Published in final edited form as:

Chem Phys. 2013 August 30; 422: 156–164. doi:10.1016/j.chemphys.2013.03.006.

Proton affinity of the histidine-tryptophan cluster motif from the influenza A virus from *ab initio* molecular dynamics

Arindam Bankura, Michael L. Klein, and Vincenzo Carnevale

Institute for Computational Molecular Science, Temple University, Philadelphia, Pennsylvania 19122, United States

Vincenzo Carnevale: vincenzo.carnevale@temple.edu

Abstract

Ab initio molecular dynamics calculations have been used to compare and contrast the deprotonation reaction of a histidine residue in aqueous solution with the situation arising in a histidine-tryptophan cluster. The latter is used as a model of the proton storage unit present in the pore of the M2 proton conducting ion channel. We compute potentials of mean force for the dissociation of a proton from the $N\delta$ and $N\epsilon$ positions of the imidazole group to estimate the pK_a 's. Anticipating our results, we will see that the estimated pK_a for the first protonation event of the M2 channel is in good agreement with experimental estimates. Surprisingly, despite the fact that the histidine is partially desolvated in the M2 channel, the affinity for protons is similar to that of a histidine in aqueous solution. Importantly, the electrostatic environment provided by the indoles is responsible for the stabilization of the charged imidazolium.

1 Introduction

The M2 protein from the influenza A virus forms a pH-dependent channel, which selectively allows permeation of protons across the viral envelope [1–5]. Several experimental structures of M2 have shown that the transmembrane (TM) domain is a left-handed four alpha-helix bundle with a well-defined mostly hydrophobic lumen and that two conserved residues, His-37 and Trp-41, physically occlude the C-terminal end of the pore [6–20]. Sequential protonation events of the four His-37 residues are believed to be at the heart of activation and selectivity due to the ability of histidines to exchange protons with the water molecules solvating the channel lumen. The five-membered ring of the histidine side chain carries two nitrogen atoms ($N\delta$ and $N\epsilon$) and only one of them is protonated in the neutral imidazole form, which is usually stable at pH 7 or higher. At acidic pH both nitrogens can bind protons and the ring is preferentially found in the cationic imidazolium form. The Trp-41 residue, which is displaced along the peptide backbone from His-37 by one helical turn, has also been found to be crucial for channel function [17–22]. The interaction between His-37 and Trp-41, potentially affected by pH, has been suggested to play a role in

© 20XX Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

preventing backflow of protons and other cations from the interior of the viral particle to the external medium. Consistently, in the high-pH conformation of the channel, the four Trp-41 indoles form a tight hydrophobic seal preventing permeation of waters and ions.

Solid-state NMR measurements have shown that the first two histidines of the tetrad are protonated with $pK_a \sim 8.2$; the remaining two pK_a s are 6.3 and <5 [8]. Remarkably, the pK_a of the imidazole sidechain of L-histidine in aqueous solution is approximately 6.0. The series of M2 pK_a s values strongly suggest that the channel provides an environment able to increase the affinity of the imidazole moieties for protons. Given the partial desolvation of the histidine residues, the positive charge of imidazolium is not completely screened by the dipolar field of H-bonded water molecules, therefore a different mechanism for the energetic stabilization of the charged species has to be envisioned. In this respect, several hypotheses have been proposed that invoke: strong (low-barrier) H-bonds between neutral and cationic histidines, transfer of electronic density from the electron-rich indole moiety to the charged imidazolium, cation- π interactions between the benzenoid group of the tryptophan and imidazolium, and effective smearing of the positive charge by statistical delocalization of the excess proton over the four histidines and the surrounding water molecules.

The aim of the present study is to investigate the presence and, possibly, the relevance of some of these molecular properties with the ultimate goal of identifying the molecular determinants of such a large proton affinity. Accordingly, we have explored the protonation/deprotonation equilibrium by performing calculations on a model system, namely a solvated cluster of histidine and tryptophan residues, mimicking the histidine environment in the M2 channel. By comparing the results for the cluster with the case of a single histidine in aqueous solution, we were able to dissect the energetic stabilization of the positively charged state and highlight the most significant contributions.

Given its relevance as an experimental observable quantity, we focus on the acid dissociation constant, pK_a , to analyze and compare the proton dissociation equilibrium in different molecular contexts. The pK_a value of an acid in aqueous solution is directly related to the standard free energy change of the reversible proton dissociation reaction. Several methods for computing the free energy change of an acid dissociation have been used in the literature. The majority of these rely on the use of thermodynamic cycles [23–30], where the dissociation free energy is computed from the sum of the deprotonation free energy of the molecule in the gas phase and the difference between the solvation free energies of the dissociated and undissociated molecule. In the latter scheme, the electronic structure of the solute molecule is explicitly considered, while the interactions with the solvent molecules are modeled by a dielectric continuum medium. Thus, this approach provides an accurate description of the chemical deprotonation step, however the short-range intermolecular interactions with solvent molecules are only approximately taken into account. In contrast, first-principles molecular dynamics (MD) simulations can describe the processes of bond breaking and formation without a priori assumptions. Therefore, density functional theory (DFT) based *ab initio* MD simulations are in principle able to provide a detailed mechanistic picture of the dissociation process while retaining a satisfactory description of solvation the microscopic level.

To this end, we have performed *ab initio* MD simulations in the so-called Blue Moon ensemble method [31] to compute the potential of mean force (PMF) for the process of proton dissociation. pK_a values were estimated via integration of the PMF over a suitable domain using water autodissociation as a reference reaction. Accordingly, we performed simulations on three classes of systems: (i) solvated histidine-tryptophan (His-Trp) structural motif, (ii) histidine residue in aqueous solution, and (iii) pure water (Figure 1).

In order to assess the accuracy of our methods, we performed several additional simulations. Firstly, we investigated the tautomeric equilibrium of histidine [32–34] in aqueous solution. The two non-equivalent nitrogen, $N\delta$ and the $N\epsilon$, are expected to feature slightly different values of pK_a , thereby providing a stringent check of accuracy. Then, we considered the effects of system size and concentration on the estimated pK_a 's by performing simulations on systems comprising two different numbers of waters. Finally, we qualitatively characterized the impact of dispersion interactions on the proton dissociation equilibrium by introducing an empirical correction to the generalized gradient approximation used in the present DFT calculations.

Anticipating our results, we find that the intrinsic pK_a of the His-Trp cluster motif is 6.5, suggesting that the imidazole group fully retains its proton affinity in this partially desolvated environment. Moreover, when the stabilization of the protein-membrane environment is taken into account, the pK_a is shifted to 8.1, a value in remarkable agreement with two recent independent experimental estimates. Seeking an explanation for the unexpected proton affinity of the His-Trp cluster motif we analyzed the electronic structure and found that no significant transfer of charge density occurs from the indole moiety to imidazolium. Instead, the electrostatic potential generated by the four indole groups is seemingly able to interact in a favorable way with the charged imidazolium, thereby pointing to cation- π interactions as the likely source of stabilization for the excess proton.

2 Computational Methods

2.1 *Ab initio* Molecular Dynamics

All the simulations presented herein are based on the Born-Oppenheimer molecular dynamics implementation of the DFT, specifically within the CP2K [35]/QUICKSTEP [36] dynamics simulation package. CP2K uses a computationally efficient scheme based on a hybrid Gaussian basis set for the wave functions with an auxiliary plane wave (PW) basis set for the density. Kohn-Sham [37] orbitals were represented using the triple- ζ doubly polarized (TZV2P) basis set, and the electron density is expanded in plane waves with a 280 Ry cutoff. We have used the NN10 smoothing method implemented in QUICKSTEP. We used Goedecker-Teter-Hutter [38, 39](GTH) norm-conserving pseudopotentials. In all simulations the Becke-Lee-Yang-Parr (BLYP) exchange and correlation functional [40] was used. The BLYP approximation is one of the most widely used exchange and correlation (XC) functionals to describe liquid water. The Born-Oppenheimer MD simulations were performed with an integration time step of 0.5 fs. The canonical sampling via velocity rescaling (CSVR) thermostat [41] is used to keep the system at the target temperature of 330 K with a 100 fs time constant. This value of temperature was chosen to avoid the glassy behavior [42] of liquid water observed with this exchange and correlation functional at

lower temperatures; this seemingly high value of temperature is thus expected to approximate the structural and dynamical properties of liquid water at room temperature. The starting configurations for all the *ab initio* MD simulations presented here were obtained from preliminary classical simulations using the CHARMM force-field. Constrained MD simulations are performed taking a single N-H distance as reaction coordinate. One of the systems consist of a single protonated L-histidine molecule solvated by 49 water molecules, in a periodically replicated cubic supercell with fixed dimensions of $11.74 \times 11.74 \times 11.74 \text{ \AA}^3$. The volume and number of water molecules in the box provide the standard state of an ideal 1.0 M solution. Simulations have been carried out for both the N δ and the N ϵ positions of the imidazole.

To account for the effect of dispersion interactions, we have also carried out MD simulations with the addition of the semi-empirical DFT-D3 correction to DFT calculated values of energy and forces [43]. A further four independent constrained MD simulations were carried out for a larger system using a standard cubic box of size $16.44 \times 16.44 \times 16.44 \text{ \AA}^3$ containing 142 waters and one L-histidine molecule in a +1 state, which results in a 0.37 M solution. To calculate the pK_a values of the histidine-tryptophan (His-Trp) cluster motif, we considered a system comprised of four histidine (one of which is charged), four tryptophan and twelve glycine residues. The resulting cluster is solvated by 183 water molecules in a $21.0 \times 17.5 \times 21.0 \text{ \AA}^3$ box. The atoms of the peptide main-chain groups were kept restrained to the initial configuration by applying harmonic potentials.

An additional simulation was performed to investigate a proton transfer reaction involving two water molecules in aqueous environment. Indeed, a direct estimation of the pK_a value from integration of the PMF can introduce large systematic errors. To solve this issue, we have used water autodissociation as a reference reaction, by performing constrained MD simulations on an orthorhombic box of edges $L = 9.86 \text{ \AA}$ containing 64 water molecules. For the water autodissociation reaction, we have chosen as the reaction coordinate the H-O distance of an arbitrary water molecule. The computational details for all of the simulations are reported in Table 1.

For each system, we considered twenty points along the reaction coordinate in the range 0.80–1.75 $\text{\AA}$. For each of the twenty windows, constrained MD simulations were carried out for 13–15 ps while keeping the reaction coordinate ξ fixed. The initial 3 ps-long segment of trajectory of each simulation window were not considered in the computation of the mean constraint force. For each system, a trajectory spanning 0.25–0.3 ns was used to obtain a converged PMF. Overall, for this study more than 3 ns of *ab initio* simulations were performed.

2.2 Calculation of Potentials of Mean Force

The deprotonation of weak acids is clearly a rare event on the time scale of *ab initio* MD, as the reaction path involves overcoming relatively large free energy barriers. To investigate the protonation/deprotonation equilibrium two alternative approaches are usually adopted: either free energy perturbation (FEP) methods [44–47] or the constrained MD simulation protocol [48–55]. In these approaches, both the solute and the solvent are treated at the same quantum mechanical DFT level of theory, and the deprotonation free energy can be obtained

from a now standard statistical mechanical formalism. In the former method, DFT MD simulations are carried out to sample a mapping potential consisting of a linear mixing of the Born-Oppenheimer energy surfaces of dissociated and undissociated states. Free energy changes are obtained by integrating ensemble averages of vertical energy gaps along the alchemical transformation path from a dissociated and undissociated state. In the latter method, free energies can be calculated via constrained MD simulations [48–55] based on a DFT description of the electronic degrees of freedom. Free energies profiles for the deprotonation event are estimated by the thermodynamic integration of the mean force along a pathway entailing the transfer of a proton from the solvated acid to a neighboring water molecule. This method is virtually parameter-free and all relevant interactions between the solute and the solvent molecules are properly included. Using this method, the pK_a value and free energy differences can be calculated with reasonable accuracy [54, 55].

In the case of constrained MD simulation, the reaction coordinates $\xi = \xi(\mathbf{R}_I)$, defined by a given subset of atomic coordinates $\mathbf{R}_I$, is assumed to be a distance between two selected atoms. The value of the reaction coordinate is kept constant and the relative free energies (ΔF) are calculated by integrating the average constraining force ($\langle f_\xi \rangle$) via the thermodynamic integration relation:

$$\Delta F = \int_{\xi_0}^{\xi_1} \langle f_{\xi'} \rangle d\xi', \quad (1)$$

where ξ_0 and ξ_1 indicate the reactant and product state, respectively. The reaction coordinate (ξ) represents the reaction progress and can be expressed as an analytical function of the Cartesian coordinate ($\mathbf{R}_I$). For the reaction coordinate fixed at ξ' , the mean force has the following form:

$$\langle f_{\xi'} \rangle = - \frac{\langle Z^{-1/2} [\lambda - K_B T G] \rangle_{\xi'}}{\langle Z^{-1/2} \rangle_{\xi'}}, \quad (2)$$

where the bracket indicates the time-average over the MD simulation by keeping ξ to the desired value, λ is the Lagrange multiplier constant, K_B is the Boltzmann constant, T is the temperature in Kelvin, and Z and G are weight and correction factors. For a one dimensional linear distance constraint, the second equation reduces to a simple ensemble average of the Lagrange multiplier λ : $\langle f_{\xi'} \rangle = \langle \lambda \rangle_{\xi'}$.

We have used constrained MD by selecting a reaction coordinate ξ and constrained its value in a range suitable to study a proton transfer event (Figure 2 and Figure 1). In particular, the N*-H* distance between one of the imidazole nitrogen and its neighboring hydrogen is varied between 0.80 and 1.75 Å with an increment of 0.05 Å. The N atom and the H atom that are constrained are labeled N* and H*, respectively, and the oxygen atom in contact with H* is denoted as O'. For the water autodissociation reaction, we have chosen as reaction coordinate the H-O distance of a randomly selected water molecule and constrained its value between 0.80 and 1.75 Å with 0.05 Å increments.

2.3 Calculation of pK_a 's

The pK_a value for a weak acid in aqueous solution is determined from the free energy profiles of a proton dissociation reaction. The pK_a value for weak acid in aqueous solution can be estimated using the approach described by Chandler [56]. The value of Helmholtz free energy for a proton transfer reaction is directly related to the dissociation constant as:

$$K_c^{-1} = C_0 \int_0^\infty H(r) e^{-\beta \omega(r)} 4\pi r^2 dr; \quad H(r) = \begin{cases} 1 & r < R_c \\ 0 & r \geq R_c \end{cases}, \quad (3)$$

where K_c is the equilibrium constant for dissociation, C_0 is the standard concentration, $H(r)$ is a Heaviside function, $w(r) = F(r)$ is the PMF, and R_c is a distance cutoff used to distinguish between the covalently bonded reactant and dissociated product. Using equation 3, we cannot directly estimate the pK_a value from the energy difference of the PMF. Firstly, equation 3 requires a complete knowledge of $w(r)$, which is the reversible work needed for bringing the proton from an infinite distance to r . Clearly, given the size of the simulation box, we can calculate $w(r)$ only for small finite separations between the reactant and products. Secondly, equation 3 ignores the entropic contribution due to the diffusion of the solvent-bound excess proton throughout the solution. Specifically, since we apply a constraint only to the N-H bond length, the excess proton is allowed to diffuse away from the water molecule H-bonded to the nitrogen atom. At large values of the reaction coordinate, *i.e.*, when the proton is already dissociated from the imidazolium, the probability for the excess proton of being located on the proximal water molecule depends on the number of water molecules competing for proton binding and, therefore, on the total number of water molecules present in the system. This results into a missing volume-dependent term in the expression for K_a . It is worth mentioning that with a different choice of the reaction coordinate [57], the center of excess charge can be localized, thereby avoiding the aforementioned ambiguity in the position of the proton. The combination of these effects can lead to an error in the pK_a value of several units. To reduce the magnitude of these systematic errors, we have adapted the procedure used by Ivanov *et al.* [55]. As an alternative to the calculation of absolute pK_a 's, we have evaluated relative pK_a 's with respect to a deprotonation reaction for which the experimental pK_a is known. The relative pK_a values are calculated using water autodissociation as a reference reaction. By taking the ratio of dissociation constants for the reaction of weak acid and the autodissociation reference reaction, according to eq 3 we obtain:

$$\frac{K_d(His)}{K_d(H_2O)} = \frac{\int_0^{R_c} e^{-\beta \omega_{H_2O}(r)} r^2 dr}{\int_0^{R_c} e^{-\beta \omega_{His}(r)} r^2 dr}. \quad (4)$$

We have also taken into account the macroscopic pK_a value correction due to the contribution of several indistinguishable microscopic protonation states [27]. If the protonated species has n indistinguishable microscopic states and its unprotonated counterpart has m indistinguishable states, then the microscopic pK_a corresponding to the equilibrium between microscopic protonated and unprotonated species can be related to the macroscopic pK_a through the formula:

$$pK_A^{Macro} = pK_A^{Micro} + \log \frac{n}{m} \quad (5)$$

Importantly, we have estimated the effect of transferring our model His-Trp cluster from water to the M2 protein environment. Accordingly, the pK_a value in M2 protein is calculated using the free energy cycle [58–60] shown below.

Here HA and A^- represent the His-Trp cluster in the protonated and deprotonated state, respectively. The free energy change for the deprotonation in protein ($\Delta_a G_{HA}$) is determined as:

$$\Delta_a G_{HA} = \Delta_a G_{HA, model} + \Delta_{xfer} G_{A^-} - \Delta_{xfer} G_{HA},$$

where $\Delta_a G_{HA, model}$ is the free energy change for the deprotonation of the His-Trp cluster in aqueous solution (calculated from our constrained *ab initio* MD simulation) and $\Delta_{xfer} G_{A^-} - \Delta_{xfer} G_{HA}$ is the transfer free energy. We have evaluated transfer free energies and pK_a shift associated with this transfer using the APBS program package [61]. The transfer free energy ($\Delta_{xfer} G$), is the free energy changes for transferring the His-Trp cluster in the ionized (A^-) and neutral (HA) forms from solution to the protein environment. The transfer free energies are estimated from electrostatic energies, which are calculated from the numerical solution of the Poisson-Boltzmann equation using a continuous homogeneous dielectric medium to represent the aqueous environment. The transfer free energy ($\Delta_{xfer} G$) is calculated as:

$$\Delta_{xfer} G = \Delta_{xfer} G_{A^-} - \Delta_{xfer} G_{HA};$$

$$\Delta_{xfer} G_X = G_{\{protein\ with\ charged\ X\}} - G_{\{protein\ with\ uncharged\ X\}} - G_{\{charged\ X\ in\ solution\}},$$

where X can be A^- or HA, $G_{\{protein\ with\ charged\ X\}}$ is the electrostatic energy of the protein with group X bound and all charges on X set to their normal values, $G_{\{protein\ with\ uncharged\ X\}}$ is the electrostatic energy of the protein with group X bound and all charges on X set to zero, and $G_{\{charged\ X\ in\ solution\}}$ is the electrostatic energy of group X in solution with all charges set to their normal values. The pK_a shift (ΔpK_A^{shift}) associated with this transfer are computed from the transfer free energies ($\Delta_{xfer} G$) as $\Delta pK_A^{shift} = \Delta_{xfer} G / 2.303RT$. The pK_a inside the protein ($pK_A^{protein}$) is estimated based on the computed pK_a of the corresponding model compound (pK_A^{model}) in aqueous solution by adding the pK_a shift (ΔpK_A^{shift}) induced by transferring the model compound into the protein environment:

$$pK_A^{protein} = pK_A^{model} + \Delta pK_A^{shift}.$$

2.4 Quantum Chemical Calculations

We studied the energetics of the proton transfer process in isolated model systems using the BLYP functional with a TZVP basis set (GAUSSIAN 03 package [62]). Potential energy surfaces (PES) were computed for clusters containing an imidazolium cation and either three or six water molecules. We have also computed the intermolecular interaction energies between an indole group and a cluster containing an imidazolium cation and five water molecules by performing first a full geometry optimization.

3 Results and Discussion

To gain structural insights into the dissociation process of histidine, we have computed ensemble-averaged structural properties for different values of the distance constraint. The average distances of the N* and H* with its nearest neighbor O' as a function of the reaction coordinate (ξ) are shown in Figure 3. The deprotonation process occurs through three distinct phases: (i) stretching of the N*-H* covalent bond; (ii) proton transfer from N*-H* to the nearest neighbor solvating (H-bond acceptor) water molecule; and (iii) diffusion of the solvent-bound excess proton through the solution.

At small values of the reaction coordinate, H* is covalently bound to the N* and the average H*-O' distance is comparable to those typical of a H-bond. The nearest neighbor water molecule is mostly found to be four-fold coordinated, *i.e.*, it donates two H-bonds and accepts two H-bonds. At intermediate values of the reaction coordinate, the H*-O' distance smoothly decreases until it reaches the average value of an O-H bond in a hydronium ion. The bond transformation (conversion of N*-H* from chemical to hydrogen bond) takes place for values of ξ in the range of 1.2 to 1.35 Å. Simultaneously, the neighbor water of H* (O') starts to loose one of the accepted H-bonds and it becomes three-fold coordinated.

At $\xi = 1.3$ Å, the N*-H* distance coincides with the O'-H* one, the proton is shared between the imidazole N* atom and water O' atom. This behavior persists up to $\xi = 1.35$ Å, when the covalent bond between N* and H* transforms into a hydrogen bond. The average N*-O' distance gradually decreases up to this point. At $\xi = 1.4$ Å, the H* has completely been transferred to the O' water molecule, which becomes a three-fold coordinated hydronium (H₃O⁺) ion. The constrained bond at this point a hydrogen bond instead of an N*-H* covalent bond. Any further increases in the value of ξ correspond to stretching a hydrogen bond and is therefore accompanied by smaller energetic effects.

At $\xi = 1.5$ Å, proton transfer events are observed from the hydronium ion to the neighboring water molecules. The proton, at this point dissociated from the original hydronium moiety, can diffuse through a Grotthuss-type process. Up to $\xi = 1.6$ Å, these events are only transient, *i.e.*, the excess proton is mostly localized on the neighbor water molecule of H* forming a three-coordinated Eigen [63] cation (H₉O₄⁺). In some of the configurations, the proton is equally shared by two water molecules in the so-called Zundel [64] configuration (H₅O₂⁺). From $\xi = 1.65$ –1.75, full proton diffusion events are observed in which protons hop through the hydrogen bond network via the Grotthuss mechanism inducing a structural rearrangement of the network [65–68]. At this point, deprotonation is complete and the hydronium ion eventually diffuses through the solution. Consistently, the sign of the average

constraint force changes thereby providing an unambiguous definition of the transition state for the deprotonation process.

We have calculated Mulliken and Löwdin atomic charges and monitored their evolution as a function of the constrained reaction coordinate during the dissociation process.

Configurations have been selected randomly from the each constrained trajectory. In Figure 4, the Mulliken and Löwdin populations of N* and H* are presented as a function of the reaction coordinate. As expected, as the value of the reaction coordinate is increased, the charge on N* becomes more negative due to the increased electron population and the charge on H* becomes more positive as the electron population decreases. Concerning the total charge of the imidazole as a function of reaction coordinate, we note that for small values the positive charge is mostly localized on the imidazole and is progressively transferred to the solvent molecules as the deprotonation proceeds.

The free energy profiles for the proton dissociation reaction and water autodissociation arising from the thermodynamic integration of the mean force are shown in (Figure 5). The free energy differences obtained for the deprotonation reactions are summarized in Table 2. The PMFs for deprotonation from the N δ and the N ϵ positions of L-histidine at 1.03 M concentration, based on using the GGA BLYP exchange and correlation functional are shown in upper-left panel of Figure 5. We have obtained a free energy difference (ΔF) of 9.3 and 10.2 kcal/mol, respectively. A previous study by Ivanov *et al.* [54] reported free energy differences of 9.8 and 8.7 kcal/mol for the deprotonation of L-histidine from the N δ and N ϵ positions, respectively. The Free energy profile for the proton dissociation reaction from the N δ positions at the two different concentration (1.03 M and 0.37 M) are shown in the upper-right panel of Figure 5. The free energy profiles at the two different concentrations are very similar.

In the lower-left panel of Figure 5, we have compared the free energy profile obtained by including the empirical dispersion correction to that obtained using the simple GGA approach. Despite providing qualitatively similar PMFs, the two schemes (simple GGA and DFT-D3) lead to different values of the estimated free energy changes (Table 2). Simulation with the Grimme-D3 (BLYP-D3) correction yield the value of 10.0 kcal/mol for the deprotonation at 0.37 M of L-histidine from the N δ positions whereas BLYP leads to a predicted value of 9.6 kcal/mol. The calculated free energy changes for the proton dissociation reaction of histidine are thus found to be quantitatively different: the Grimme-D3 correction provides a significant stabilization of the reactant (about 0.4–0.7 kcal/mol, see also Table 2).

On the basis of our free energy estimation (Table 2), we can conclude that the histidine cation is preferentially deprotonated from the N δ position in aqueous solution. Consistently, pH-dependent ^{13}C chemical shift [69] measurements have shown that the N ϵ tautomeric form of the imidazole ring is preferred over the N δ tautomer with 4:1 ratio in the zwitterionic form of histidine. Furthermore, in recent Solid-State NMR experiments [34], only the N ϵ tautomeric form of histidine was observed near pH 6. The N δ tautomer was shown to be undetectable also in NMR studies of small histidine compounds at less basic pH conditions [33, 34]. These experiments clearly indicate that deprotonation occurs

preferentially from the N δ position of the histidine cation. In both these cases, the barrier for the proton transfer reaction is either very small or absent; therefore we can conclude that the kinetics of the reverse reaction is mostly diffusion limited.

The estimation of the pK_a values of weak acids from the potential of mean force depends critically on the cutoff radius (R_c) used to discriminate between the covalently bound and the dissociated H* atom. In the approach proposed by Davies *et al.* [52], the parameter R_c is adjusted to reproduce the experimentally determined dissociation constant ($pK_w = 14$ and) of water. The pK_a of a weak acid is subsequently estimated from the potential of mean force using this R_c cutoff value ($R_c = 1.22$ Å when the BLYP functional is used). It is worth noting that a change of cutoff value (from 1.22 to 1.40 Å) shifts the estimated pK_a value by 2–3 pK_a units. Therefore in this approach the computed pK_a carry a crucial dependence on R_c , thereby making the choice of this parameter a critical step in the calculation. The most reliable approach adopted to date is that of Ivanov *et al.* [55] in which the dependence of the calculated relative pK_a values on the chosen value of R_c is small. For instance, changing the value of R_c from 1.3 to 1.5 Å shifts the value of pK_a by less than 0.1 units.

In view of the above discussion, we have adopted the approach of Ivanov *et al.* [55] to calculate the pK_a values in our systems, *i.e.* by integrating the PMF according to eq 4 and setting the cutoff radius R_c to 1.35 Å. The resulting calculated pK_a values are reported on Table 2. From the GGA BLYP calculation, we obtain pK_a estimates of 7.6 (7.5) and 8.3 (7.7) for the deprotonation of L-histidine at a concentration of 1.03 M (0.37 M) from the N δ and the N ϵ positions, respectively. Simulations with the Grimme-D3 correction provide the values 7.3 (6.8) and 7.7 (7.1) for the same systems. Very recent experimental estimates based on solid-state NMR spectroscopy and standard electrochemical titration give a pK_a value of 6.1 [32–34].

The theoretically determined pK_a value is ~1 unit higher than that of the experimentally measured value. In our relative pK_a calculation method, we have used the ratio of dissociation constants for the reaction under consideration and a reference reaction, thereby approximately removing the error associated with the aforementioned missing entropic contribution. However, the contributions from the zero point energies associated with histidine deprotonation and possible quantum tunneling of the proton are neglected. Ivanov *et al.* [55] have shown that zero point energy contributions are able to decrease the estimated pK_a . The small number of solvent molecules in the simulation box is also likely to have an impact on the computed pK_a , although in our case the simulation box is large enough to include both the first and second solvation shells of the histidine and therefore the response of the solvent in the vicinity of the acidic group is expected to be correctly modeled. Inaccuracy of the DFT functionals and the statistical uncertainty associated with thermodynamic integration are also likely to generate larger errors in the computed pK_a values.

In the His-Trp cluster, the four imidazoles are in close contact. Four equivalent sites can accept a proton, thus all the histidines have the same affinity. We focused on deprotonation from the N δ positions of the imidazoles that are exposed to the solvent. In simulations using the GGA-BLYP scheme (without dispersion corrections), the excess proton is observed to

diffuse from one imidazole group to another through the water molecules via Grotthuss-type structural diffusion process. Such proton transfer events occur within a time-scale of 1–2 ps, a time-interval too short to get a converged value of the mean constraint force. Remarkably, this behavior was not observed when the dispersion correction is introduced.

The structural properties (Figure 3) and the free energy profile (lower-right panel of the Figure 5) obtained for the deprotonation of His-Trp cluster are found to be qualitatively similar to those of histidine deprotonation. For small values of the reaction coordinate, the excess proton is bound to the N δ atom of imidazole and the first solvation shell water molecule is four-fold coordinated. Hu *et al.* [8] have proposed that the excess proton is shared in imidazole-imidazolium dimer possibly through a low-barrier H-bond. We do not observe any imidazole-imidazolium dimer formation events, possibly as a result of the chosen initial configuration. At large values of the reaction coordinate, although a gradual decrease of the mean constraint force is observed, it never becomes negative (lower-right panel of the Figure 5). The excess proton (δH) is found to be preferentially associated with the first solvation shell water molecules forming an Eigen' or a Zundel cation. Due to the high proton affinity of the cluster, the hydronium ion does not diffuse in solution, however the excess-proton is observed to diffuse towards another imidazole group through the solvating water molecules.

We have calculated the ensemble averages Mulliken and Löwdin charges for the constrained MD simulations. In Figure 6, the Mulliken and Löwdin populations of N* and H* are presented as a function of the reaction coordinate. The total charge of the imidazole (the one initially protonated) and indole (the one closer to the initially protonated imidazole) are also shown. As expected in the undissociated state, the positive charge is mostly localized on the imidazole, while in the dissociated state the charge is found on the neighboring solvent molecules. It is important to note that, according to this population analysis, no significant charge transfer occurs from the imidazole to the nearest indole group of the tryptophan residue.

The free energy profile associated with the deprotonation from the N δ positions of the His-Trp cluster, calculated using the BLYP functional and the dispersion correction, is shown in Figure 5. The free energy difference between the protonated and the deprotonated forms of the His-Trp cluster in aqueous solution is 9.1 kcal/mol (Table 1). Remarkably, this value is much smaller than the one relative to the L-histidine in aqueous solutions. By integrating the free energy profile applying the correction for the multiple titratable sites, we obtain a pK_a value of 6.5 [70].

To investigate the cause of the different affinity for protons between the His-Trp cluster and L-histidine, we computed the potential energy surface (PES) for the deprotonation process of a partially solvated imidazolium cation. The PES along the deprotonation reaction coordinate is shown in Figure 7. Deprotonation energies of 34.2 and 45.4 kcal/mol were found for the cluster containing one imidazolium cation and either three or six water molecules, respectively. When both the protons of imidazolium cation are partially solvated by water molecules, the protonated state is more stable by the 11.2 kcal/mol. Interestingly, replacement of one of the water molecules by an indole group results in a destabilization of

3.2 kcal/mol. Therefore, we conclude that the stability of the protonated state depends crucially on the solvation of both the protons of imidazolium.

In the His-Trp cluster, δH is fully solvated through the hydrogen bonded network of water molecules, whereas the eH is not solvent-exposed and mostly desolvated. The eH is directed towards the pore and only two water molecules are found in the cluster cavity. One of the first solvation shell water molecules accept a hydrogen bond from the eH and donates one hydrogen bond to the second water molecule through one of its hydrogen atoms. The other hydrogen is pointing toward the nearest neighbor indole of tryptophan. In the L-histidine system, both δH and eH are fully solvated by the water molecules. Therefore it is expected that the protonated state of L-histidine is much more stable than that of His-Trp cluster system. Surprisingly, we find that the protonated state of His-Trp cluster system is less stable by only 1 kcal/mol. Furthermore, our gas phase calculation shows that the presence of an indole group instead of a water causes a net destabilization of the imidazolium cation of about 3.2 kcal/mol, *i.e.*, larger than the observed 1 kcal/mol. Seeking for a source of stabilization of the imidazolium cation, we analyzed the electrostatic potential generated by the four indole groups. The electrostatic potential surfaces portrayed in Figure 8 strongly suggest that the indole groups of the histidine-tryptophan cluster provide an electrostatic environment able to stabilize the imidazolium cation, therefore we conclude that, in analogy with other biomolecular systems [71, 72], cation- π like interactions play a crucial role and are responsible for the affinity for protons of this structural motif.

The estimated pK_a shift for transferring the His-Trp cluster from solution to the M2 protein environment is 1.6 units [73]. Including this pK_a shift value, we obtain an absolute pK_a value of 8.1 in the M2 protein. Our estimated pK_a value is remarkably similar to the experimental value. Based on the solid state ^{15}N nuclear magnetic resonance (ssNMR) experiments performed on the TM domain of M2, the pK_a value for the binding of the first proton was found to be 8.2 [8]. Furthermore recently Hu *et. al* have estimated a pK_a of 7.6 based on independent ssNMR experiments on the M2 TM domain [5]. These experiments indicate that the first proton is bound to the His residues in the tetrad with surprisingly high affinity compared to histidine in solution. From our theoretical calculation, we have found that the protonated state of M2-TM is mostly stabilized by the electrostatic interactions between the indole groups and the charged imidazolium, and by the favorable dipolar field of the TM helices.

4 Conclusions

The deprotonation reaction of imidazole in different environments has been studied here via *ab initio* MD simulations using the constrained MD approach. Trajectories spanning time intervals of hundreds of ps allowed us to compute free energy profiles with an unprecedented statistical accuracy. We have analyzed a number of structural and electronic properties along MD trajectories generated at different distance constraints and computed accurate values of the pK_a of histidine. Our calculated pK_a 's are highly consistent with the experimentally estimated ones. A similar accuracy was reported by previous studies of the histidine deprotonation reaction [54, 55] and in this sense this aspect of our work is not new. However, the present investigation differs in one important aspect from the previous works

[54, 55]. In refs [54] and [55] the deprotonation reaction was studied with constrained MD simulations using the GGA-BLYP approach, whereas in the present work, in addition to the GGA description of exchange and correlation energy, dispersion interactions are also taken into account, albeit in an approximate fashion. Including the dispersion correction effect, the pK_a values are shifted to the lower side by 0.4–0.7 units with respect to the standard GGA approach.

The main focus of this work is, however, to present mechanistic details and the free energy profile of the proton dissociation activity of His-Trp cluster in the aqueous solutions. Using *ab initio* molecular dynamics simulations, we have found that the protonated state is stabilized by 9.1 kcal/mol. A pK_a value of 6.5 is found for the deprotonation of this cluster. Although histidine is partially desolvated in this environment, the protonated state is stabilized by electrostatic interaction with the indole groups of the tryptophans. By estimating the pK_a shift due to the transfer of the His-Trp cluster from water to the M2 transmembrane domain, we determine the pK_a value of 8.1 in for the first protonation event in the M2 channel, a value in remarkable agreement with recent experimental measurements [5]. It is important to mention that the first two protonation events were hypothesized to be cooperative in M2 [8]. Since we do not investigate the second protonation event, cooperativity is beyond the scope of the present investigation. Importantly, we show that this relatively high pK_a value does not require the formation of imidazole-imidazolium pairs as previously hypothesized [8], instead it results from favorable electrostatic interactions between the imidazolium moiety and the protein environment.

Acknowledgments

We thank Bill DeGrado for his input and continuing interest in this work. The research was supported by the National Institutes of Health, the National Center for Computational Sciences under grant CHM045, and through XSEDE resources provided by National Institute for Computational Science under Grant Nos. TG-MCA93S020.

References

1. Pinto LH, Holsinger LJ, Lamb RA. Cell. 1992; 69:517. [PubMed: 1374685]
2. Schnell JR, Chou JJ. Nature. 2008; 451:591. [PubMed: 18235503]
3. Pielaka RM, Schnelli JR, Choua JJ. Proc Natl Acad Sci USA. 2009; 106:7379. [PubMed: 19383794]
4. Hu F, Luo W, Hong M. Science. 2010; 330:505. [PubMed: 20966251]
5. Hu F, Schmidt-Rohr K, Hong M. J Am Chem Soc. 2012; 134:3703. [PubMed: 21974716]
6. Acharya R, Carnevale V, Fiorin G, Levine BG, Polishchuk AL, Balannik V, Samish I, Lamb RA, Pinto LH, DeGrado WF, Klein ML. Proc Natl Acad Sci USA. 2010; 107:15075. [PubMed: 20689043]
7. Wang C, Lamb RA, Pinto LH. Biophys J. 1995; 69:1363. [PubMed: 8534806]
8. Hu J, Fu R, Nishimura K, Zhang L, Zhou H-X, Busath DD, Vijayvergiya V, Cross TA. Proc Natl Acad Sci USA. 2006; 103:6865. [PubMed: 16632600]
9. Cady SD, Schmidt-Rohr K, Wang J, Soto CS, DeGrado WF, Hong M. Nature. 2010; 463:689. [PubMed: 20130653]
10. Cady SD, Mishanina TV, Hong M. J Mol Biol. 2009; 385:1127. [PubMed: 19061899]
11. Stouffer AL, Acharya R, Salom D, Levine AS, Di Costanzo L, Soto CS, Tereshko V, Nanda V, Stayrook S, DeGrado WF. Nature. 2008; 451:596. [PubMed: 18235504]
12. Pinto LH, Dieckmann GR, Gandhi CS, Papworth CG, Braman J, Shaughnessy MA, Lear JD, Lamb RA, DeGrado WF. Proc Natl Acad Sci USA. 1997; 94:11301. [PubMed: 9326604]

13. Wang J, Kim S, Kovacs F, Cross TA. *Protein Sci.* 2001; 10:2241. [PubMed: 11604531]
14. Cady SD, Goodman C, Tatko CD, DeGrado WF, Hong M. *J Am Chem Soc.* 2007; 129:5719. [PubMed: 17417850]
15. Balannik V, Carnevale V, Fiorin G, Levine BG, Lamb RA, Klein ML, DeGrado WF, Pinto LH. *Biochemistry.* 2010; 49:696. [PubMed: 20028125]
16. Carnevale V, Fiorin G, Levine BG, DeGrado WF, Klein ML. *J Phys Chem C.* 2010; 114:20856.
17. Sharma M, Yi M, Dong H, Qin H, Peterson E, Busath D, Zhou HX, Cross TA. *Science.* 2010; 330:509. [PubMed: 20966252]
18. Nishimura K, Kim S, Zhang L, Cross TA. *Biochemistry.* 2002; 41:13170. [PubMed: 12403618]
19. Okada A, Miura T, Takeuchi H. *Biochemistry.* 2001(40):6053. [PubMed: 11352741]
20. Witter R, Nozairov F, Sternberg U, Cross TA, Ulrich AS, Fu R. *J Am Chem Soc.* 2008; 130:918. [PubMed: 18163621]
21. Tang Y, Zaitseva F, Lamb RA, Pinto LH. *J Biol Chem.* 2002; 277:39880. [PubMed: 12183461]
22. Luo W, Mani R, Hong M. *J Phys Chem B.* 2007(111):10825. [PubMed: 17705425]
23. Warshel A. *Biochemistry.* 1981; 20:3167. [PubMed: 7248277]
24. Lim C, Bashford D, Karplus M. *J Phys Chem.* 1991; 95:5610.
25. Sham YY, Chu ZT, Warshel A. *J Phys Chem B.* 1997; 101:458.
26. Kamerlin SCL, Haranczyk M, Warshel A. *J Phys Chem B.* 2009; 113:1253. [PubMed: 19055405]
27. Lopez X, Schneider M, Dejaegere A, Karplus M. *J Am Chem Soc.* 2002; 124:5010. [PubMed: 11982365]
28. Lyne PD, Karplus M. *J Am Chem Soc.* 2000; 122:166.
29. Schuurmann G, Cossi M, Barone V, Tomasi J. *J Phys Chem A.* 1998; 102:6706.
30. Major DT, Laxer A, Fischer B. *J Org Chem.* 2002; 67:790. [PubMed: 11856021]
31. Carter EA, Ciccotti G, Hynes JT, Kapral R. *Chem Phys Lett.* 1989; 156:472.
32. Jones SF, Wong WYL, Gutbeii WG, Bachovchin WW. *J Am Chem Soc.* 1993; 115:6813.
33. Henry B, Tekely P, Delpuech JJ. *J Am Chem Soc.* 2002; 124:2025. [PubMed: 11866617]
34. Li S, Hong M. *J Am Chem Soc.* 2011; 133:1534. [PubMed: 21207964]
35. <http://cp2k.berlios.de/>
36. VandeVondele J, Krack M, Mohamed F, Parrinello M, Chassaing T, Hutter J. *Comp Phys Comm.* 2005; 167:103.
37. Kohn W, Sham LJ. *Phys Rev A.* 1965; 140:1133.
38. Goedecker S, Teter M, Hutter J. *Phys Rev B.* 1996; 54:1703.
39. Hartwigsen C, Goedecker S, Hutter J. *Phys Rev B.* 1998; 58:3641.
40. Becke AD. *Phys Rev A.* 1988; 38:3098. Lee, C; Yang, W.; Parr, R. G. *Phys. Rev B* **1988**, 37, 785. [PubMed: 9900728]
41. Bussi G, Donadio D, Parrinello M. *J Chem Phys.* 2007; 126:014101. [PubMed: 17212484]
42. VandeVondele J, Mohamed F, Krack M, Hutter J, Sprik M, Parrinello M. *J Chem Phys.* 2005; 122:014515.
43. Grimme S, Antony J, Ehrlich S, Krieg H. *J Chem Phys.* 2010; 132:154104. [PubMed: 20423165]
44. Sulpizi M, Sprik M. *Phys Chem Chem Phys.* 2008; 10:5238. [PubMed: 18728866]
45. Cheng J, Sulpizi M, Sprik M. *J Chem Phys.* 2009; 131:154504. [PubMed: 20568869]
46. Cheng J, Sprik M. *J Chem Theory Comput.* 2010; 6:880.
47. Mangold M, Rolland L, Costanzo F, Sprik M, Sulpizi M, Blumberger J. *J Chem Theory Comput.* 2011; 7:1951.
48. Sprik M, Ciccotti G. *J Chem Phys.* 1998; 109:7737.
49. Trout BL, Parrinello M. *Chem Phys Lett.* 1998; 288:343.
50. Trout BL, Parrinello M. *J Phys Chem B.* 1999; 103:7340.
51. Sprik M. *Chem Phys.* 2000; 258:139.
52. Davies JE, Doltsinis NL, Kirby AJ, Roussev CD, Sprik M. *J Am Chem Soc.* 2002; 124:6594. [PubMed: 12047179]

53. Doltsinis NL, Sprik M. *Phys Chem Chem Phys*. 2003; 5:2612.
54. Ivanov I, Klein ML. *J Am Chem Soc*. 2002; 124:13380. [PubMed: 12418879]
55. Ivanov I, Chen B, Raugei S, Klein ML. *J Phys Chem B*. 2006; 110:6365. [PubMed: 16553455]
56. Chandler, D. *Introduction to Modern Statistical Mechanics*. Oxford University Press; Oxford: 1987. A second generation multistate empirical valence bond model for proton transport in aqueous systems.
57. Day TJF, Soudackov AV, Cuma M, Schmitt UW, Voth GA. *J Chem Phys*. 2003; 117:5839.
58. Bashford D, Karplus M. *Biochemistry*. 1990; 29:10219. [PubMed: 2271649]
59. Antosiewicz J, McCammon JA, Gilson MK. *Biochemistry*. 1996; 35:7819. [PubMed: 8672483]
60. Nielsen JE, Vriend G. *Proteins*. 2001; 43:403. [PubMed: 11340657]
61. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. *Proc Natl Acad Sci USA*. 2001; 98:10037. [PubMed: 11517324]
62. Frisch, MJ. *Gaussian 03*. Gaussian, Inc; Pittsburgh, PA:
63. Eigen M. *Angew Chem Int Ed Engl*. 1964; 3:1.
64. Schiöberg D, Zundel G. *J Chem Soc Faraday Trans II*. 1973(69):771.
65. Marx D, Tuckerman ME, Hutter J, Parrinello M. *Nature (London)*. 1999; 397:601.
66. Tuckerman M, Laasonen K, Sprik M, Parrinello M. *J Phys Chem*. 1995; 99:5749.
67. Tuckerman M, Laasonen K, Sprik M, Parrinello M. *J Chem Phys*. 1995; 103:150.
68. Marx D, Chandra A, Tuckerman M. *Chem Rev*. 2010; 120:2174. [PubMed: 20170203]
69. Reynolds WF, Peat IR, Freedman MH, Lyerla JR. *J Am Chem Soc*. 1973; 95:328. [PubMed: 4687673]
70. The relative pK_a values (pK_a^{Micro}) of 5.9 was estimated by integration of the PMF; 0.6 units of pH account for the multiplicity of the indistinguishable microscopic protonation states.
71. Gallivan JP, Dougherty DA. *Proc Natl Acad Sci USA*. 1999; 96:9459. [PubMed: 10449714]
72. Dougherty DA. *Science*. 1996; 271:163. [PubMed: 8539615]
73. The free energy changes for transferring the His-Trp cluster in the ionized ($\chi_{\text{fer}}G_{A-}$) and neutral ($\chi_{\text{fer}}G_{HA}$) forms from aqueous solution to the protein environment are -57.85 and -60.06 kcal/mol, respectively. Thus, the transfer free energy difference ($\chi_{\text{fer}}G$) is 2.21 kcal/mol, corresponding to 1.6 pH units.

The estimated pK_a is in agreement with the experimental one.

The affinity for protons is similar to that of a histidine residue in aqueous solution.

The electrostatic environment is responsible for the stabilization of the charged imidazolium moiety.

Figure 1.

Shown is a snapshot is taken from an unconstrained dynamics trajectory associated with the His-Trp cluster system. For visual clarity, only the water molecules in the first solvation shell of $N\delta$ and $N\epsilon$ are shown.

Figure 2.

Shown are snapshots taken from the unconstrained *ab initio* MD simulations trajectory where a protonated L-histidine ion is solvated by 69 water molecules. The nitrogen atoms at δ and ε positions of the imidazole are color-coded as orange and green, respectively.

Figure 3.

Average bond distances as a function of the reaction coordinate ξ . Left-hand side: The H*-O' distances for histidine deprotonation from the N δ (solid black) and N ϵ (solid red) positions and for the deprotonation of the His-Trp cluster systems from the N δ position (solid green) are shown. The N*-H* distances are shown by the blue line. Right-hand side: For the same system, the N*-O' distances are shown by the dotted black, red and green lines, respectively.

Figure 4. Mulliken charges (solid line) and Löwdin charges (dotted line) plotted as a function of the reaction coordinate ξ for N* (black), H* (red) and the histidine imidazole, excluding the charge of H* atom, (green) during the histidine deprotonation process (system 2b).

Figure 5.

Shown are the free energy profiles for the proton dissociation reaction taken from thermodynamic integration of the PMF along the distance constraint. Upper-left panel: deprotonation from the N δ (black) and N ϵ (red) positions of L-histidine in 1.03 M aqueous solution plus the water autodissociation reaction (green) using the BLYP functional. Upper-right panel: Free energy profiles for 1.03 M (black) and 0.37M (red) aqueous solutions. Lower-left panel: The effect of dispersion forces for BLYP (black) and for BLYP-D3 (red). Lower-right panel: Free energy profile for the deprotonation of L-histidine (black), His-Trp cluster (red) and the water autodissociation reaction (green) using the BLYP-D3 scheme.

Figure 6.

Mulliken charges (solid line) and Löwdin charges (dotted line) plotted as a function of the reaction coordinate ξ : N* (black), H* (red), histidine-imidazole excluding the charge of H* atom (green) and nearest neighbour tryptophan indole (blue) for the His-Trp cluster systems are plotted as a function of the reaction coordinate ξ .

Figure 7. The potential energy surface along the deprotonation reaction coordinate for the imidazole cation with either three water molecules (black), or six water molecules (red), respectively.

Figure 8. Shown is the ab initio electrostatic potential surface for four indoles. The blue and red surfaces represent positive and negative isovalues, respectively. Four imidazoles are superimposed to emphasize the cation- π interaction.

Table 1

Details of the computational setups of the *ab initio* MD simulations. The table reports for each system the number of water molecules ($N_{\text{H}_2\text{O}}$), the concentration (C_0), the titrable site considered in the deprotonation event ($\text{N}\delta$ or $\text{N}\epsilon$), and the computational approach (BLYP/BLYP-D3).

System	Solute	$N_{\text{H}_2\text{O}}$	C_0	Isomer	Method	F	$\text{p}K_a$
1a	L-histidine	49	1.03M	$\text{N}\delta$	BLYP	9.3	7.6
1b	L-histidine	49	1.03M	$\text{N}\delta$	BLYP-D3	10.1	7.3
1c	L-histidine	49	1.03M	$\text{N}\epsilon$	BLYP	10.2	8.3
1d	L-histidine	49	1.03M	$\text{N}\epsilon$	BLYP-D3	10.6	7.7
2a	L-histidine	142	0.37M	$\text{N}\delta$	BLYP	9.6	7.5
2b	L-histidine	142	0.37M	$\text{N}\delta$	BLYP-D3	10.0	6.8
2c	L-histidine	142	0.37M	$\text{N}\epsilon$	BLYP	10.0	7.7
2d	L-histidine	142	0.37M	$\text{N}\epsilon$	BLYP-D3	10.5	7.1
3	His-Trp	183	0.22M	$\text{N}\delta$	BLYP-D3	9.1	6.5
4		64	56.0M		BLYP	18.2	
5		64	56.0M		BLYP-D3	19.5	

Table 2

Free energy differences (ΔF in kcal/mol) and pK_a values of all the deprotonation events considered in this work.

System	Isomer	Method	ΔF	pK_a
1a	$N\delta$	BLYP	9.3	7.6
1b	$N\delta$	BLYP-D3	10.1	7.3
1c	$N\epsilon$	BLYP	10.2	8.3
1d	$N\epsilon$	BLYP-D3	10.6	7.7
2a	$N\delta$	BLYP	9.6	7.5
2b	$N\delta$	BLYP-D3	10.0	6.8
2c	$N\epsilon$	BLYP	10.0	7.7
2d	$N\epsilon$	BLYP-D3	10.5	7.1
3	$N\delta$	BLYP-D3	9.1	6.5
4		BLYP	18.2	
5		BLYP-D3	19.5	