NMR chemical shifts in the low-pH form of alpha-chymotrypsin. A QM/MM and ONIOM-NMR study

ARTICLE *in* THEORETICAL CHEMISTRY ACCOUNTS · MARCH 2003

Impact Factor: 2.23 · DOI: 10.1007/s00214-002-0415-1

3 AUTHORS, INCLUDING:

Jan Halborg Jensen
University of Copenhagen
116 PUBLICATIONS **6,597** CITATIONS

Theor Chem Acc (2003) 109:100–107
DOI 10.1007/s00214-002-0415-1

Regular article

NMR chemical shifts in the low–pH form of α-chymotrypsin. A QM/MM and ONIOM–NMR study

Pablo A. Molina, R. Steven Sikorski, Jan H. Jensen

Department of Chemistry, The University of Iowa, Iowa City, IA 52242, USA

Received: 26 September 2001 / Accepted: 6 September 2002 / Published online: 21 January 2003

Abstract. The relationship between hydrogen bonding and NMR chemical shifts in the catalytic triad of low-pH α-chymotrypsin is investigated by combined use of the effective fragment potential [(2001) *J Phys Chem A* 105:293] and ONIOM–NMR [(2000) *Chem Phys Lett* 317:589] methods. Our study shows that while the His57 $N^{\delta 1}$–H bond is stretched by a relatively modest amount (to about 1.060 Å) this lengthening, combined with the polarization due to the molecular environment, is sufficient to explain the experimentally observed chemical shifts of 18.2 ppm. Furthermore, the unusual down-field shift of $H^{\varepsilon 1}$ (9.2 ppm) observed experimentally is reproduced and shown to be induced by interactions with the C=O group of Ser214 as previously postulated. The free-energy cost of moving $H^{\delta 1}$ from His57 to Asp102 is predicted to be 5.5 kcal/mol.

Keywords: Quantum mechanics/molecular mechanics – Serine protease – Low-barrier hydrogen bonding – Nuclear magnetic resonance – Ab initio

Introduction

NMR chemical shifts provide a powerful experimental probe of interactions within protein, in particular hydrogen bonding. The study of the interactions within the catalytic triad of serine proteases provides a good example. An unusual low-field proton chemical shift ($\delta_{H\delta 1}$) of 18 ppm was originally observed by Robillard and Shulman [1] for chymotrypsin and chymotrypsinogen at pH 4, and subsequently for trypsin, trypsinogen, subtilisin, and α-lytic protease [2]. The assignment of $\delta_{H\delta 1}$ to the (N–)$H^{\delta 1}$ proton of His57 (see Fig. 1b for atom labels) was confirmed by Bachovchin [3], who measured the $^{15}N^{\delta 1}$–H spin-coupling constant in low-pH α-lytic protease.

Frey et al. [4] observed that the $\delta_{H\delta 1}$ of 18 ppm is within the 16–20-ppm range observed for low-barrier hydrogen bonds (LBHBs) in simple compounds [5], as are isotope effects on the NMR shifts ($\delta_{H\delta 1}-\delta_{D\delta 1}$) and fractionation factors measured for these enzymes. However, Ash et al. [6] noted that the $N^{\delta 1}$–H spin-coupling constant and differences in the ^{15}N chemical shifts ($\delta_{N\epsilon 2}-\delta_{N\delta 1}$) of His57 seem inconsistent with a markedly stretched $N^{\delta 1}$–H bond in low-pH α-lytic protease.

Finally, an unusual low-field chemical shift ($\delta_{H\epsilon 1}$) of 9.2 ppm has been measured for the (C–)$H^{\epsilon 1}$ proton of His57 in the low-pH forms of α-chymotrypsin [7], subtilisin BPN'97 [8], and α-lytic protease [8], in accord with earlier studies on the zymogens and suicide-inhibitor complexes [9, 10]. The deshielding of the $H^{\epsilon 1}$ proton has been ascribed to an interaction with the carbonyl oxygen of Ser214, possibly a CH···O=C HB as proposed by Derewenda et al. [11] based on the evolutionary conserved position of the C=O group. Calculations by Lin et al. [9] using an empirical NMR model [12] predicted a more shielded chemical shift on the basis of X-ray structures of chymotrypsin and chymotrypsinogen. Ash et al. [8] have proposed a "ring flip mechanism" by which the CH···O=C bond is replaced by an NH···O=C bond during the catalytic cycle of α-chymotrypsin.

Several computational studies have addressed the relationship between unusual low-field proton chemical shifts and SSHBs in model compounds[1]. Calculations by Garcia-Viloca et al. [13] on hydrogen maleate and

Contribution to the Proceedings of the Symposium on Combined QM/MM Methods at the 222nd National Meeting of the American Chemical Society, 2001

Correspondence to: J. H. Jensen
e-mail: jan-jensen@uiowa.edu

[1]The term SSHB refers to HBs with heteroatom separations of less than about 2.6 Å. LBHBs are a special case of SSHBs where the zero-point energy of the proton is near the barrier on the double-well potential and where the proton is almost or completely equally shared between the heteroatoms.

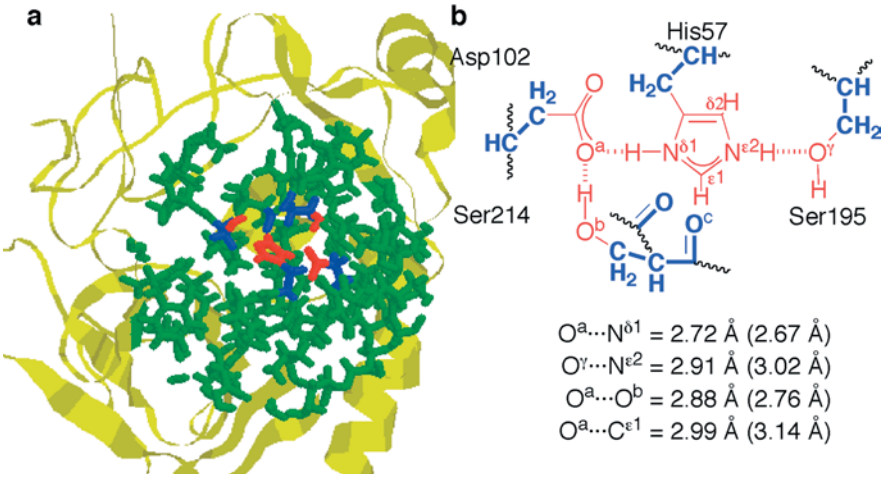


Fig. 1. a Ab initio/buffer/effective fragment potential (*EFP*) regions (*red*/*blue*/*green*) used in this study. The EFP describes the protein environment within 13 Å of the active site, while the rest (*yellow ribbon structure*) is not included in this study. **b** Schematic representation of the ab initio and buffer (*bold*) regions. Key distances for the HisH geometry are compared to values (in *parentheses*) from the X-ray geometry

related compounds supported earlier suggestions by Warshel and Papazyan [14] and Guthrie [15] that an extremely low-field proton chemical shift is not conclusive evidence for LBHB, i.e. a nearly equally shared proton. A further study by Kumar and McAllister [16] on formic acid complexed with substituted formate and enolate anions found a linear correlation between gas-phase HB strength and proton chemical shift, but the authors noted that the exact correlation depends on the nature of the system. Del Bene et al. [17] have confirmed this finding for small neutral and positively charged model compounds containing Cl–H–Cl, Cl–H–N, O–H–O, N–H–O, and N–H–N HBs.

Wei at al. [18] have investigated the effect of hydrogen bonding to carboxylate groups on the $^{15}N(-H\cdots O)$ chemical shift in histidine-containing compounds in a combined experimental/computational study. The experimental values from solid-state NMR and the computational results for imidazolium–acetate dimer at various separations were in good agreement and consistent with the $\delta_{N\delta1}$ measured by Ash et al. [6].

Finally, a recent computational study by Scheiner et al. [19] found C–H proton chemical shielding downfield shifts in the range 1.0–1.5 ppm for $F_nH_{3-n}CH\cdots H_2O$, $HOCH_3$, H_2CO.

Quantum mechanics/molecular mechanics (QM/MM) studies on subtilisin [20, 21], trypsin [22], and elastase [23] all predict that $H^{\delta1}$ is predominantly located on the histidine in the triad. Another QM/MM study [4] on the hydrogen bonding in citrate synthase reached a similar conclusion. However, these studies did not attempt to validate these findings by comparing computed and experimentally observed chemical shifts, since chemical shift predictions using QM/MM have been implemented only very recently [25].

In this paper we investigate the relationship between hydrogen bonding in the catalytic triad of low-pH α-chymotrypsin and NMR chemical shifts of select atoms in the triad. The effective fragment potential (EFP) [26] method, a hybrid QM/MM method, is used to generate a potential-energy surface (PES) for the proton transfer between His57 and Ap102 in the low-pH form of α-chymotrypsin to find the optimum NH distance. NMR chemical shifts of the $(N-)H^{\delta1}$ proton, and the $^{15}N^{\delta1}-^{15}N^{\epsilon1}$ chemical shift difference are computed along this PES, and additional proton chemical shifts on the His57 are computed at the optimum NH distance, using the recently developed ONIOM–NMR approach [27].

The paper is organized as follows. First, the EFP and ONIOM–NMR methods used in this study are described. Second, the dependence of the $\delta_{H\delta1}$ and $\delta_{N\epsilon2}-\delta_{N\delta1}$ chemical shifts on the $N^{\delta1}-H$ distance are investigated, using a comparatively small model of the protein, and the origin of the deshielding of $H^{\delta1}$ is discussed. Third, the chemical shifts of the hydrogens and $\delta_{N\epsilon2}-\delta_{N\delta1}$ of His57 computed for a larger model of the protein are presented and compared to experiment. In addition, the origin of the deshielding of $H^{\epsilon1}$ is discussed. Fourth, the energy cost associated with proton transfer from His57 to Asp102 is predicted. Finally, the results are summarized and future directions are discussed.

Computational methodology

Computational model of α-chymotrypsin

The crystal structure of resting α-chymotrypsin dimer has been determined with 1.67-Å resolution by Blevins and Tulinski [28] and was obtained from the Protein Data Bank (entry 5CHA). Hydrogen atoms were added to one of the monomers (the other was deleted) with the CHARMm program [29] and their positions were optimized, while the coordinates of the heavy atoms were fixed.

Our computational model of α-chymotrypsin consists of three parts (Fig. 1a):

1. The catalytic triad of α-chymotrypsin (Asp102, His57, Ser195) plus Ser214 is treated at the MP2/6-31 + G(d,p)//RHF/6-31G(d) level of theory.
2. All residues within 13 Å of the His57 C^γ are treated using an EFP, consisting of a distributed multipole expansion (charges through octupoles at all atomic centers and bond midpoints)

and dipole polarizability tensors for each valence localized molecular orbital (LMO). The multipole expansion is obtained using Stone's distributed multipole analysis [30], while the LMO polarizabilities are calculated using a perturbative approach due to Minikis et al.[31]. Thirty five residues or roughly 600 atoms were used to simulate the molecular environment of the active site (roughly 14% of the total number of residues in the enzyme). The 13-Å sphere was divided into five spatially distinct fragments consisting of the following residues: 192–197, 212–216, 227–229, 94–103, 41–44_52–60. The EFP parameters for the first three fragments are obtained by three separate RHF/6-31G(d) single-point calculations. The last two fragments are too big for single-point calculations so the EFP parameters for each of these fragments are obtained by a divide-and-conquer approach [31]. EFP parameters for fragment 94–103 are obtained from ab initio calculations on two overlapping subfragments 93–98 and 99–104, where the overlap occurs at the peptide bond between Thr98 and Ile99 (Fig. 2). Similarly, EFP parameters for fragment 41–44_54–60 were obtained from two subfragments joined by an overlapping disulfide bridge between Cys42 and Cys58 (Fig. 3). As long as all monopoles are scaled to ensure integer charge, this divide-and-conquer approach is accurate to within 0.2 kcal/mol [11].

3. The ab initio region is separated from the protein EFP by a buffer region [32] composed of frozen LMOs corresponding to the C^{α}–C^{β} bonds of Asp102, His57, Ser195, and Ser214 and the associated CH and core LMOs, as well as the two neighboring backbone C=O groups. Previous work [32] has shown that placing the buffer region at the C^{α}–C^{β} bond yields proton affinities within 0.5 kcal/mol of the all-ab initio reference value for the tripeptide glycyl–lysyl–glycine. The C=O buffers are needed to describe short-range interactions with His57. The buffer LMOs are generated by an RHF/6-31G(d) calculation on a subset of the system (shown in Fig. 4), projected onto the buffer atom basis functions [33], and subsequently frozen in the EFP calculations by setting select off-diagonal MO Fock matrix elements to zero [34, 35]. The ab initio/buffer region interactions are calculated ab initio, and thus include short-range interactions.

Geometry optimizations and Hessians

The geometry of the ab initio region is optimized using RHF/6-31G(d), using two different starting geometries: one in which His57 is doubly protonated; the other where the aspartic acid/histidine proton is transferred to Asp102. Two different stationary points (HisH and AspH) were found, and were verified as minima by numerically calculating the Hessian using the method proposed by Head [36], and verifying the absence of imaginary frequencies upon diagonalization. Briefly, in Head's method only a subset of the atoms (in our case the atoms in the ab initio region) are displaced during a numerical Hessian calculation, to calculate a "partial Hessian". Head showed that diagonalization of this partial Hessian leads to vibrational frequencies for surface adsorbates that compared well with experimental values. Further studies by Li and Jensen [37] have shown that vibrational energy and entropy changes for proton abstraction reactions calculated using frequencies obtained in this manner are within 0.2 kcal/mol of conventional values.

Linear least motion path

A linear least motion path (LLMP, Fig. 6) connecting HisH and AspH was constructed as follows:

1. **Z** matrices with identical connectivity were constructed for HisH and AspH, and were used to calculate the changes in bond lengths, bond angles, and dihedral angles on going from HisH to AspH.
2. Five intermediate points were generated by adding one sixth, one third, one half, two thirds, and five sixths of each change in internal coordinate to the corresponding HisH value. Three additional points were generated by scaling the coordinates by the factors necessary to increase the $N^{\delta 1}$–H distance to 1.060 and 1.090 Å and to reduce it to 0.999 Å.

MP2 single-point energies

MP2/6-31+G(d,p) single-point energies were evaluated at all points on the LLMP. The MP2 energy was calculated based on RHF orbitals computed using 6-31+G(d,p) for the ab initio region in the presence of the 6-31G(d) buffer and EFP. Excitations from the core MOs in the ab initio region and all buffer MOs were neglected. A previous study [32] has shown that MP2 corrections to deprotonation energies in Gly–Lys–Gly compare well to all-MP2 values.

NMR chemical shifts

NMR chemical shifts were calculated with the GIAO approach [38]. $\delta_{H\delta 1}$ and $\delta_{N\epsilon 2}-\delta_{N\delta 1}$ were calculated along the LLMP (Fig. 6) by RHF/6-31+G(d,p)//RHF/6-31G(d) calculations on a system that includes all atoms within a roughly 5-Å radius of the His57 C^{γ} (Fig. 4). This level of theory reproduces the measured [27] (N–)H chemical shift of 4-methylimidazolium (in $CDCl_3$) to within 0.1 ppm.

For the optimum NH distance of 1.06 Å the NMR chemical shifts were recalculated using the ONIOM–NMR method of Karadov and Morokuma [27] for calculating chemical shifts in large molecules. Here we used a two-layered approach (using the notation from Ref. [27]),

$$\sigma^2[\mathrm{ONIOM2(HF/L:HF/S)}] = \sigma^2\,[\mathrm{HF/S}, 7\text{Å}] + \sigma\,[\mathrm{HF/L}, 5\text{Å}] - \sigma^2[\mathrm{HF/S}, 5\text{Å}], \quad (1)$$

where σ[HF/L, 5 Å] denotes the calculations described in the previous paragraph, while σ^2 [HF/S, 5 Å] and σ2[HF/S, 7 Å] denote RHF/STO3G chemical shift calculations on the 5-Å system and a

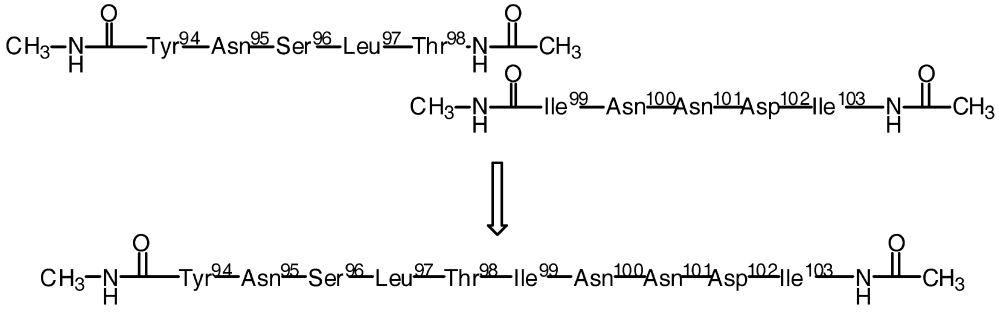


Fig. 2. Schematic representation of the EFP construction for fragment 94–103, by combining EFP parameters generated for subfragments 94–98 and 99–103, excluding common parameters in the region of overlap

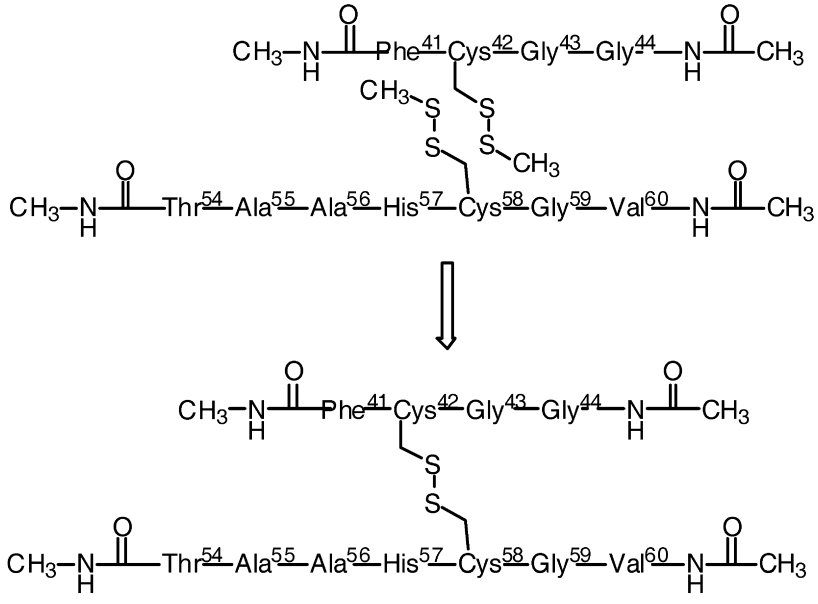


Fig. 3. Schematic representation of the EFP construction for fragment 41–44_54–60

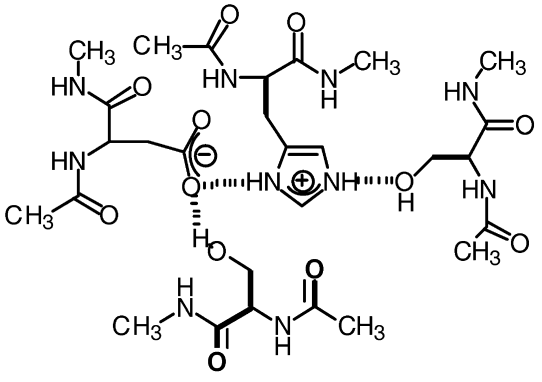


Fig. 4. Subsystem of α -chymotrypsin used to obtain the buffer region (*bold*) used in this study. The system, corresponding to the protein environment within roughly 5 Å of the active site, is also used for all– ab initio NMR calculations

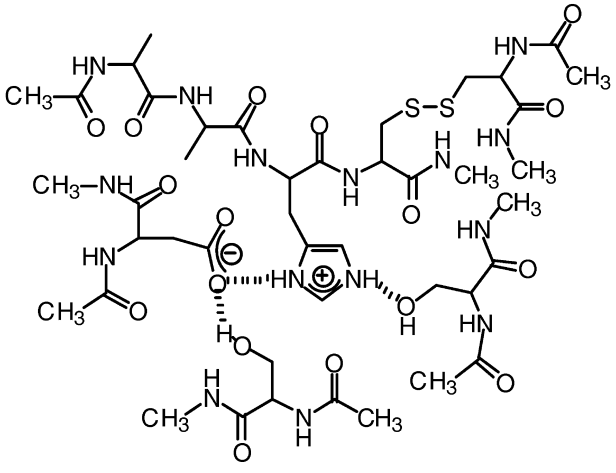


Fig. 5. MP2/6-31 + G(d,p)//RHF/6-31G(d) linear least motion path connecting the HisH and AspH minima, plus an extrapolated point corresponding to an $N^{\delta 1}$–H distance calculated for 4-methylimidazolium ion. RHF/6-31 + G(d,p)//RHF/6-31G(d) NMR chemical shift of the aspartic acid/histidine proton (δ_H) and the difference in ^{15}N chemical shift of His57 ($\delta_{N\varepsilon 2}-\delta_{N\delta 1}$) for each point

similar 7-Å system (Fig. 5), respectively[2]. The use of an STO-3G layer to model chemical shifts in large systems has been used previously by Mennucci at al. [40] to reproduce experimental solvent-induced shifts on chemical shielding tensors.

Miscellaneous

The Foster–Boys procedure was used to generate localized orbitals. The core orbitals were included in the orbital localization. The GAMESS program [42] was used for all calculations, except the NMR calculations, which were performed using GAUSSIAN98 [43].

[2]The σ^2 [HF/S, 7 Å]–σ^2 [HF/S, 5 Å] term evaluates the change in the cusp values of the active site density induced by atoms more than 5 Å away.

Results and discussion

$\delta_{H\delta 1}$ and $\delta_{N\varepsilon 2}-\delta_{N\delta 1}$ chemical shift along the LLMP

RHF/6-31G(d) geometry optimization of the ab initio region results in two minima that differ mainly in the position of the aspartic acid/histidine proton. The minimum with lowest energy (HisH, Fig. 1b) has an $N^{\delta 1}$–H bond length of 1.030 Å, only slightly elongated compared to the 0.999 Å RHF/6-31G(d) optimized geometry of 4-methylimidazolium ion. Other key distances are in good agreement with the X-ray structure (Fig. 1b). Given the 1.67-Å resolution of the X-ray geometry,

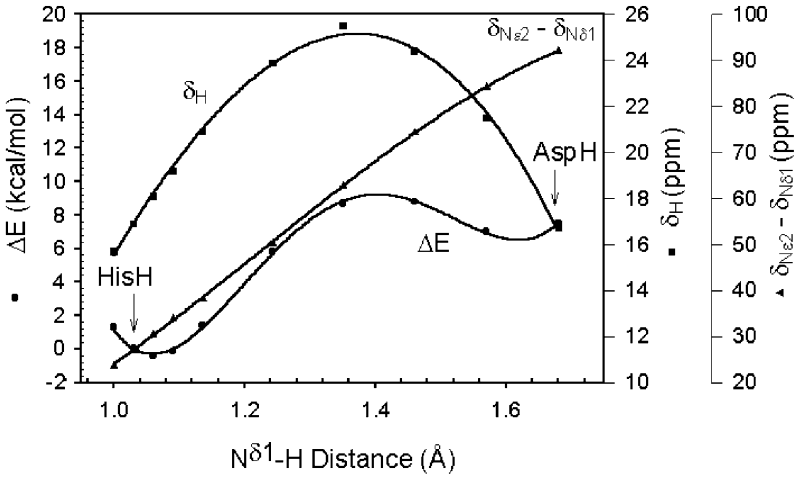


Fig. 6. Subsystem of α-chymotrypsin corresponding to the protein environment within roughly 7 Å of the active site

the uncertainty in the measured distances is roughly 0.17–0.51 Å (0.1–0.3 times the resolution [44]) so the computed distances are well within experimental error. The structure of substilisin has been determined at 0.78-Å resolution [45], where the O^a–$N^{\delta 1}$ distance is 2.62 Å, 0.1 Å shorter than the distance computed here. However, this difference is still within the experimental uncertainty in the distance of 0.08–0.23 Å.

In the other minimum (AspH) the aspartic acid/histidine proton is bonded to an Asp102 oxygen by an O–H distance of 1.004 Å.

An MP2/6-31 + G(d,p)//RHF/6-31G(d) LLMP connecting the two minima (Fig. 6) shows an asymmetric double-well potential, with a "barrier" to AspH of about 8.8 kcal/mol. It is apparent from the position of the energy minima on the LLMP that RHF/6-31G(d) underestimates the $N^{\delta 1}$–H and O^a–H bond lengths somewhat: better estimates are 1.060 and 1.120 Å, respectively. The former bond length is significantly shorter than the 1.2 Å $N^{\delta 1}$–H distance measured for substilisin at 0.78-Å resolution, but is consistent with a 2.2-Å neutron diffraction structure of trypsin [46], which indicated that a deuterium shared by aspartic acid and histidine is associated solely with histidine. However, in the former study the overall protonation state of histidine could not be conclusively determined, while in the latter study the lower zero-point energy of the N–D bond could have a significant effect on the bond length.

The chemical shift of the aspartic acid/histidine proton along the LLMP was computed and the results, displayed in Fig. 6, show that the 1.060 Å $N^{\delta 1}$–H bond results in a $\delta_{H\delta 1}$ of 18.1 ppm, in good agreement with the most recent experimental value of 18.2 ppm [7].

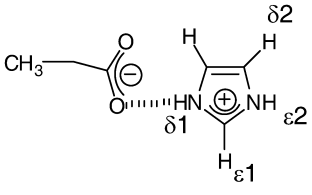
The source of the 18.1-ppm downfield shift of the aspartic acid/histidine proton compared to, for example, the corresponding 4-methylimidazolium ion value of 9.4 ppm (computed for the gas phase) is twofold:

1. About 2.4 ppm is due to the 0.06 Å $N^{\delta 1}$–H bond lengthening since an extrapolation of the LLMP to a point where the $N^{\delta 1}$–H distance is that of 4–methylimidazolium ion (the first point in Fig. 6) decreases $\delta_{H\delta 1}$ to 15.7 ppm.
2. The remaining 6.2 ppm must, therefore, be due to polarization of the $N^{\delta 1}$–H bond by Asp102 and the rest of the protein environment. Removing all but the acetate group (Scheme 1) leaves $\delta_{H\delta 1}$ unchanged, so most of the shift is predicted to be induced by polarization due to the negative charge on the carboxylate group. This is consistent with the results of Ash et al. [6] for *cis*-urocanic acid.

We note that a lengthening of the $N^{\delta 1}$–H to the value of 1.2 Å observed for substilisin is predicted to increase $\delta_{H\delta 1}$ to about 22 ppm, which differs significantly from the experimental value.

It is noteworthy that the computed $\delta_{H\delta 1}$ for AspH is also in the > 16-ppm range associated with LBHBs. To further distinguish the two geometries, we calculated the difference in the ^{15}N chemical shift, $\delta_{N\epsilon 2}$–$\delta_{N\delta 1}$, at each point on the LLMP. The data, shown in Fig. 6, indicate that $\delta_{N\epsilon 2}$–$\delta_{N\delta 1}$ is a roughly linearly (R^2 = 0.9961) increasing function of the $N^{\delta 1}$–H distance. At an $N^{\delta 1}$–H distance of 1.060 Å, $\delta_{N\epsilon 2}$–$\delta_{N\delta 1}$ = 30.5 ppm, which compares reasonably well with the experimentally observed value of 12.4 ppm for low-pH α-lytic protease [3] (a similar value for low-pH α-chymotrypsin is not known). The computed value of $\delta_{N\epsilon 2}$–$\delta_{N\delta 1}$ for AspH (92.0 ppm) is more consistent with the absolute value for singly protonated His57 (61.4 ppm) measured for high-pH α-lytic protease [3].

The discrepancy between the calculated and experimental value of $\delta_{N\epsilon 2}$–$\delta_{N\delta 1}$ is more likely due to the chemical differences between α-lytic protease and



Scheme 1.

α-chymotrypsin. For example, the $\delta_{H\delta1}$ of low-pH α-lytic protease is 17 ppm, indicating that the $N^{\delta1}$–H bond may be shorter, thereby decreasing $\delta_{N\epsilon2}-\delta_{N\delta1}$. For low-pH α-chymotrypsin, $\delta_{H\delta1}=16.9$ ppm corresponds to an $N^{\delta1}$–H length of 1.03 Å and a decrease in $\delta_{N\epsilon2}-\delta_{N\delta1}$ of 3.0 ppm. Furthermore, long-range interactions can change $\delta_{N\epsilon2}-\delta_{N\delta1}$ by several parts per million, as discussed in the following subsection.

Chemical shifts at the optimum $N^{\delta1}$–H distance

The chemical shifts calculated using σ[HF/L, 5 Å] and Eq (1) are summarized in Fig. 7 together with pertinent experimental data. Comparison of the 5-Å and 7-Å results shows that longer-range interactions decrease $\delta_{H\delta1}$ and $\delta_{N\epsilon2}-\delta_{N\delta1}$ by 0.2 and 3.5 ppm, respectively. Thus, $\delta_{H\delta1}$ and $\delta_{N\epsilon2}-\delta_{N\delta1}$ remain in good agreement with experiment. $\delta_{H\epsilon2}$,$\delta_{H\delta2}$, and $\delta_{H\epsilon1}$ are within 2.3, 0.4, and 0.1 ppm of the experimental values. The poor prediction of $\delta_{H\epsilon2}$ is likely due to the neglect of solvation effects since that part of the active site is solvent-exposed, and additional hydrogen bonding should deshield the proton further.

The large deshielding predicted for $H^{\epsilon1}$ is especially interesting since it has been implicated in a possible HB with the C=O group of Ser214. In our computational model short-range interactions between C–H and this group, as well as the C=O group of Val213, had to be treated using frozen orbitals in order to avoid geometrical collapse during geometry optimizations. Here we present further support of a CH···O=C interaction by changing the Ser214 C=O group to a methylene group (Scheme 2) and recomputing a $\delta_{H\epsilon1}$ (using σ[HF/L, 5 Å]) of 8.2 ppm.

For comparison, the $\delta_{H\epsilon1}$ of 4-methylimidazolium ion is 8.6 ppm. Thus, our calculations support the experimental interpretation of this unusual shift.

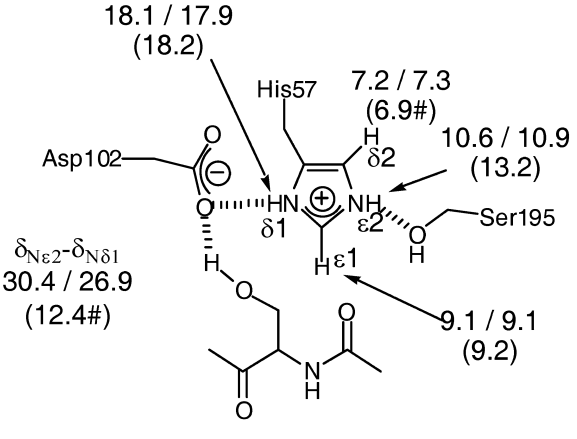


Fig. 7. Calculated, σ2[HF/L, 5 Å]/Eq (1), and experimental (in *parentheses*) NMR chemical shifts in parts per million for low-pH chymotrypsin [7] or α-lytic protease (Refs. [6, 8], marked by a *hash*, for $\delta_{N\epsilon2}-\delta_{N\delta1}$ and $\delta_{H\delta2}$ respectively)

Relative energy of AspH and HisH

The change in energy due to proton transfer from His57 to Asp102 has been discussed previously in the literature (see later), and is pursued in further detail here.

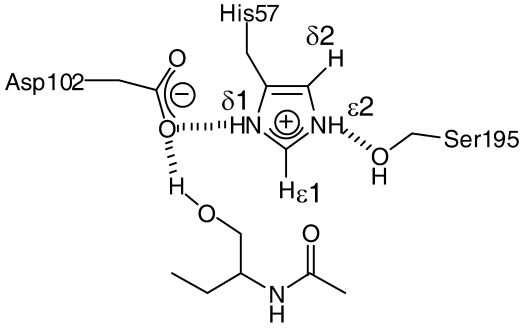
The HisH geometry is 7.5 kcal/mol lower in energy (ΔE) than the AspH configuration at the MP2/6-31+G(d,p)//RHF/6-31G(d) level of theory. Addition of the rigid rotor/harmonic oscillator ΔG^{298} correction [37] yields a free-energy difference of 7.1 kcal/mol. The difference in energy between the two low-energy points on the MP2 surface (Fig. 5) is 7.4 kcal/mol, very similar to the 7.5 kcal/mol ΔE using the RHF/6-31G(d) optimized geometry.

Neglecting polarization of the EFP region reduces ΔE to 1.0 kcal/mol, and this term is thus crucial for an accurate ΔE. Analysis of the polarizability contributions indicates that the HisH and AspH geometries are stabilized by 15.6 and 10.0 kcal/mol, respectively. Neglect of the octupoles introduces an error of only 0.3 kcal/mol, which demonstrates that the static multipolar representation of the molecular environment is converged.

EFP and all–ab initio calculations can be combined in an ONIOM-like approach to get a better estimate of ΔE,

$$\Delta G\ [\mathrm{MP2:AI}] = \Delta G\ [MP2:EFP, 13\ \text{Å}] + \Delta E\ [HF, 7\ \text{Å}] - \Delta \mathrm{E[HF:EFP, 7\ Å]} \quad (2)$$

Here, ΔG [MP2:EFP, 13 Å] is the 7.1 kcal/mol value obtained by the EFP method as discussed previously. ΔE [HF, 7 Å] denotes an all– ab initio RHF/6-31G(d) ΔE calculated using the 7-Å system shown in Fig. 5 (this value is 3.8 kcal/mol). Finally, ΔE [HF:EFP, 7 Å] denotes a 7-Å EFP calculation with ab initio and buffer regions identical to those used in the 13-Å calculations, except that the ab initio region is treated at the RHF/6-31G(d) level for the energy calculations (this value is 5.4 kcal/mol). Combining these values via Eq (2) leads to a new prediction for the free-energy difference between AspH and HisH of 5.5 kcal/mol (=7.1+3.8–5.4). For comparison, we note that a PM3/AMBER study of elastase predicted a gas-phase ΔH of 5.4 kcal/mol [23]. The introduction of solvent effects is expected to increase the asymmetry further. Indeed, Warshel [21]



Scheme 2.

has predicted a ΔG^{298} of 12 kcal/mol for solvated subtilisin based on a protein dipoles–Langevin dipoles simulation.

Though the lengthening of the $N^{\delta 1}$–H bond is smaller than that in a LBHB, it does result in a roughly 2-kcal/mol decrease in energy (Fig. 6). A similar decrease of the activation energy would have a significant effect on the rate of peptide hydrolysis.

Conclusions and future directions

We investigated the relationship between hydrogen bonding and NMR chemical shifts in the catalytic triad of low-pH α-chymotrypsin by combined use of the EFP and ONIOM–NMR methods. The NMR shift of the $H^{\delta 1}$ ($\delta_{H\delta 1}$, Fig. 1b) is shown to have a parabolic dependence on the N–$H^{\delta 1}$ distance, while the difference in chemical shift of $N^{\delta 1}$ and $N^{\epsilon 2}$ ($\delta_{N\epsilon 2}-\delta_{N\delta 1}$) is linear (Fig. 6).

The optimum N–$H^{\delta 1}$ distance is determined to be 1.06 Å, and the NMR shifts of the His57 ^{1}H and ^{15}N atoms using this geometry are in good agreement with experimental values (Fig. 7). In particular, the large downfield shift of $\delta_{H\delta 1}$ of 18.2 ppm is reproduced to within 0.3 ppm, and is shown to be predominantly due to electronic polarization. Furthermore, the unusual downfield shift of $\delta_{H\epsilon 1}$ observed experimentally is reproduced and is shown to be induced by interactions with the C = O group of Ser214 as previously postulated.

The free-energy cost of moving the $H^{\delta 1}$ from His57 to Asp102 is predicted to be about 5.5 kcal/mol using a combination of EFP and all–ab initio models of the enzyme. This asymmetric double-well potential and the modest 0.06-Å lengthening of the N–$H^{\delta 1}$ are both inconsistent with the LBHB hypothesis (though not the SSHB hypothesis). However, the lengthening of the $N^{\delta 1}$–H bond does result in a roughly 2-kcal/mol decrease in energy. A similar decrease of the activation energy would have a significant effect on the rate of peptide hydrolysis.

The agreement between theory and experiment is encouraging for this complex system, and several studies necessary for a more complete understanding of hydrogen bonding in serine proteases are planned. For example, two other important experimental probes of hydrogen bonding within proteins, deuterium isotope effects on proton chemical shifts and fractionation factors, will be investigated. Preliminary calculations indicate that it will be necessary to perform geometry optimizations at the density functional theory level for accurate results. Furthermore, other serine proteases, such as substilisin, α-lytic protease, and elastase, will be studied in their resting state and complexed with suicide inhibitors and natural substrates. Finally, the effect of solvation and molecular dynamics must be addressed for a full understanding of these very interesting enzymes.

Acknowledgements. This work was supported by a Research Innovation Award from the Research Corporation and a type G starter grant from the Petroleum Research Fund. The calculations were performed on IBM RS/6000 workstations obtained through a CRIF grant from the NSF (CHE-9974502) and on supercomputers at the National Center for Supercomputer Applications at Urbana-Champaign. The authors are indebted to Visvaldas Kairys for help with the CHARMM program, and to Daniel Quinn for many helpful discussions.

References

1. Robillard G, Shulman RG (1972) J Mol Biol 71:507
2. Robillard G, Shulman RG (1978) J Mol Biol 86:519
3. Bachovchin WW (1985) Proc Natl Acad Sci USA 82:7948
4. Frey PA, Whitt SA, Tobin JB (1994) Science 264:1927
5. Hibbert F, Emsley J (1990) Adv Phys Org Chem. 26:255
6. Ash EL, Sudmeier JL, De Fabo EC, Bachovchin WW (1997) Science 278:1128
7. Bao D, Huskey WP, Kettner CA, Jordan F (1999) J Am Chem Soc 121:4684
8. Ash EL, Sudmeier JL, Day R.M, Vincent M, Torchilin EV, Haddad KC, Bradshaw EM, Sanfrod DG, Bachovchin WW (2000) Proc Natl Acad Sci USA 97:10371
9. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA (1998) Proc Natl Acad Sci USA 95:14664
10. (a) Porcuban MA, Neves DE, Rausch SK, Markley JL (1978) Biochemistry 17:4640; (b) Markley JL, Ibanez IB (1978) Biochemistry 17:4627
11. Derewenda ZS, Derewenda U, Kobos PM (1994) J Mol Biol 241:83
12. Williamson MP, Asakura T (1993) J Magn Reson Ser B 10:63
13. Garcia-Viloca M, Gelabert R, Gonzalez-Lafont A, Moreno M, Lluch JM (1997) J Phys Chem A 101:8727
14. Warshel A, Papazyan A (1996) Proc Natl Acad Sci USA 93:13665
15. Guthrie JP (1996) Chem Biol 3:163
16. Kumar GA, McAllister MA (1998) J Org Chem 63:6968
17. Del Bene JE, Perera SA, Bartlett RJ (1999) J Phys Chem A 103:8121
18. Wei Y, de Dios A, McDermott AE (1999) J Am Chem Soc 121:10389
19. Scheiner Sl, Gu Y, Kar T (2000) J Mol Struct (THEOCHEM) 500:441
20. Benzien J, Muller RP, Florian J, Warshel A (1998) J Phys Chem B. 102:2293
21. Warshel A (1991) Computer modeling of chemical reactions in enzymes and proteins. Wiley, New York
22. (a) Perakyla M, Kollman PA (2000) J Am Chem Soc 122:3436; (b) Stanton RV, Perakyla M, Bakowies D, Kollman PA (1998) J Am Chem Soc 120:3448
23. Monard G, Loos M, Thery V, Baka K, Rivail J-L (1996) Int J Quantum Chem 58:153
24. Mulholland AJ, Lyne PD, Karplus M (2000) J Am Chem Soc 122:534
25. Cui Q, Karplus M (2000) J Phys Chem 104:3721
26. (a) Gordon MS, Freitag MA, Bandyopadhyay P, Jensen JH, Kairys V, Stevens WJ (2001) J Phys Chem A 105:293; (b) (b) Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Kraus M, Garmer D, Basch H, Cohen D (1996) J Chem Phys 105:1968
27. Karadakov PB, Morokuma K (2000) Chem Phys Lett 317:589
28. Blevins RA, Tulinski A (1985) J Biol Chem 260:4264
29. Molecular Simulations Inc (1998) QUANTA/CHARMm. Molecular Simulations, San Diego, CA
30. Stone AJ (1981) Chem Phys Lett 83:233
31. Minikis RM, Kairys V, Jensen JH (2001) J Phys Chem A 105:3829
32. Kairys V, Jensen JH (2000) J Phys Chem A 104:6656
33. King HF, Stanton RE, King H, Wyatt RE, Parr RG (1967) J Chem Phys 47:1936
34. Stevens WJ, Fink WH (1987) Chem Phys Lett 139:15
35. Bagus PS, Hermann K, Bauschlicher CW Jr (1984) J Chem Phys 80:4378
36. Head JD (1997) Int J Quantum Chem 65:827

37. Li H, Jensen JH (2002) Theor Chem Acc 107:211
38. (a) McWeeny R (1962) Phys Rev 126:1028; (b) Ditchfield R (1974) Mol Phys 27:789; (c) Dodds JL, McWeeny R, Sadlej AJ (1980) Mol Phys 41:1419; (d) Wolinski K, Hilton JF, Pulay P (1990) J Am Chem Soc 112:8251
39. Pouchert CS (ed) (1993) The Aldrich library of NMR spectra, 2nd edn, vol 2. Aldrich Chemical Company
40. Menucci B, Martinez JM, Tomasi J (2001) J Phys Chem A 105:7287
41. (a) Boys SF (1966) In: Lowdin PO (ed) Quantum science of atoms, molecules and solids. Academic, New York, pp ;(b) Edmiston C, Ruedenberg K (1963) Rev Mod Phys 35:457
42. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
43. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T,. Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98, revision A.6. Gaussian, Pittsburgh, PA
44. Lipscomb WN (1980) In: Darnall DW, Wilins RG (eds) Methods for determining metal ion environments in proteins. Elsevier, New York, pp
45. Kuhn P, Knapp M, Soltis SM, Ganshaw G, Thoene M, Bott R (1998) Biochemstry 37:13446
46. Kossiakoff AA, Spencer SA (1981) Biochemistry 20:6462