small but non-negligible contribution to the computed spectroscopic properties. The calculated proton chemical shift is 16.6 ppm (versus 16.8 ppm if one neglects the first excited state). Note that our computed $\delta_H$ cannot be used as sole evidence of a LBHB because $\delta_H >$ 16 ppm can also be attributed to "strong"[8] or "very strong"[7] HBs. As previously discussed, NMR measurements of $\delta_H$ in carboxylic acids such as Hp are complicated because of rapid proton exchange with water.[18] Thus, experimentally, the absence of a downfield peak does not rule out a LBHB. A reasonable comparison between our computed $\delta_H$ and an experimental aqueous $\delta_H$ value of a system analogous to Hp is given by a study of the $(HCO_2)H^-$ complex for which $\delta_H$ = 14.1 ppm.[37] This system is somewhat similar to the aspartyl dyad, and the $r_{OO}$ distance is "expected to be ca. 2.50 Å or slightly less". An upper limit for $\delta_H$ is given by the intramolecular HB of maleic acid. The experimental proton chemical shift of this acid in dimethylsulfoxide was reported as $\delta_H = 20$ ppm.[31] Notice that our computed $\delta_H$ is not as downfield as $\delta_H$'s found in the $N^{\delta1}-H-O^{\delta1}$ HBs of serine proteases (which range from 16.9 to 18.9 ppm).[38] The proton in Hp might be somewhat shielded by the nonbonding electrons of the four surrounding oxygens. The electrostatic potential shown in Figure 10 depicts qualitatively the rich electron density areas around the carboxylic oxygens. The

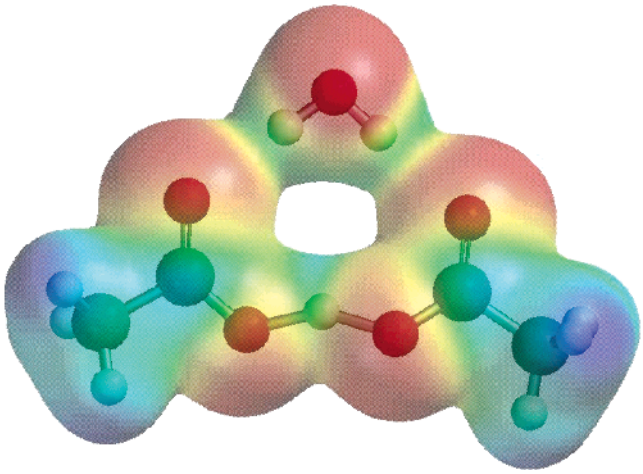

***Figure 10.*** Small model of the active site showing the electrostatic potential when the interatomic distances are those of the expectation values.

computed $\delta_H$ value can also be interpreted by examining the probability distribution function (Figure 5a). The shape of $\psi_0^2$ displays a broad region of high probability across the central region, with a dip in the center and two peaks located at about 0.05 Å from the minima. The same potential—but with a sharp peak in the center—would result in a more downfield proton chemical shift. Notice also that thermal averaging with $\psi_1^2$ (which has a node of amplitude in the center, Figure 5b) moves $\delta_H$ upfield by 0.2 ppm.

As aforesaid, the barrier causes the $E_0$ of D and T to rise. Thus, the first excited states lie fairly close to the ground states; $E_1 - E_0$ values are 0.44 and 0.19 kcal/mol for D and T, respectively, which results in first excited states having significant population fractions at 298.15 K (0.32 for D and 0.42 for T). The isotope effects on the chemical shifts are $\Delta\delta_{H-D} = 0.49$ ppm and $\Delta\delta_{H-T} = 0.72$ ppm. Because there is no experimental data for these spectroscopic properties in unbound HIV protease, we compare our calculations with

other relevant literature values. The NMR study of the $(HCO_2)H^-$ complex in water discussed above gave a $\Delta\delta_{H-D} = 0.64 \pm 0.14$ ppm. Bearing in mind the large error of this experimental value, our computed value of $\Delta\delta_{H-D} = 0.49$ for unbound Hp compares quite well. Note that our calculated $\Delta\delta_{H-T}/\Delta\delta_{H-D}$ ratio of 1.46 is in excellent agreement with the 1.44 ratio predicted by mass considerations.[39] $\Delta\delta_{H-T}$ is a very accurate parameter of HB strength because of $\delta_T$'s narrow line width in the NMR spectrum. The stronger the HB, the larger $\Delta\delta_{H-T}$ is. This parameter provides sound evidence for the presence of LBHBs in transition-state analogues of chymotrypsin. In these complexes, the experimentally measured $\Delta\delta_{H-T}$ ranges from 0.63 to 0.68 ppm.[7] Thus, our calculated $\Delta\delta_{H-T}$ of 0.72 ppm offers solid support for the presence of a LBHB in unbound Hp.

## Conclusion

We present in this work a model to characterize low-barrier double-well HBs of biological interest that should also be transferable to other HB systems of a similar nature. We apply the model to the $O^{\delta1}-H-O^{\delta1}$ HB in unbound Hp. Our study traces the entire shape of the HB's potential energy curve and yields a double-well potential with a ground-state vibrational energy that is 0.87 kcal/mol below the barrier. Our work is in agreement with Piana and Carloni's results on unbound Hp and reveals some unexpected aspects of this HB such as the presence of the first excited states of deuterium and tritium below the energy barrier. Experimentally, $\Delta\delta_{H-T}$ values can be used to gauge HB strength. Our computed $\Delta\delta_{H-T}$ of 0.72 ppm is consistent with experimental measurements in transition-state analogues of chymotrypsin and strongly points toward the presence of a LBHB in unbound Hp. Northrop has proposed to relate inhibitory efficacy in aspartyl proteases to the LBHB character of the $O^{\delta1}-H-O^{\delta1}$ HB. The driving force of this work was to develop a theoretical model to investigate such a relationship. An analysis of changes in structural and spectroscopic properties could help in the screening of drug candidate compounds. We believe that the present model can be used to study inhibitory efficacy in aspartyl proteases by evaluating how the nature of the $O^{\delta1}-H-O^{\delta1}$ HB differs between unbound and bound enzymes. Our current efforts are pointing in that direction.

**Acknowledgment.** This work was supported by the National Science Foundation—Experimental Program to Stimulate Competitive Research and the Committee on Institutional Studies and Research—Murray State University. The authors thank Dr. John Porter and Carl Woods for their aid with the potential curves. Our study was done using computational resources at Murray State University, the University of Iowa, the University of Kentucky, and the Advanced Biomedical Computing Center.

### References

(1) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.

(2) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529−25532.

(3) Perrin, C. L.; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511−544.

(4) Ash, E. L.; Sudmeier, J. L.; DeFabo, E. C.; Bachovchin, W. W. *Science* **1997**, *278*, 1128−1132.

(5) Schutz, C. N.; Warshel, A. *Proteins* **2004**, *55*, 711−23.

(6) Schiott, B. *Chem. Commun.* **2004**, 498−9.

(7) Frey, P. A. *J. Phys. Org. Chem.* **2004**, *17*, 511−520.

(8) Molina, P. A.; Jensen, J. H. *J. Phys. Chem. B* **2003**, *107*, 6226−6233.

(9) Frey, P. A. *Magn. Reson. Chem.* **2001**, *39*, S190−S198.

(10) Westler, W. M.; Frey, P. A.; Lin, J.; Wemmer, D. E.; Morimoto, H.; Williams, P. G.; Markley, J. L. *J. Chem. Am. Soc.* **2002**, *124*, 4196−4197.

(11) Coates, L.; Erskine, P. T.; Wood, S. P.; Myles, D. A.; Cooper, J. B. *Biochemistry* **2001**, *40*, 13149−57.

(12) Hong, L.; Turner, R. T., III; Koelsch, G.; Shin, D.; Ghosh, A. K.; Tang, J. *Biochemistry* **2002**, *41*, 10963−7.

(13) Ghosh, A. K.; Bilcer, G.; Harwood, C.; Kawahama, R.; Shin, D.; Hussain, K. A.; Hong, L.; Loy, J. A.; Nguyen, C.; Koelsch, G.; Ermolieff, J.; Tang, J. *J. Med. Chem.* **2001**, *44*, 2865−8.

(14) Andersson, H. O.; Fridborg, K.; Lowgren, S.; Alterman, M.; Muhlman, A.; Bjorsne, M.; Garg, N.; Kvarnstrom, I.; Schaal, W.; Classon, B.; Karlen, A.; Danielsson, U. H.; Ahlsen, G.; Nillroth, U.; Vrang, L.; Oberg, B.; Samuelsson, B.; Hallberg, A.; Unge, T. *Eur. J. Biochem.* **2003**, *270*, 1746−58.

(15) Piana, S.; Carloni, P. *Proteins* **2000**, *39*, 26−36.

(16) Piana, S.; Bucher, D.; Carloni, P.; Rothlisberger, U. *J. Phys. Chem. B* **2004**, *108*, 11139−11149.

(17) Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790−7.

(18) Trylska, J.; Grochowski, P.; McCammon, J. A. *Protein Sci.* **2004**, *13*, 513−28.

(19) Silverstein, R.; Webster, T. Proton Magnetic Resonance Spectroscopy. In *Spectroscopic Identification of Organic Compounds*, 7th ed.; Rose, N., Ed.; John Wiley: New York, 1998; p 166.

(20) Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2000**, *104*, 6656−6665.

(21) Levine, I. N. The Harmonic Oscillator. In *Quantum Chemistry*, 5th ed.; Prentice Hall: New Jersey, NJ, 2000; pp 78−79.

(22) McKeever, B. M.; Navia, M. A.; Fitzgerald, P. M.; Springer, J. P.; Leu, C. T.; Heimbach, J. C.; Herbert, W. K.; Sigal, I. S.; Darke, P. L. *J. Biol. Chem.* **1989**, *264*, 1919−21.

(23) Krause, R.; Nielsen, J. E.; Vriend, G. The WHAT IF Web Interface. http://swift.cmbi.kun.nl/WIWWWI (accessed August 7, 2006).

(24) Minikis, R. M.; Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2001**, *105*, 3829−3837.

(25) Molina, P. A.; Li, H.; Jensen, J. H. *J. Comput. Chem.* **2003**, *24*, 1971−9.

(26) Schmidt, K. The Applet. http://fermi.la.asu.edu/Schroedinger/html/node2.html#SECTION00020000000000000000 (accessed August 7, 2006)

(27) Rablen, P. R.; Pearlman, S. A.; Finkbiner, J. *J. Phys. Chem. A* **1999**, *103*, 7357−7363.

(28) Markley, J. L.; Ibanez, I. B. *Biochemistry* **1978**, *17*, 4627−40.

(29) 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. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−63.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.

(31) Hibbert, F.; Emsley, J. *Adv. Phys. Org. Chem.* **1990**, *26*, 255−379.

(32) Hao, M. H. *J. Chem. Theory Comput.* **2006**, *2*, 863−872.

(33) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−30.

(34) Belot, J. A.; Clark, J.; Cowan, J. A.; Harbison, G. H.; Kolesnikov, A. I.; Kye, Y.; Schultz, A. J.; Silvernail, C.; Zhao, X. *J. Phys. Chem. B* **2004**, *108*, 6922−6926.

(35) Benoit, M.; Marx, D. *ChemPhysChem* **2005**, *6*, 1738−1741.

(36) Noggle, J. H. Quantum Theory. In *Physical Chemistry*, 3rd ed.; Harper Collins: New York, 1996; pp 678−679.

(37) Fenn, M. D.; Spinner, E. *J. Phys. Chem. B* **1984**, *88*, 3993−3997.

(38) Harris, T. K.; Mildvan, A. S. *Proteins* **1999**, *35*, 275−282.

(39) Altman, L. J.; Laungani, D.; Gunnarsson, G.; Wennerstrom, H.; Forsen, S. *J. Am. Chem. Soc.* **1978**, *100*, 8264−8266.

CT600200S