JCTC Journal of Chemical Theory and Computation

Histidine−Aromatic Interactions in Proteins and Protein−Ligand Complexes: Quantum Chemical Study of X-ray and Model Structures

Emilie Cauët,[†] Marianne Rooman,[‡] René Wintjens,[§] Jacques Liévin,[†] and Christophe Biot*,[‡]

Service de Chimie quantique et Photophysique, Université Libre de Bruxelles, CP 160/09, 50 Avenue F.D. Roosevelt, B-1050 Bruxelles, Belgium, Unité de Bioinformatique génomique et structurale, Université Libre de Bruxelles, CP 165/61, 50 Avenue F.D. Roosevelt, Institut de Pharmacie, B-1050 Bruxelles, Belgium, and Service de Chimie générale, Université Libre de Bruxelles, CP 206/04, Bld du Triomphe, B-1050 Bruxelles, Belgium

Received November 23, 2004

Abstract: His−aromatic complexes, with the His located above the aromatic plane, are stabilized by $\pi-\pi$, $\delta^+-\pi$ and/or cation$-\pi$ interactions according to whether the His is neutral or protonated and the partners are in stacked or T-shape conformations. Here we attempt to probe the relative strength of these interactions as a function of the geometry and protonation state, in gas phase, in water and protein-like environments (acetone, THF and CCl_4), by means of quantum chemistry calculations performed up to second order of the Møller−Plesset pertubation theory. Two sets of conformations are considered for that purpose. The first set contains 89 interactions between His and Phe, Tyr, Trp, or Ade, observed in X-ray structures of proteins and protein−ligand complexes. The second set contains model structures obtained by moving an imidazolium/imidazole moiety above a benzene ring or an adenine moiety. We found that the protonated complexes are much more stable than the neutral ones in gas phase. This higher stability is due to the electrostatic contributions, the electron correlation contributions being equally important in the two forms. Thus, $\pi-\pi$ and $\delta^+-\pi$ interactions present essentially favorable electron correlation energy terms, whereas cation$-\pi$ interactions feature in addition favorable electrostatic energies. The protonated complexes remain more stable than the neutral ones in protein-like environments, but the difference is drastically reduced. Furthermore, the T-shape conformation is undoubtedly more favorable than the stacked one in gas phase. This advantage decreases in the solvents, and the stacked conformation becomes even slightly more favorable in water. The frequent occurrence of His−aromatic interactions in catalytic sites, at protein−DNA or protein−ligand interfaces and in 3D domain swapping proteins emphasize their importance in biological processes.

Introduction

Cation$-\pi$ interactions were shown over the past 15 years to play important roles in many chemical and biological processes.[1−3] In proteins, for example, cation$-\pi$ interactions between an aromatic π ring (i.e. Phe, Tyr or Trp) and an organic cation (i.e. Lys or Arg) are frequently observed.[4−9] Their stabilizing nature was demonstrated by both in silico energy calculations and experimental means. In particular,

* Corresponding author e-mail: christophe.biot@ensc-lille.fr. Current address: Laboratoire de Catalyse de Lille − UMR CNRS 8010, ENSCL, Bâtiment C7, Université des Sciences et Technologies, B.P. 90108, 59652, Villeneuve d'Ascq Cedex, France.

[†] Service de Chimie quantique et Photophysique.

[‡] Unité de Bioinformatique génomique et structurale.

[§] Service de Chimie générale.

10.1021/ct049875k CCC: $30.25
Published on Web 03/09/2005

a specific cation−π interaction between a phenyl ring and a protonated amine in an α-helix was measured as stabilizing the protein structure by 0.4 kcal/mol.[10]

Several biomolecular association processes such as ligand-antibody binding and receptor−ligand interactions also feature cation−π interactions.[11,12] Cation−π interactions have also been identified at the protein−DNA[13,14] and protein−ligand interface,[15,16] where the nucleic acid bases play the role of aromatic partners. The existence of cation−π interactions between aromatic side chains and metal ions is more debated; note however the interaction between a Na^+ and a Trp ring experimentally evidenced in hen egg white lysozyme.[17]

His side chains are also commonly observed as potential cation−π partners in protein structures but have much less been investigated, due to their ubiquitous nature. Indeed, the imidazole cycle of His has a pK_a value of 6.1 and can therefore occur as neutral or protonated under physiological conditions. Two cases can be distinguished according to whether the imidazole ring is located above a positively charged or aromatic side chain. In the former, the His is obviously neutral and plays the role of aromatic moiety of the cation−π pair, whereas in the latter, the His may be either neutral or protonated. When neutral, it forms aromatic−aromatic interactions with its aromatic partner as well as δ^+−π interactions, due to polarization. When protonated, it plays the role of cation in the cation−π pair. We focus in the present paper on His residues interacting with aromatic side chains and analyze the effect of protonation on the strength of the interaction.

In barnase, experimental evidence indicates that protonated His18 interacts with Trp94 more strongly than the neutral form, by about 1 kcal/mol, and increases the pK_a value of His18; this suggests that the His−aromatic cation−π interactions play an important role in the stabilization of the structure.[18] The stabilizing role of protonated His interacting with aromatic side chains has also been experimentally evidenced in α-helices[19] and in apoflavodoxin.[20] Moreover, His-involving cation−π interactions have been described as having a functional role, by affecting ligand binding[21,22] or catalysis in ion channels,[23−25] in G-protein-coupled receptors,[26] in transporters[27] and in enzymes.[28]

To our knowledge, no in silico characterization of His−aromatic interactions have been performed up to now. To fill this lacuna, we searched X-ray structures of proteins and protein−ligand complexes for His−aromatic interactions and probed their interaction free energy contribution in water and in different solvents mimicking protein environments (CCl_4, THF and acetone) by means of quantum chemistry computations. The free energies calculated for the protonated and neutral forms of His were compared so as to evaluate the relative weight of the cationic and aromatic natures of His in stabilizing the His−aromatic interaction (δ^+−π or cation−π). To complete the analysis, we also considered model structures to allow a more thorough investigation of the stacked and T-shape geometries of His−Phe and His−Ade pairs. We end by discussing the potential functional role of these interactions.

Methods

1. Sets of X-ray Cation−π Complexes. Two sets of high-resolution X-ray structures with less than 25% sequence identity were considered. One comprehends 141 protein chains[30] and the other 188 complexes of proteins and nucleobase containing ligands.[15] The first set was searched for cation−π interactions between an His and an aromatic side chain (Trp, Tyr, Phe) and the second for His−nucleic acid base (Ade, Gua, Cyt, Thy).

Cation−π interactions were identified according to a distance and an angle criterion.[13] The distance criterion requires that at least one of the atoms of the aromatic ring is located no further than 4.5 Å from one of the atoms carrying the positive charge of the His, considered to the $N_{\delta 1}$, $N_{\epsilon 2}$, and $C_{\epsilon 1}$-atoms in the absence of H-atoms. The angle criterion demands the latter atom to be situated above the plane defined by the aromatic ring, more precisely, inside a cylinder of height 4.5 Å, whose base includes the ring and has a radius equal to twice the radius of the ring.

The cation−π pairs were simplified for computational study. Ligands were reduced to the nucleobase(s) they contain. Phe, Tyr and Trp were represented as benzene, phenol and indole moieties, respectively, and His as an imidazole ring (protonated and unprotonated). The H-atoms were added by construction. For neutral His, the two tautomeric forms are considered, with an H-atom linked to either $N_{\delta 1}$ or $N_{\epsilon 2}$.

The table containing the list of His−aromatic interactions is given as Supporting Information.

2. Geometrical Parameters Characterizing the Model Imidazole−Aromatic Complexes. Figure 1 illustrates the geometrical parameters used to investigate how interaction energies evolve with the relative positions of the molecular fragments. Stacked and T-shape conformations of the imidazole−benzene complexes are represented in parts (a) and (b) of the figure respectively, and part (c) illustrates a stacked imidazole−adenine complex. Parameters α, β, d and θ have the same meaning in all complexes: α and β are the molecular planes of both interacting partners, d is the distance from the imidazole ring center to the β plane and θ defines the orientation of imidazole within α with respect to benzene or adenine. More details about the parameters are given in the legend of Figure 1 and in the next sections detailing the calculations. Note that all calculations were performed with the geometrical parameters of the fragments frozen to their HF/6-31G(d, p) equilibrium values.

3. Quantum Chemistry Energy Calculations. All energy calculations were carried out using the Gaussian 03 suite of programs.[31] In a first step, the nucleic acid bases and the benzene, phenol, indole and neutral and protonated imidazole rings were optimized separately, using the Hartree−Fock method and the 6-31G(d, p) basis set. No further geometry optimizations of the internal degrees of freedom of the individual partners of the complexes were carried out. When dealing with X-ray structures, the neutral or protonated imidazole and the aromatic moiety were replaced by the individually optimized partners using the U3BEST algorithm.[32]

The gas-phase interaction energies ΔE_{MP2} of the His−aromatic systems were calculated at second-order level of

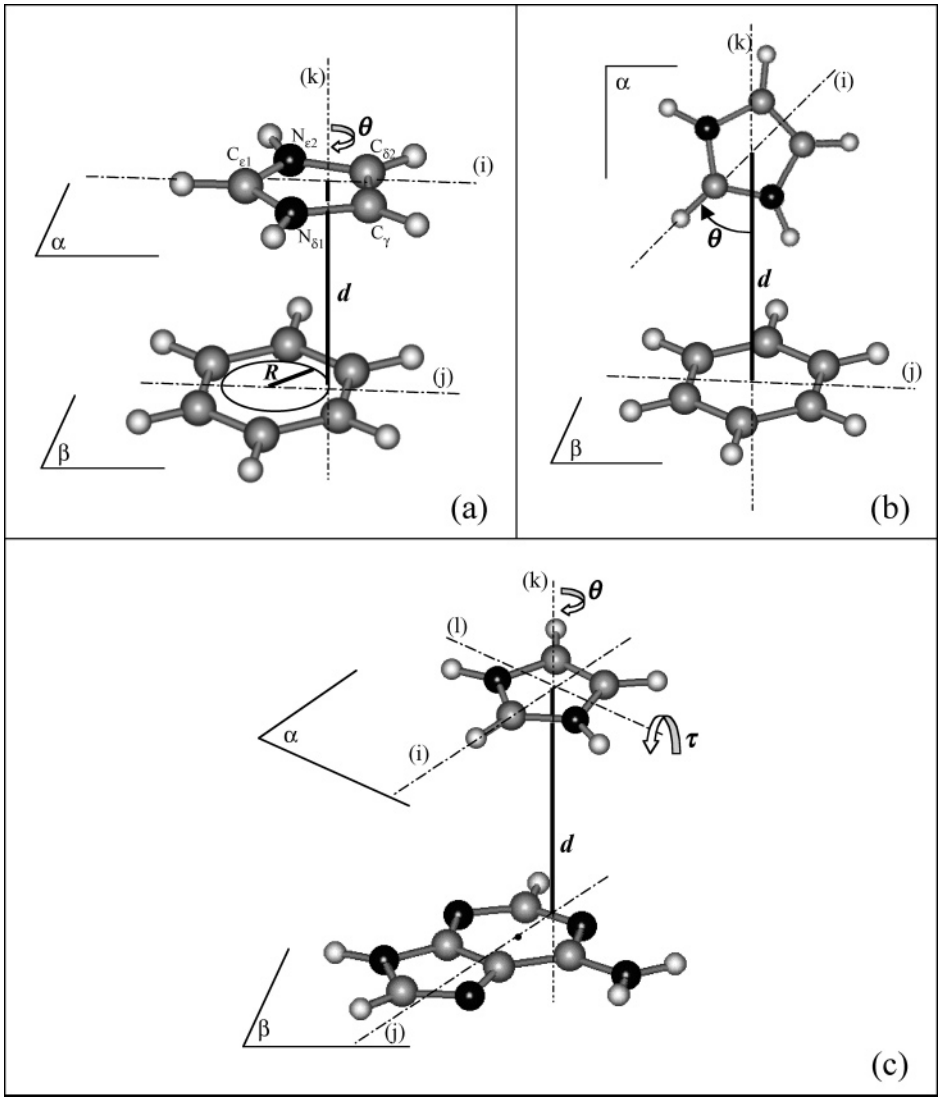

Figure 1. Geometrical parameters defining imidazole−aromatic (benzene or adenine) complexes. Only the protonated form of imidazole is illustrated, but the same parameters hold for the neutral species. According to standard labeling, one defines the nitrogen atoms of imidazole as $N_{\epsilon 2}$ and $N_{\delta 1}$, with carbon $C_{\epsilon 1}$ between them. $N_{\epsilon 2}$ is unprotonated in neutral imidazole. Hydrogens are labeled with a subscript following the heavy atom to which they are bonded ($H_{\epsilon 1}$ bonded to $C_{\epsilon 1}$ for instance). Carbon, nitrogen and hydrogen atoms are drawn in gray, black and white, respectively. The molecular planes of the imidazole and of the aromatic species are α and β respectively. (a) stacked imidazole−benzene complex with $\alpha//\beta$. (b) T-shape imidazole−benzene complex $\alpha\perp\beta$. (c) stacked imidazole−adenine complex. *d* is the distance between β and the center of the imidazole ring. *R* is the radius of a circle in β centered on the benzene ring center and containing the orthogonal projection in β of the imidazole center. θ defines the rotation of (i), the line passing through the $C_{\epsilon 1}-H_{\epsilon 1}$ bond, around the imidazole ring center. Parts a and c show the orientation of imidazole with $\theta = 0°$. τ is the tilt angle, which is equal to zero for a parallel conformation.

Møller−Plesset pertubation theory (MP2),[33] as the sum of the Hartree−Fock (HF) energy contribution ΔE_{HF} and a correlation interaction energy ΔE_{cor}, evaluated as the second-order term of the Moller−Plesset perturbation expansion:

$$\Delta E_{MP2} = \Delta E_{HF} + \Delta E_{cor} \quad (1)$$

The standard counterpoise (CP) method was applied to correct the interaction energies for the basis set superposition error (BSSE).[34] The basis set used is 6-31G(2d (0.8,0.2),p), which corresponds to the standard 6-31G(d,p) basis set where the Gaussian α_d-exponent of the d-polarization functions on the heavy atoms C, N and O is equal to 0.8, with an additional α_d-exponent equal to 0.2. It has indeed been shown that this extended description of the d-polarization functions allows a more accurate description of cation−π energies.[35]

Finally, the interaction free energy of complexes in the presence of a solvent, noted $\Delta\mathbf{G}$, was evaluated as

$$\Delta\mathbf{G} = \Delta E_{MP2} + \Delta\Delta G_{solv} \quad (2)$$

where $\Delta\Delta G_{solv}$ is the solvation free energy contribution of the complex. The latter were estimated using the integral equation formalism version (IEF) of the polarized continuum model (PCM) implemented in the Gaussian 03 program.[36] It is a continuum solvation model in a quantum mechanical framework, where the solvent is mimicked by a polarizable continuum surrounding a cavity having the shape and

Table 1. Vacuum Interaction Energy (in kcal/mol) of the 89 X-ray His−Aromatic Pairs, Calculated at HF and MP2 Levels of Theory, with 6-31G(2d (0.8,0.2), p) Basis Set[a]

		protonated His			neutral His		
	N	ΔE_{HF}	ΔE_{cor}	ΔE_{MP2}	ΔE_{HF}	ΔE_{cor}	ΔE_{MP2}
His−Phe	34	−3.3 ± 2.4 (−8.5)	−3.3 ± 2.1 (−8.5)	−6.7 ± 1.7 (−11.0)	1.6 ± 2.9 (−1.8)	−3.8 ± 2.1 (−9.4)	−2.2 ± 1.4 (−4.7)
His−Tyr	28	−3.4 ± 2.3 (−10.0)	−3.8 ± 1.7 (−6.9)	−7.2 ± 2.5 (−13.9)	2.2 ± 1.4 (0.2)	−4.2 ± 1.7 (−7.6)	−1.9 ± 1.3 (−5.4)
His−Trp	13	−4.7 ±3.5 (−11.2)	−4.0 ± 2.0 (−8.1)	−8.7 ± 3.1 (−13.4)	1.0 ± 1.9 (−1.6)	−4.5 ± 1.9 (−8.9)	−3.4 ± 0.5 (−4.2)
His−Ade	14	−3.5 ± 2.5 (−7.0)	−4.2 ± 2.2 (−8.1)	−7.7 ± 4.0 (−12.0)	0.9 ± 1.8 (−2.0)	−4.6 ± 2.4 (−8.6)	−3.6 ± 0.9 (−5.1)

[a] N denotes the number of occurrences in the sets of proteins and protein−ligand complexes. The mean energy values ± the standard deviation are given, with the minimum energy values in parentheses.

dimension of the solute molecule. The cavity is described by interlocking spheres centered on solute atoms; we used the default values for the atomic radii (UATM), multiplied by a default factor that accounts for the fact that the distance between the solvent and solute atoms is normally somewhat larger than the van der Waals radii.

By use of a PCM, the molecular free energy G of the solute embedded in a continuum medium can be expressed as the sum:

$$G = G_{el} + G_{dis-rep} + G_{cav} + G_{mm} \quad (3)$$

G_{el} collects the components of electrostatic origin, $G_{dis-rep}$ is a term related to the solute−solvent dispersion energy, and G_{cav} is the work needed to form the cavity where the solute must be accommodated. G_{mm} includes all the contributions due to molecular motions. Neglecting the entropic contributions due to molecular motions, the solvation free energy is obtained by subtracting the gas-phase energy E_{gas}

$$\Delta G_{solv} = \Delta G_{el} + G_{nel} \quad (4)$$

with

$$\Delta G_{el} = G_{el} - E_{gas} \text{ and } G_{nel} = G_{dis-rep} + G_{cav}$$

We thus find

$$\Delta\Delta G_{solv} = \Delta\Delta G_{el} + \Delta G_{nel} \quad (5)$$

where $\Delta\Delta G_{el}$ and ΔG_{nel} are the differences of electrostatic and nonelectrostatic contributions between the complex and its isolated partners.

The IEF−PCM calculations were performed at the HF/6-31G(2d (0.8,0.2), p) level, and some tests were performed at the MP2/6-31G(2d (0.8,0.2), p) level. Note that the quantity that emerges from the MP2/PCM calculations is G_{el}. The nonelectrostatic terms must be then added to it.

Results

1. Geometries of Molecular Fragments. To probe the His−aromatic complexes using quantum chemistry energy calculations, Phe, Tyr, Trp and protonated/neutral His were reduced to benzene, phenol, indole, and imidazolium/imidazole, respectively; Ade was kept as such. The geometries of the molecular fragments were optimized at the HF/6-31G(d, p) level of approximation. Both imidazole species were found to be planar, but some in plane geometry changes occur upon protonation. In particular, changes of +0.02 and −0.04 Å are observed for the $C_{\epsilon1}-N_{\epsilon2}$ and $C_{\epsilon1}-N_{\delta1}$ bond lengths respectively (see Figure 1a for the atom labeling), and changes of 1° to 4° for the aromatic ring bond angles.

The natural bond orbital (NBO) atomic charges calculated at the MP2/6-31G(2d (0.8,0.2), p) level of theory for the protonated (neutral) imidazole species are as follows: $C_{\epsilon1}$: 0.26 (0.17); $H_{\epsilon1}$: 0.28 (0.22); $N_{\delta1}$: −0.49 (−0.57); $H_{\delta1}$: 0.49 (0.44); $N_{\epsilon2}$: −0.49 (−0.50); C_{γ}: −0.05 (−0.11); H_{γ}: 0.29 (0.23); $C_{\delta2}$: −0.05 (−0.10); $H_{\delta2}$: 0.29 (0.23); $H_{\epsilon2}$: 0.49. In the case of protonated imidazole, a delocalization of the positive charge is observed over the $C_{\epsilon1}$, $H_{\epsilon1}$, $H_{\epsilon2}$ and $H_{\delta1}$ atoms. These atoms are bearing together a total charge of 0.69, which means that the positive charge of the imidazole cation is globally located in the vicinity of the $C_{\epsilon1}-H_{\epsilon1}$ bond. In the case of neutral imidazole, a partial positive charge of 0.51 is found over the same $C_{\epsilon1}-H_{\epsilon1}$ bond. Hence, when the imidazole is protonated, it can form a cation−π interaction with an aromatic ring, and when neutral, it can form interactions of the type C−H−π[29] or $\delta^+-\pi$.[37−39]

2. His−Aromatic Interactions in X-ray Structures. A total of 75 His−aromatic interactions were identified in the set of protein X-ray structures: 34 His−Phe, 28 His−Tyr and 13 His−Trp. In the set of protein−ligand complexes, 14 His−Ade complexes were found. No His−Gua, His−Cyt and His−Thy were observed, probably because of the infrequency of these nucleic acid bases in ligands. The average angle between the aromatic plane is equal to 42 degrees, reflecting a slight preference for stacked compared to T-shape conformations (see Supporting Information).

The vacuum interaction energy of these complexes, with the separate molecular fragments replaced by the optimized ones (see Methods), were computed at HF and MP2 theory levels with the 6-31G(2d (0.8,0.2), p) basis set. Three species of His were considered: the protonated imidazole cycle (imidazolium) and the two neutral ring forms resulting from the tautomeric equilibrium of His. The results are summarized in Table 1. For the protonated His form, the minimal HF interaction energies were found to be comprised between −7 kcal/mol for His−Ade pairs and −11 kcal/mol for His−Trp. The electronic correlation contributions are of the same order of magnitude: from −7 kcal/mol for His−Tyr up to −9 kcal/mol for His−Phe. The most favorable HF interaction energies, known to be dominated by electrostatic forces, appears thus in the Trp double cycle, whereas the most favorable electronic correlation is found with the benzene ring of Phe, the only aromatic moiety that does not carry heteroatoms.

The relative strengths of the ΔE_{HF} and ΔE_{cor} contributions are different in the protonated and neutral complexes. Indeed,

Table 2. HF/6-31G(2d (0.8,0.2), p) Solvation Free Energies $\Delta\Delta G_{solv}$ and Total Interaction Free Energies ΔG (in kcal/mol) with the Vacuum Contributions Evaluated at the MP2/6-31G(2d (0.8,0.2), p) Level, in Four Solvents Characterized by Their Dielectric Constant ϵ (IEF−PCM Model), for the 89 X-ray His−Aromatic Pairs[a]

solvent ϵ	*His−X*	$\Delta\Delta G_{solv}$	ΔG
water	*F*	8.1 ± 1.7 (4.8)	1.4 ± 1.6 (−1.4)
78.4	*Y*	8.3 ± 2.0 (4.8)	1.2 ± 1.2 (−1.6)
	W	10.2 ± 3.3 (5.2)[b]	1.5 ± 1.2 (−1.3)[b]
	Ade	7.5 ± 2.5 (4.1)	2.9 ± 0.5 (1.9)
acetone	*F*	7.2 ± 1.4 (4.6)	0.6 ± 1.7 (−2.3)
20.7		[7.0 ± 1.3 (4.7)]	
	Y	7.6 ± 1.5 (5.2)	0.5 ± 1.6 (−3.0)
		[7.6 ± 1.4 (5.3)]	
	W	9.2 ± 2.4 (5.2)	0.5 ± 1.7 (−3.4)
		[9.0 ± 2.3 (5.4)]	
	Ade	8.9 ± 2.7 (3.2)	1.2 ± 1.7 (−1.4)
		[8.9 ± 2.6 (3.4)]	
THF	*F*	7.1 ± 1.3 (4.9)	0.5 ± 1.8 (−2.6)
7.6	*Y*	7.6 ± 1.6 (5.0)	0.5 ± 2.5 (−6.4)
	W	9.3 ± 2.9 (5.1)	0.6 ± 1.7 (−3.2)
	Ade	9.0 ± 2.9 (3.4)	1.3 ± 1.6 (−1.0)
CCl_4	*F*	5.5 ± 1.2 (2.7)	−1.2 ± 2.1 (−8.3)
2.2	*Y*	6.0 ± 0.7 (4.4)	−1.2 ± 2.1 (−6.6)
	W	6.5 ± 2.1 (0.6)	−2.3 ± 2.2 (−5.5)
	Ade	7.3 ± 1.9 (3.6)	−0.4 ± 2.5 (−3.6)
gas phase	*F*		−6.7 ± 1.7 (−11.0)
MP2	*Y*		−7.2 ± 2.5 (−13.9)
1.0	*W*		−8.7 ± 3.1 (−13.4)
	Ade		−7.7 ± 4.0 (−12.0)

[a] All His are considered to be protonated. Results for neutral His are given in the Supporting Information. The mean energy values ± the standard deviation are given, with the minimum values in parentheses. For acetone, $\Delta\Delta G_{solv}$ values calculated at MP2 level are given in brackets. The last row displays the gas phase $\Delta G_{gas} = \Delta E_{MP2}$ values. [b] The complex between W131 and H88 in 2gcr was removed from the analysis due to abnormal $\Delta\Delta G_{solv}$ value in water.

the ΔE_{cor} values are roughly the same, while the ΔE_{HF} energies stabilize the protonated species and slightly destabilize the neutral ones. This is not surprising, as the electrostatic contributions almost vanish in the neutral complexes and HF contributions are mainly of electrostatic nature.

For both neutral and charged complexes, the conformations presenting the most favorable electronic correlation interaction energies have a rather stacked geometry, stabilized by significative dispersion energy contributions.

The solvation free energies $\Delta\Delta G_{solv}$ of the complexes immersed in water, acetone, THF, and CCl_4 were estimated using the IEF−PCM method (Table 2 and Supporting Information). In the first stage, the calculations for acetone were performed at both HF and MP2 levels. For the charged complexes, the $\Delta\Delta G_{solv}$ values are found to be almost identical within 0.2 kcal/mol (see Table 2), as previously reported for Arg−Ade complexes in water.[16] The same agreement (within 0.1 kcal/mol) is observed for the neutral form (see Supporting Information). On the basis of this result, and in order to save computer time, we limited further solvent calculations to the HF level, unless otherwise stated. The total interaction free energy values were thus calculated as $\Delta\mathbf{G} = \Delta E_{MP2} + \Delta\Delta G_{solv(HF)}$.

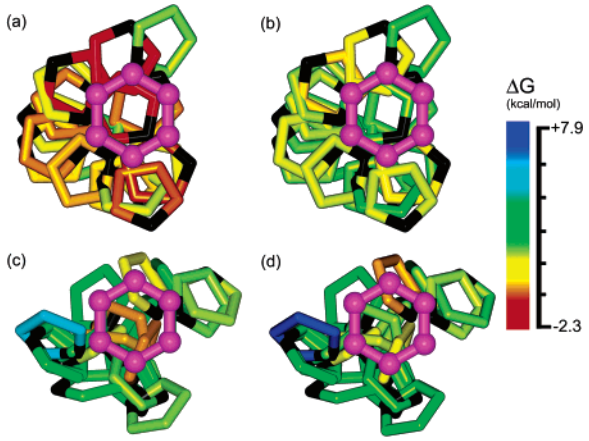

Figure 2. Distribution of the imidazole rings of His residues around the phenyl cycle of Phe amino acids, observed in the X-ray protein set, as a function of their calculated interaction free energy values $\Delta\mathbf{G}$ in acetone. The phenyl groups are superimposed, and all the imidazole rings have been located "under" the Phe planes by suitable projection, to facilitate comparison. The calculated free energy values are colored according to a rainbow ramp where the most favorable conformations are symbolized in red and the less favorable in blue. Carbon atoms of the Phe rings are colored in purple and Cγ atoms of the imidazole cycle are colored in black. The images were generated using Insight II (Accelrys, Inc.) Stacked conformation of the protonated (a) and neutral (b) imidazole rings. T-shape conformations of the protonated (c) and neutral (d) imidazole cycles.

The penalty upon immersing His−aromatic complexes in the solvent increases with the dielectric constant (Table 2), especially from $\epsilon = 2.2$ to 7.6. It is also much larger for the protonated than for the neutral form. Indeed, the $\Delta\Delta G_{solv}$ values for the former are about 4 kcal/mol larger than those for the latter, on the average (see the Supporting Information). As a consequence, the larger stability of protonated His−aromatic complexes in a vacuum relative to neutral ones is reduced by the solvents. Nevertheless, they remain more stable in the four solvents considered. In summary, the interaction free energy $\Delta\mathbf{G}$ increases from gas to water, with minimal values remaining negative except for the Ade−His complexes in water, and remaining slightly more negative for protonated than for neutral form.

We now analyze in more detail the His−Phe and His−Ade complexes by means of X-ray and model structures.

3. Detailed Analysis of His−Phe Interactions. (a) X-ray Structures. The 34 His−Phe X-ray pairs were divided into stacked and T-shape conformations, according to whether the angle between the His and benzene planes is lower or higher than 45°. The stacked and T-shape groups contains 19 and 15 His−Phe pairs, respectively. For each of these two groups, the interaction free energies $\Delta\mathbf{G}$ in acetone as a function of the relative position and orientation of the His and Phe moieties are depicted in Figure 2.

As clearly seen in Figure 2, the stacked conformations have more favorable interaction free energies than T-shape conformations on average, the protonated forms have more

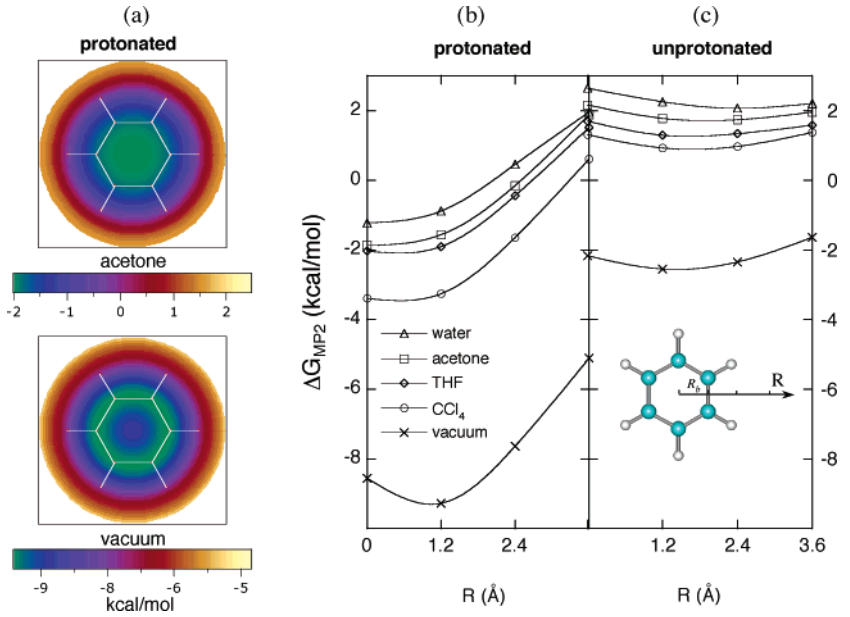

Figure 3. Model structures of the stacked imidazole−benzene complex: maps of $\Delta\mathbf{G}$ in acetone as a function of the position of imidazole in the α plane. All calculations are performed at the MP2/6-31G(2d (0.8,0.2), p) level and with the IEF/PCM approach for solvents. Imidazole is located in a plane parallel to the aromatic ring of benzene at an optimal distance of $d = 3.5$ Å (3.7 Å) for the protonated (neutral) species. The maps presented in panel (a) for the protonated complex in the vacuum and in acetone are obtained by moving the projection of the imidazole center in β on circles of radii R = 0 to 3.6 Å (3 times the radius R_b of the inscribed circle to the benzene hexagon). See Figure 1a for a visual definition of the parameters. Note that the colored scale covers the same energy range in both maps, but the minimum is shifted. Panels (b) and (c) show one-dimensional cuts ΔG(R) in the 2D surfaces (see figure inset).

favorable interaction free energies than the neutral forms, and the most favorable interaction free energies are observed above the benzene ring.

For several His−Phe pairs, the $\Delta\mathbf{G}$ value of the neutral form is equal to or even slightly more favorable than that of the protonated form. Considering that the protein X-ray structures contain a mixture of protonated and neutral His residues, it can be argued that the complexes computed as more favorable in neutral form actually contain a neutral His. This could be a way to determine which His residues in a folded structure are charged and which are not. This possibility should however be considered with care, because of the residual flexibility in folded proteins and the limited resolution of the structures, which do not exclude local inaccuracies in atomic positioning.

(b) Model Conformations: Stacked. To probe in detail the evolution of the interaction free energy as a function of the position and orientation of imidazolium/imidazole with respect to benzene, we analyzed ensembles of model structures in which the molecular planes α and β were kept parallel (perfectly stacked conformations). Three parameters define the relative positions of both species (Figure 1a): the interplane distance d, the radius R of a circle centered on the benzene ring center and containing the orthogonal projection in β of the imidazole center, and the angle θ defining the rotation of the $C_{\epsilon 1}-H_{\epsilon 1}$ bond around the imidazolium/imidazole ring center.

In the first step, we determined the minimum of the vacuum interaction energy as a function of the distance d, by approaching both species while keeping their centers superimposed (R=0). The optimal distances, calculated at the MP2/6-31G(2d (0.8,0.2), p) level of theory, are of 3.5 and 3.7 Å for protonated and neutral imidazoles, respectively.

In the second step, having fixed the interplane distances d to its optimal value, we studied the effect of moving the imidazole with respect to the benzene in α. The projection in β of the imidazole center was displaced on circles of radius R (see Figure 1a), R being varied from 0 to 3 times the radius R_b of the inscribed circle to the benzene hexagon ($R_b = 1.2$ Å). The energetic effect of the imidazole orientation parameter θ was investigated at different relative positions of the fragments, but its effect on the interaction energies was found to be small ($\Delta E < 0.5$ kcal/mol), and the preferred orientation of $\theta = 0°$ was adopted in all further calculations.

At this stage, we added the solvation contributions in water, acetone, THF or CCl_4 and evaluated the evolution of the interaction free energy $\Delta\mathbf{G}$ as a function of the position of the stacked imidazole with respect to benzene, when the imidazole moves away from the benzene in a linear displacement parallel to the benzene plane (see Figure 3).

In agreement with the results on the X-ray complexes, the protonated complex is found significantly more stable than the neutral one (see Figure 3). This can be explained by the enhancement of the interaction induced by the positive charge carried by imidazole. The difference in interaction energy between protonated and neutral species thus provides a direct measure of this charge contribution. It amounts to −7 kcal/mol in a vacuum and is significantly weakened in solution: −4 kcal/mol in CCl_4 and −3 kcal/mol in THF, acetone and water, due to the screening of the electrostatic interactions by the solvent. Another finding is the existence of pronounced minima over the benzene ring in the protonated case,

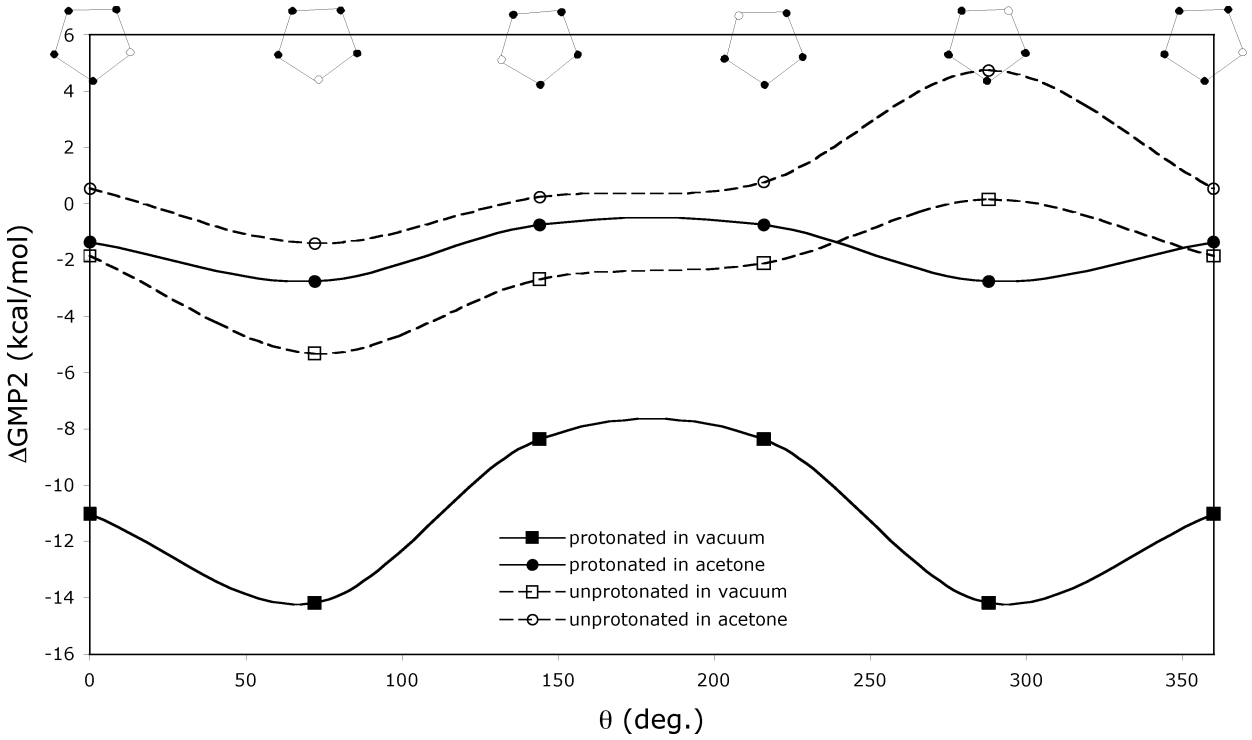

Figure 4. Model conformations of T-shape imidazole−benzene complex: evolution of $\Delta\mathbf{G}$ in acetone as a function of angle θ (see Figure 1b). All calculations are performed at the MP2/6-31G(2d (0.8,0.2), p) level and with the IEF/PCM approach for representing the acetone solvent environment. Drawings of the imidazole kernel (hydrogens are omitted) show its orientation in the T-shape structure, the lowest atom pointing toward the benzene ring. Color code for atoms is gray for carbon, white for protonated nitrogen and black for unprotonated nitrogen.

while very flat profiles characterize the neutral complex. One can again invoke the cation$-\pi$ interaction, when the positively charged imidazolium approaches the π-electron benzene cloud, inducing a deepening of the potential well.

The (free) energy maps in a vacuum and acetone, shown in Figure 3, exhibit a central symmetry in which the 6-fold symmetry of the benzene electronic structure does not appear, as a consequence of the aromatic π-electron delocalization. The energy changes along a given circle are indeed negligible ($<$0.02 kcal/mol). A comparison of both maps shows that the effect of the solvent on the shape of the potential energy surface is minor. Indeed, the two maps exhibit a flat minimum covering all the benzene ring and a destabilization of the same amplitude when the imidazole moves away from the benzene ring ($R >$ 1.2 Å). However, the vacuum energy surface is shifted by about +7 kcal/mol in the presence of the solvent. Furthermore, the existence of a very shallow minimum at a value of R=R_b distinguishes the gas-phase map from the solvated one. This minimum corresponds to a position of the imidazole center just above the C−C bonds of the benzene. The positively charged atoms of imidazole ($C_{\epsilon 1}$, $H_{\epsilon 1}$, $H_{\epsilon 2}$ and $H_{\delta 1}$) are then located above the aromatic ring of the benzene, thus in a favorable position for cation$-\pi$ interactions. Similarly shifted optimal conformations have already been reported in benzene[40,41] and naphthalene dimers.[42]

(c) Model Conformations: T-Shape. In the case of the perpendicular conformation, two parameters, illustrated in Figure 1b, were explored: the distance d between the β plane and the center of the imidazole ring and the angle θ defining the orientation of imidazole within the α plane perpendicular to β.

Optimal values of the parameters were obtained in a vacuum at the MP2 level of theory, as in the case of the stacked conformation. The following three step procedure has been adopted: i) optimization of θ for a d distance fixed to 5 Å; ii) optimization of d at the previous optimal θ-value; and iii) reoptimization of θ at the previous d-value. The optimal θ-value is found to be the same ($\theta \sim 72°$) in the first and last step, indicating the decoupling of the parameters d and θ. The optimal d values are equal to 4.2 and 4.5 Å for the protonated and neutral complexes, respectively. The interaction (free) energy profiles in a vacuum and acetone as a function of θ are shown in Figure 4.

Let us first remark that all the interaction (free) energy curves present a minimum when the N−H bond points toward the π-electron cloud of the benzene, leading thus to a couple of symmetric minima ($\theta \sim 72°$ and 288°) for the protonated species and a single minimum ($\theta \sim 72°$) for the neutral one. At these minima, the MP2 binding energies reach −14 and −5 kcal/mol in a vacuum for the protonated and neutral complexes, respectively. The corresponding values are −6 and −2 kcal/mol respectively in CCl_4, −3 and −1 kcal/mol in THF, −3 and −1 kcal/mol in acetone and 0.8 and 0.9 kcal/mol in water. The contribution of the positive charge to the complex stability, obtained from differences between the above cited energies for protonated and neutral complexes, are as follows: −9 kcal/mol in a vacuum, −4 kcal/mol in CCl_4, −2 kcal/mol in THF, −2 kcal/mol in

Table 3. Interaction Energies (in kcal/mol) for the Modeled Benzene−Imidazole Complexes in Protonated and Neutral Forms[a]

solvent	stacked				T-shape			
	protonated		neutral		protonated		neutral	
ϵ	ΔG	$\Delta\Delta G_{solv}$	ΔG	$\Delta\Delta G_{solv}$	ΔG	$\Delta\Delta G_{solv}$	ΔG	$\Delta\Delta G_{solv}$
water	−0.9	8.4	2.3	4.8	0.8	15.0	0.9	6.2
78.4	(4.9)	(8.9)	(7.5)	(5.1)	(6.0)	(15.3)	(5.3)	(6.3)
	[−5.8]	[−0.5]	[−5.2]	[−0.3]	[−5.2]	[−0.3]	[−4.4]	[−0.1]
acetone	−1.9	7.4	1.3	3.8	−2.8	11.4	−1.4	3.9
20.7	(3.6)	(7.6)	(6.4)	(4.0)	(2.3)	(11.6)	(2.9)	(3.9)
	[−5.5]	[−0.2]	[−5.1]	[−0.2]	[−5.1]	[−0.2]	[−4.3]	[0.0]
THF	−1.6	7.7	1.8	4.3	−3.2	11.0	−1.2	4.1
7.6	(3.9)	(7.9)	(6.9)	(4.5)	(1.9)	(11.2)	(3.1)	(4.1)
	[−5.5]	[−0.2]	[−5.1]	[−0.1]	[−5.1]	[−0.2]	[−4.3]	[0.0]
CCl_4	−3.3	6.0	0.9	3.5	−6.5	7.7	−2.2	3.1
2.2	(2.1)	(6.1)	(5.9)	(3.5)	(−1.5)	(7.8)	(2.1)	(3.1)
	[−5.4]	[−0.1]	[−5.0]	[0.0]	[−5.0]	[−0.1]	[−4.3]	[0.0]
gas phase	−9.3		−2.5		−14.2		−5.3	
1.0	(−4.0)		(2.4)		(−9.3)		(−1.0)	
	[−5.3]		[−4.9]		[−4.9]		[−4.4]	

[a] Results obtained at the MP2/6-31G(2d (0.8,0.2), p) level are given for the gas phase and in solvents of increasing ϵ dielectric constants (IEF/PCM model). Values in parentheses are from HF/6-31G(2d (0.8,0.2), p) calculations. Values in brackets are the corresponding MP2 correlation energy contributions. ΔG values are corrected for the BSSE calculated at the corresponding level of theory.

acetone and −0.1 kcal/mol in water. These values confirm the trends already discussed for the stacked complex. Eventually note the striking parallelism between the gas phase and acetone energy curves.

(d) Comparison of Stacked and T-Shape Complexes. Table 3 collects the HF and MP2 interaction (free) energies in a vacuum and in the different solvents for the most stable stacked and T-shape conformations of the model structures. The results show the importance of the correlation energy contributions to the stabilities of both the stacked and T-shape complexes, for both the protonated and neutral species, in agreement with the results obtained with X-ray structures. The balance between electrostatic and dispersion energy contributions differs however. Indeed, the correlation effect is slightly larger (∼0.5 kcal/mol) for the stacked than the T-shape geometries and is also larger (∼0.8 kcal/mol) for the protonated than for the neutral species.

The calculations seem to indicate that the perpendicular arrangement of imidazole−benzene complexes is energetically favored as well in a vacuum as in a condensed phase, except for water. This result seems in contradiction with those obtained on the basis of the X-ray structures, where the stacked conformations displayed less favorable interaction energies in a vacuum but more favorable interaction free energies in the solvents, on the average. Note that the $\Delta\mathbf{G}$ value of the most favorable of all stacked X-ray structures is lower that that of all stacked model structures; it is however higher than that of the most favorable T-shape model structure.

These results lead us to conclude that the restriction to model or X-ray structures defines two different, but incomplete, samplings of the conformational space, the first because of the neglect of degrees of freedom and the second due to the limited number of "randomly" chosen structures. The latter point also suggests that the T-shaped free energy minima are more narrow than the stacked ones, thereby explaining the higher average interaction free energies of the random (X-ray) structures.

Another point is the minor importance of the electron correlation effect in the IEF−PCM treatment (<0.5 kcal/mol). Nevertheless, this effect increases systematically with ϵ (see values in brackets), which means that HF IEF−PCM calculations tend to overestimate the solvation free energies. Recalling that $\Delta\Delta G_{solv}$ is the sum of an electrostatic $\Delta\Delta G_{el}$ and nonelectrostatic ΔG_{nel} contribution (see Methods), we would also like to stress that the relative magnitude of these contributions is far from being the same for the different complexes. Taking results in acetone as an example, the following values (in kcal/mol) of ($\Delta\Delta G_{el}$, ΔG_{nel}) are found: (9.2, 2.2), (4.7, 2.7), (1.3, 2.6) and (0.7, 3.1) for the protonated T-shape, protonated stacked, neutral stacked and neutral T-shape complexes, respectively. The corresponding ratio $\Delta\Delta G_{el}/\Delta G_{nel}$ thus exhibits the following drastic changes: 4.2, 1.7, 0.2 and 0.5, respectively. These changes essentially arise from $\Delta\Delta G_{el}$ and reflect the strength of the screening of the solute electrostatic interactions by the solvent.

4. Detailed Analysis of His−Ade Interactions. (a) X-ray Structures. Among the 14 His−Ade X-ray pairs, 6 are rather T-shaped and 8 stacked. The positioning of the imidazole moieties above the Ade plane are depicted in Figure 5 and colored according to their interaction free energies $\Delta\mathbf{G}$ in acetone. As in the case of His−Phe, the $\Delta\mathbf{G}$ values are on the average more favorable for the complexes with protonated than with neutral His. Moreover, the imidazole is more favorable above the 6-cycle than above the 5-cycle. Note also that the minimal values of ΔG_{HF}, ΔG_{cor} and $\Delta\mathbf{G}$ are slightly less favorable for the His−Ade pairs than for His−Phe (Tables 1 and 2), but the number of His−Ade pairs in our set is too low to allow firm conclusions.

(b) Model Structures. We focused on stacked His−Ade conformations, slightly more frequent among the X-ray

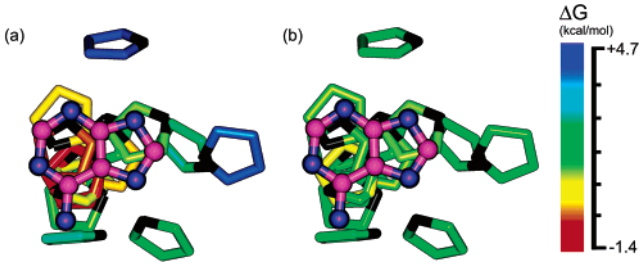

Figure 5. Distribution of the imidazole rings of His residues around the Ade ring systems, observed in the protein−ligand set, as a function of their calculated free energy values in acetone. The Ade bases are superimposed and the imidazole rings have been located "under" the Ade planes, for a better comparison, by projection. The calculated free energy values are colored according to a rainbow ramp where the most favorable conformations are symbolized in red and the less favorable in blue. Carbon and nitrogen atoms of the Ade rings are colored in purple and blue respectively, and Cγ atoms of the imidazole cycle are colored in black. The images were generated using Insight II (Accelrys, Inc.) Structures of the protonated (a) and neutral (b) imidazole rings.

structures. The relative positions of both species were specified by three parameters, the distance d between the β plane and the center of imidazole ring, the orientation angle θ defining the rotation of the $C_{\epsilon 1}-H_{\epsilon 1}$ bond around the imidazole center, and the tilt angle τ defining the rotation of this bond around the perpendicular line passing through the imidazole center (Figure 1c). On the basis of the protein−ligand X-ray structures, we chose to fix d=3.5 Å and explored the other two parameters.

In the first step, we kept both species parallel (τ=0°) and studied the effect of moving the imidazolium parallel with respect to Ade. To explore the entire plane above the Ade system, the orthogonal projection of the imidazole center in the Ade plane was placed at 12 positions, corresponding to the centers of the two Ade rings and to the middles of all bonds constituting these rings. At each position, the imidazole orientation parameter θ was optimized at the vacuum MP2/6-31G(2d (0.8,0.2), p) level of theory. All interaction energies ΔE_{MP2} were found to be comprised between −7 and −12 kcal/mol. Whereas ΔE_{MP2} is rather insensitive to θ (0.9 kcal/mol on the average), as found for the imidazole−benzene complex, it strongly depends on the position of the imidazole center. According to whether the imidazole center is situated above the 5- or 6-cycle of Ade, ΔE_{MP2} is in the range of −7 to −10 kcal/mol or −10 to −12 kcal/mol, respectively. The minimum ΔE_{MP2} corresponds to a position of the imidazole center above the bond between the two cycles. At this position, ΔE_{MP2} reaches −12 kcal/mol for the protonated complex.

In a second step, having fixed the orientation angle θ to its position-dependent optimal value, we have studied the effect of tilting the imidazole. Therefore, we varied the τ angle from 0° to ± 45°, the positive values of τ corresponding to the situations where the positive charge carried by the His approaches the Ade ring (Figure 1c). The calculated ΔE_{MP2} values show that, except for one position, the protonated complex is more stable when τ is comprised between +15 and +30°, which can be explained by the enhancement of the cation−π interaction induced by the approach of the positively charged atoms. One observes that this stabilizing effect is on average 3 times larger when the projection of the imidazole center is situated on the 5-cycle, reaching then −1.8 kcal/mol.

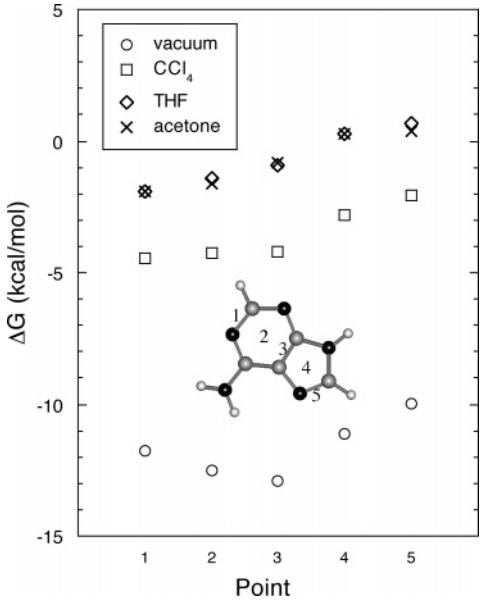

Figure 6. Stacked imidazole−adenine complex: evolution of ΔG in various solvents as a function of the position of imidazole with respect to adenine. The five points considered (see inset of the figure) correspond to the projection of the imidazole center in the adenine aromatic plane α. Parameters θ and τ (see Figure 1c) defining the orientation of imidazole in α and the tilted orientation of α have been optimized at each point.

At this stage, we evaluated the contributions of solvation to the interaction energies by using the IEF−PCM model. Figure 6 shows the interaction energies ΔE_{MP2} and the interaction free energies $\Delta\mathbf{G}$ in CCl_4, THF, acetone and water, when the orthogonal projection of the imidazole center into the aromatic plane is situated at some characteristic positions located on a line across the two aromatic ring centers. At each position, the preferred parameters of θ and τ were adopted. The positions with the lowest free energies are located in the 6-cycle. The interaction (free) energies become less favorable when moving away above the 5-cycle. The existence of a pronounced minimum above the bond common to the 2 cycles distinguishes the gas phase and solvated energy profiles. At this minimum, ΔE_{MP2} reaches −13 kcal/mol. The corresponding $\Delta\mathbf{G}$ values are −4. kcal/mol in CCl_4, −1 kcal/mol in THF and acetone and 1 kcal/mol in water. The contribution of the positive charge to the stability of the complex are −10 kcal/mol in a vacuum, −5 kcal/mol in CCl_4, −2. kcal/mol in THF and acetone, and −0.5 kcal/mol in water.

The correlation effects contribute significantly to the stability of the complex, like for His−Phe. The ΔE_{cor} value in the vacuum amounts to −7 and −9 kcal/mol for points 1 and 3 respectively, which corresponds to 56 and 68% of the gas phase $\Delta\mathbf{E}$ values, respectively (−12 and −13 kcal/mol).

The excess of correlation at point 3 clearly induces the observed gas-phase minimum. The large values of the solvation energies (10 and 12 kcal/mol for points 1 and 3 in acetone) compensate however the correlation energy contributions, with as a result a shift of the minimum from points 3 to 1. As in the case of the benzene, one finds that the balance between the $\Delta\Delta G_{el}$ and ΔG_{nel} solvation terms follow the changes occurring in the electronic structure of the solute. At positions 1 and 3 ($\Delta\Delta G_{el}$, ΔG_{nel}) are (in kcal/mol) (8.6, 1.7) and (10.7, 1.4) respectively, the corresponding ratio $\Delta\Delta G_{el}/\Delta G_{nel}$ being 5 and 8.

Note also the systematic overestimation of the solvation energies by HF IEF−PCM calculations (0.2 and 0.5 kcal/mol at point 1 and 3 respectively), as in the Phe−His case. The larger discrepancy for point 3 probably reflects the enhancement of the correlation effect pointed out above for this geometry

5. Biological Role of His−Aromatic Interactions. There are several arguments in favor of the biological role of His−aromatic interactions in proteins. First, these interactions are often found in catalytic sites. This is the case, among the proteins of our data set, for His48-Tyr52 in phospholipase A2 (4p2p),[43] His54-Trp16 in D-xylose isomerase (7xia),[44] His195-Tyr25 in chloramphenicol acetyltransferase (3cla),[45] and His57-Phe54 in serine proteinases (2apl and 2sga).[46] All these His−aromatic pairs were determined as essential to the enzymatic reactions, and the aromatic partner was proven to participate indirectly in the stabilization of the catalytic site.[43,47,48] This suggests that His−aromatic interactions are key components of the catalytic mechanism of these enzymes, either in providing critical constraints on the precise conformation of the catalytic site or in participating in transition-state stabilization.

Second, His−aromatic cation−π interactions are recurrently observed at crucial positions in 3D domain swapping proteins.[49] Indeed, using the data set of domain swapping proteins of Dehouck et al.,[49] we find several His−π pairs located in the hinge loop of either the closed or swapped form, for example, His116-Tyr102 in the closed form of Spo0A-N (1qmp) and His384-Phe530 in swapped diphtheria toxin (1ddt). In other proteins, the same interaction is observed in the two forms, involving residues that are part of the same chain in the closed form but of two different chains in the swapped form and connecting the chain pieces across the closed interface. Such interactions seem thus to zip up the closed interface of domain swapped oligomers. Examples of this type are His86-Tyr42 in cystatine C (1g96), His90-Trp49 in cyanovirin-N (2ezm, 3ezm) and His18-Trp94 in barnase (1brn, 1yvs). We computed the free energy in acetone for the His18-Trp94 pair of the closed form of barnase. As usual, the protonated form appears more stable ($\Delta\mathbf{G}=-1.6$ kcal/mol in acetone) than the two neutral forms ($\Delta\mathbf{G}=+0.9$ and $+1.7$ kcal/mol). So, a positive charge on the His ring stabilizes barnase by about 2.5 to 3.3 kcal/mol, in agreement, though slightly overestimated, with the experimental values of 1 to 2 kcal/mol.[50,51]

Hence, because of their frequent occurrence at specifically important positions for 3D domain swapping and their stabilizing properties, His−aromatic interactions can be expected to contribute in a pH-dependent way to the stability of the closed and/or swapped forms and to modulate the conversion between the two forms.

Third, at the interface between protein and DNA, stair motifs involving simultaneously cation−π, H-bond and base stacking interactions are often observed.[52] They consist of two successive bases along the DNA stack, one in cation−π interaction with an amino acid side chain that carries a net or partial positive charge, and the other H-bonded with the same side chain. In some proteins, like zinc fingers, several successive stair motifs are observed. When His is considered as potential cation/δ^+−π partner in addition to Arg, Lys, Asn and Gln, the number of succesive stairs is sometimes impressive. In the zinc finger 1ag1, in particular, we find the multiple stair: Thy(B5)$\therefore$His(A149)$\vee$Gua(B6)$\therefore$Arg(A146)$\vee$Gua(B7)$\therefore$Arg(A124)$\vee$Gua(B8)$\therefore$Asn(A121)$\vee$Cyt(B9), where $\therefore$ denotes cation−π and $\vee$ H-bond. The exact structural or functional role of these motifs is not known, but their recurrent occurrence suggests that they play a role in protein−DNA recognition.

Conclusions

Our quantum chemistry analysis of both X-ray and model structures of His−aromatic interactions concur to indicate that the positively charged complex is more stable than the neutral one by about -6 to -9 kcal/mol in gas phase, -4 to -5 kcal/mol in CCl_4, -2 to -3 kcal/mol in THF, -1 to -3 kcal/mol in acetone and 0 to -2 kcal/mol in water (Tables 1 and 3 and Supporting Information). Our results are thus in agreement with the experimental finding that in protein environments, best mimicked by THF and acetone, the protonation of His has a stabilizing effect of about 1 to 2 kcal/mol.[18]

The substantial difference in interaction free energy between the protonated and neutral forms of His−aromatic complexes, and the property of His to become protonated near physiological conditions, makes this type of interactions likely to be involved in the biological function. This hypothesis is supported by the frequent occurrence of His−aromatic interactions in 3D domain swapping proteins, at protein−DNA interfaces, and among catalytic residues.

Our results do not allow to clearly settle whether His−aromatic interactions in protein environments are more favorable in T-shape or stacked arrangement. Both conformations are roughly equally frequently observed in X-ray structures, and their computed relative stabilities are similar, with some differences according to the precise geometry, the type of partners and the solvent. However, some trends can be distinguished. In particular, T-shaped structures are much more stable in gas phase, and this advantage is reduced in the solvents considered. In water, the stacked conformation appears even slightly more favorable. We can thus tentatively conclude that T-shaped conformations are more favorable in gas phase and nonpolar solvents and that the stacked conformations become more favorable as the polarity of the solvent increases, in agreement with molecular dynamics studies[53] and the more frequent observation of stacked conformations at the protein surface. Finally note that the

favorable nature of stacked conformations in protein environments must not be related with the higher accessibility of H-bond forming groups as sometimes stated,[54] at least for His−aromatic complexes, since the two groups ($N_{\delta 1}$ and $N_{\epsilon 2}$) in the His cycle that are able to form H-bonds are roughly evenly accessible in the stacked and T-shape conformations. We verified indeed that the number of His involving H-bonds in the complexes studied are independent of their conformation.

Acknowledgment. The Communauté Française de Belgique (Action de Recherche Concertée), the Cost P9 Action, the European Community and The Belgian National Fund for Scientific Research are acknowledged for support. Eric Buisine is thanked for stimulating discussions.

Supporting Information Available: The set of 89 His−aromatic interactions identified (Table S1) and solvation free energies in acetone for the 89 X-ray His−aromatic cation−π pairs (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dougherty, D. A. *Science* **1996**, *271*, 163−168.

(2) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303−1324.

(3) Zacharias, N.; Dougherty, D. A. *Trends Pharmacol. Sci.* **2002**, *23*, 281−287.

(4) Burley, S. K.; Petsko, G. A. *FEBS Lett.* **1986**, *203*, 139−143.

(5) Flocco, M. M.; Mowbray, S. L. *J. Mol. Biol.* **1994**, *235*, 709−717.

(6) Gallivan, J. P.; Doughert, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459−9464.

(7) Minoux H.; Chipot C. *J. Am. Chem. Soc.* **1999**, *121*, 10366−10372.

(8) Gromiha, M. M.; Thomas, S.; Santhosh, C. *Prep. Biochem. Biotechnol.* **2002**, *32*, 355−362.

(9) Gromiha, M. M. *Biophys. Chem.* **2003**, *103*, 251−258.

(10) Tsou, L. K.; Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 14917−14921.

(11) Zhong, W.; Gallivan, J. P.; Zhang, Y.; Li, L.; Lester, H. A.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12088−12093.

(12) Peterson, E. J.; Choi, A.; Dahan, D. S.; Lester, H. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 12662−12663.

(13) Wintjens, R.; Liévin, J.; Rooman, M.; Buisine, E. *J. Mol. Biol.* **2000**, *302*, 395−410.

(14) Biot, C.; Wintjens, R.; Rooman, M. *J. Am. Chem. Soc.* **2004**, *126*, 6220−6221.

(15) Biot, C.; Buisine, E.; Kwasigroch, J. M.; Wintjens, R.; Rooman, M. *J. Biol. Chem.* **2002**, *277*, 40816−40822.

(16) Biot, C.; Buisine, E.; Rooman, M. *J. Am. Chem. Soc.* **2003**, *125*, 13988−13994.

(17) Wouters, J. *Protein Sci.* **1998**, *7*, 2472−2475.

(18) Loewenthal, R.; Sancho, J.; Fersht, A. R. *J. Mol. Biol.* **1992**, *224*, 759−770.

(19) Fernández-Recio, J.; Vazquez, A.; Civera, C.; Sevilla, P.; Sancho J. *J. Mol. Biol.* **1997**, *267*, 184−197.

(20) Fernández-Recio, J.; Romero, A.; Sanjo, J. *J. Mol. Biol.* **1999**, 290, 319−330.

(21) Feinberg, H.; Torgersen, D.; Drickamer, K.; Weis, W. I. *J. Biol. Chem.* **2000**, *275*, 35176−35184.

(22) Williams, S.; Bledsoe, R. K.; Collins, J. L.; Boggs, S.; Lambert, M. H.; Miller, A. B.; Moore, J.; McKee, D. D.; Moore, L.; Nichols, J.; Parks, D.; Watson, M.; Wisely, B.; Willson, T. M. *J. Biol. Chem.* **2003**, *278*, 27138−27143.

(23) Li, H. L.; Galue, A.; Meadows L.; Ragsdale D. S. *Mol. Pharmacol.* **1999**, *55*, 134−141.

(24) Okada, A.; Miura, T.; Takeuchi, H. *Biochemistry* **2001**, *40*, 6053−6060.

(25) Takeuchi, H.; Okada, A.; Miura, T. *FEBS Lett.* **2003**, *552*, 35−38.

(26) Inoue, Y.; Nakamura, N.; Inagami, T. *J. Hypertens.* **1997**, *15*, 703−714.

(27) Chen, X. Z.; Steel, A.; Hediger, M. A. *Biochem. Biophys. Res. Commun.* **2000**, *272*, 726−730.

(28) Argiriadi, M. A.; Morisseau, C.; Hammock, B. D.; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10637−10642.

(29) Spiwok, V.; Lipovova, P.; Skalova, T.; Buchtelova, E.; Hasek, J.; Kralova, B. *Carbohydr. Res.* **2004**, *339*, 2275−2280.

(30) Wintjens, R.; Rooman, M.; Wodak, S. J. *J. Mol. Biol.* **1996**, *255*, 235−253.

(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.

(32) Kabsch W. *Acta Crystallogr. A* **1978**, *34*, 827−828.

(33) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618−622.

(34) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(35) Wintjens, R.; Biot, C.; Rooman, M.; Liévin, J. *J. Phys. Chem. A* **2003**, *107*, 6249−6258.

(36) Tomasi, J.; Mennucci, B.; Cancès, E. *J. Mol. Struct. (THEOCHEM)* **1999**, *464*, 211−226.

(37) Meyer E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 1210−1250.

(38) Meyer E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 4120.

(39) Wlodawer, A.; Walter, J.; Huber, R.; Sjolin, L. *J. Mol. Biol.* **1984**, *180*, 301−329.

(40) Hunter, C. A. *Chem. Soc. Rev.* Meldola Lecture, **1994**, 101−109.

(41) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104−112.

(42) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2004**, *120*, 647−659.

(43) Soares, A. M.; Giglio, J. R. *Toxicon* **2003**, *42*, 855−868.

(44) Whitaker, R. D.; Cho, Y.; Cha, J.; Carrell, H. L.; Glusker, J. P.; Karplus, P. A.; Batt, C. A. *J. Biol. Chem.* **1995**, *270*, 22895−22906.

(45) Lewendon, M. R.; Murray, I. A.; Shaw, W. V. *Biochemistry* **1994**, *33*, 1944−1950.

(46) Fuhrmann, C. N.; Kelch, B. A.; Ota, N.; Agard, D. A. *J. Mol. Biol.* **2004**, *338*, 999−1013.

(47) Lambeir, A. M.; Lauwereys, M.; Stanssens, P.; Mrabet, N. T.; Snauwaert, J.; van Tilbeurgh, H.; Matthyssens, G.; Lasters, I.; De Maeyer, M.; Wodak, S. J.; Jenkins, J.; Chiadmi, M.; Janin, J. *Biochemistry* **1992**, *31*, 5459−5466.

(48) Murray, I. A.; Lewendon, A.; Shaw, W. V. *J. Biol. Chem.* **1991**, *266*, 11695−11698.

(49) Dehouck, Y.; Biot, C.; Gilis, D.; Kwasigroch, J. M.; Rooman, M. *J. Mol. Biol.* **2003**, *330*, 1215−1225.

(50) Sali, D.; Bycroft, M.; Fersht, A. R. *Nature* **1988**, *335*, 740−743.

(51) Loewenthal, R.; Sancho, J.; Fersht, A. R. *Biochemistry* **1991**, *30*, 6775−6779.

(52) Rooman, M.; Liévin, J.; Buisine, E.; Wintjens, R. *J. Mol. Biol.* **2002**, *319*, 67−76.

(53) Gervasio, F. L., Chelli, R., Marchi, M., Procacci, P., Scettino, V. *J. Phys. Chem. B* **2001**, *105*, 7835−7846.

(54) Mitchell, J. B.; Nandi, C. L.; McDonald, I. K.; Thornton, J. M.; Price, S. L. *J. Mol. Biol.* **1994**, *239*, 315−331.

CT049875K