J. Phys. Chem. A **1997,** *101,* 7885−7892

Ab Initio Study of 4(5)-Methylimidazole in Aqueous Solution

G.-S. Li,[†] M. F. Ruiz-López,* and B. Maigret

Laboratoire de Chimie théorique, URA CNRS 510, Institut Nancéien de Chimie Moléculaire,
Université Henri Poincaré-Nancy I, BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France

Received: April 24, 1997; In Final Form: July 18, 1997[⊗]

The properties of the two tautomers of 4(5)-methylimidazole in aqueous solution have been studied through theoretical calculations using ab initio methods and various solvent models. We focus this study on the environmental effects due to cooperative solute−solvent interactions. The modifications of the tautomeric equilibrium and pK_a value through hydrogen-bond formation and electrostatic solvation effects are discussed. The computations show that the polarizable continuum model of the solvent is able to give reasonable results for these properties. The prediction of absolute values of pK_a is difficult, but relative values are reproduced quite well. We also consider a discrete-continuum model for the solution, but we show that this model leads to slightly stable (or unstable) complexes. Our results stress the importance of nonadditive energy contributions, which are presumably fundamental in order to explain the mechanism of several biological processes involving histidine residues.

1. Introduction

Nowadays, theoretical chemistry provides very accurate results for small and isolated system properties. The fast development of combined quantum mechanics and molecular mechanics techniques[1] is rendering it possible, on the other hand, to afford the study of very large molecules, such as proteins, and other complex systems such as solution and surface processes, etc. These techniques will certainly allow the realistic study of enzymatic reactions soon. Meanwhile, the ab initio study of small or medium size model systems is opportune since it can be carried out at high computational levels and can be used in turn to judge the approximations made in a simplified theoretical model.

We describe in this paper some fundamental properties of simple imidazole derivatives in different molecular environments. Imidazole is a five-membered heterocyclic system with basic and acid properties and is the key functional group of histidine. It plays an important role in many biochemical processes such as proton transfer[2] at the active site of enzymes, and therefore it has received a great deal of attentions over the years.[3]

Many experimental works[4] have been reported on the study of properties of imidazole and imidazole derivatives. The tautomer ratio of substituted imidazole is of great biophysical importance[5] but difficult to study experimentally. For instance, the equilibrium constant of 4(5)-methylimidazole, first measured by Pyman[6] by separating *N*-methylated 4(5)-methylimidazole tautomer, presents a range of variation going from 0.45 to 1.5,[6−8] which corresponds to a small free energy change in the tautomer equilibrium that may be positive or negative. The influence of the medium is certainly important.

On the other hand, the knowledge of the pK_a value of protein residues is fundamental for understanding the mechanisms of enzyme-catalyzed reactions, especially those involving proton transfer.[2,9] Unfortunately, the experimental determination of pK_a's is made difficult by uncertainties in several factors such as pH dependence,[10] ionic strength dependence,[11] ion-pairing and aggregation effects,[12] etc. For this reason, the accurate theoretical determination of absolute values of pK_a would be very useful. However, there are considerable problems in achieving this, as we show below.

Many theoretical studies have been also reported.[8,13−30] Computations for isolated molecules at the semiempirical AM1, PM3, and MNDO levels[27,28] show 5-methylimidazole to be slightly more stable than 4-methylimidazole, in line with small basis sets (STO-3G and 3-21G) ab initio calculations.[17,27] However, ab initio results at higher levels[17,26,29−30] predict a small preference for 4-methylimidazole. Other authors have investigated the influence of the solvent on the equilibrium constant, in particular in aqueous solution.[8] The theoretical evaluation of the pK_a has also received some attention.[18]

We focus the present study on the influence of the solvent, water in our case, on the properties of 4- and 5-methylimidazole (MeIm), in an attempt to analyze those factors that can play a role in explaining the biochemical activity of histidine. Several points will be examined. The effect of hydrogen bonds with water is compared to that of electrostatic solvation, and cooperative effects are discussed. The computation of the hydration energy using either a continuum or a discrete-continuum model is then presented. The effect of solvation on the tautomeric equilibrium is also analyzed. Finally, we present some results on pK_a values and discuss the limitations of simple solvent models in such kinds of computations.

2. Methodology

Ab initio calculations in the gas phase have been carried out at two computational levels. Level A corresponds to computations at the MP2/6-31G* level both for the geometries and total energy of the systems. Level B corresponds to computations at the MP2/6-311++G** level for the total energy at MP2/6-31G* optimized geometries. This last computational scheme is also referred as MP2/6-311++G**//MP2/6-31G* calculations. The Gaussian 94 package[31] have been used. The choice of the basis sets and correlation level has been made on the basis of previous calculations for these systems.[29] The MP2/6-31G* level was shown to yield good results for the geometry of the systems and also for the tautomeric equilibrium.[8,29] The MP2/6-311++G** level is necessary to compute protonation energies and pK_a's, as shown in our previous study.[29] Following

---

[†] On leave from: Department of Chemical Physics, University of Science and Technology of China, 230026 Hefei, P. R. China.

[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1997.

S1089-5639(97)01390-X CCC: $14.00

the optimizations at the MP2/6-31G* level, analytical frequency calculations are done to determine the nature (minimum or saddle point) of the stationary points found and to obtain contributions to the free energy that has been evaluated using standard procedures[32] at 298.15 K and 1 atm conditions.

Computations in aqueous solution have been carried out using continuum and discrete-continuum models. In the continuum model, the solute is placed in a general cavity fitting as best as possible the molecular shape. The cavity is embedded in a polarizable dielectric continuum, the relative dielectric permittivity of which is taken to be equal to that of water at 298 K (78.4). The continuum is polarized by the solute charge distribution, and the reaction field resulting from this polarized continuum polarizes in turn the solute's molecule. The relaxation of the system and the free energy of solvation is obtained by doing a self-consistent reaction field calculation[33] in which the Fock operator includes the solute−solvent electrostatic interaction that is obtained using a multipole development of the reaction field potential:

$$F_{\mu\nu} = F^{\circ}_{\mu\nu} + \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \langle R_l^m \rangle \langle \mu | M_l^m | \nu \rangle \tag{1}$$

Here $F^{\circ}_{\mu\nu}$ is the operator for the isolated molecule, $\langle R_l^m \rangle$ are the components of the polarized solvent potential, electric-field, electric-field gradient, etc., at a point inside the cavity (normally the center of positive charges) and M_l^m are the components of the multipole moments of the solute of order (l,m) computed at the same center. Development 1 converges rapidly in standard cases, and one needs to compute a few terms only (we compute here terms up to the sixth order).

To evaluate the other terms entering in the solvation energy (referred here as the nonelectrostatic contribution), we employ a semiempirical formula. Thus, the total solvation free energy is written as

$$\Delta G_{\text{sol}} = \Delta G_{\text{sol}}^{\text{EP}} + \Delta G^{\text{NE}}{}_{\text{sol}} \tag{2}$$

$\Delta G_{\text{sol}}^{\text{EP}}$ includes the electrostatic plus polarization energy. $\Delta G_{\text{sol}}^{\text{NE}}$ accounts for all the other contributions to the solvation free energy (arising mainly from dispersion and cavitation energies) and is obtained by[34]

$$\Delta G_{\text{sol}}^{\text{NE}} = \sum_i \gamma_i S_i \tag{3}$$

where S_i is the accessible surface area of atom i in the solute's molecule and γ_i are parameters depending on atom type that have been determined at a given computational level in order to fit as well as possible the experimental free energy of solvation of a reference set of molecules. Accessible surface areas are calculated adding the solvent radius (1.4 Å for water) to the van der Waals ones, using an analytical algorithm.[34] Note that the experimental free energy of solvation contains other small contributions (such as variation of zero-point energy and thermal corrections) that are not explicitly considered in the calculation but are to some extent included in the parametric term $\Delta G_{\text{sol}}^{\text{NE}}$. Computations of ΔG_{sol} have been done using the program SCRFPAC[35] updated to the Gaussian 94[31] environment.

In the case of the solvated systems, the geometry has been optimized for 4(5)-MeIm and for the corresponding protonated form (MeImH$^+$). However, it was not possible to systematically optimize the geometry of the complexes with water molecules in a dielectric environment. Thus, the geometry of the complex formed by 4-MeIm with a single water molecule (imidazole being the proton acceptor) was optimized only, allowing us to discuss electrostatic solvent effects on the geometry of the complexes and cooperativity phenomena. Then, the vibrational frequencies were computed numerically using analytical first derivatives of the energy.

An important question when studying hydrogen-bonded systems is the role played by BSSE (basis set superposition errors) both in the equilibrium geometry of the complex and in the association energy. We have not carried out a systematic evaluation of BSSE, but some estimations are given below in order to quantify their significance in the present case. These computations have been carried out at level A using the counterpoise method[36] in which the BSSE correction is given by

$$\text{BSSE(A−B)} = E_{\text{a}}(\text{A}) + E_{\text{b}}(\text{B}) - E_{\text{ab}}(\text{A}) - E_{\text{ab}}(\text{B}) \tag{4}$$

where E_λ(Z) represents the energy of system Z (A or B) at the complex A−B geometry with the basis set λ associated to it (a or b) or to the complex (ab).

3. Results and Discussion

3.1. Cooperative Effects. We start the discussion of our results by making an analysis of cooperative effects in MeIm−water hydrogen-bonded complexes. The cooperativity phenomenon was described by Frank and Wen[37] who introduced the concept of "flickering clusters". It consists in the strengthening of hydrogen bonds with increasing number of H-bonds in an agregate. One of the characteristics of cooperativity is the lengthening of the donor Y−H bonds. It has been shown experimentally that such nonadditive effects cannot be neglected.[38] A pioneering ab initio calculation was reported by Del Bene and Pople,[39] demonstrating the nonadditivity of hydrogen-bond energy in the water trimer. From that time, many quantum-mechanical works have been devoted to its study, mainly for water clusters but also for other systems. It is not possible to provide here a detailed list of references in the literature, but see ref 40 for some recent works in which further relevant bibliography may be found. Cooperativity between hydrogen-bonding and electrostatic interactions with the environment is another interesting aspect of the phenomenon. Thus, the structure of the water dimer in liquid water has been also described using a discrete-continuum approach.[41] Actually, a better understanding of cooperativity is crucial in order to develop sophisticated models for studying complex systems such as liquids or macromolecules. Apart from the lengthening of the donor Y−H bonds, a number of quantities have been employed in order to probe the magnitude of the effect. We shall consider here the Y−H bond critical points[42] and Mayer indices[43] as well as the cooperativity factor proposed by Kleeberg et al.[44] The later is based on the work reported by Huyskens[45] and represents the ratio between the shift of the Y−H stretch in the trimer (the dihydrated complex here) and the dimer (the monohydrated complex here). We also extend this concept to the case of an electrostatically solvated dimer.

Methylimidazole−Water Hydrogen Bonds. We first consider the structure of the hydrogen-bonded complexes formed by 4- and 5-methylimidazole with water molecules. The imidazole ring is a bifunctional group that may behave either as a proton donor or a proton acceptor. Therefore, we have examined hydrogen bonding of methylimidazole with one or two water molecules. In Tables 1 and 2, we summarize the NO distances and association energies obtained for the mono- and dihydrated complexes, respectively. Figure 1 displays the geometry of the systems.

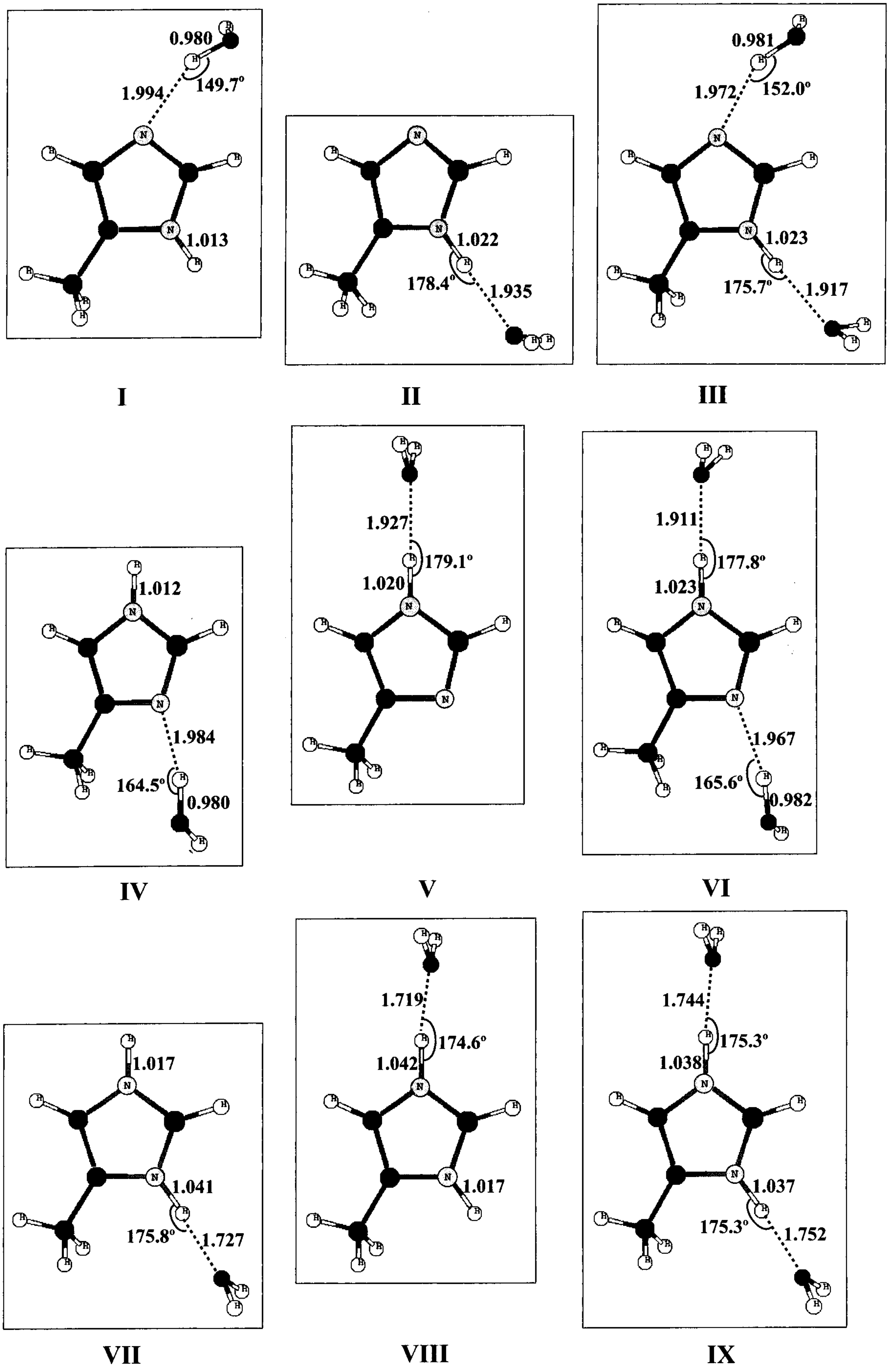


Figure 1. Geometry parameters for hydrogen-bonded complexes of 5-, 4-, and protonated methylimidazole with one or two water molecules. Distances in angstroms and angles in degrees. Values are obtained at the MP2/6-31G* level.

The donating HY distances (Y=N or O) may be compared to those obtained for the isolated systems. The NH bond lengths for isolated 5-MeIm and 4-MeIm are 1.013 and 1.012 Å, respectively. The OH bond length in the isolated water

TABLE 1: Computed Properties of the Hydrogen-Bonded Complexes of 4(5)-MeIm with One Water Molecule[a]

system	d_{N-O}	E_{assoc} A	E_{assoc} B	ΔZPE A	D_0 A	D_0 B
proton acceptor						
5-MeIm, **I**	2.883	−9.76	−7.65	2.23	−7.53	−5.42
4-MeIm, **IV**	2.941	−9.82	−7.78	2.25	−7.57	−5.53
proton donor						
5-MeIm, **II**	2.957	−8.34	−7.45	1.47	−6.87	−5.98
4-MeIm, **V**	2.948	−8.20	−7.47	1.19	−7.01	−6.28

[a] The geometries (see Figure 1) have been optimized at the MP2/6-31G* level. The association energies are computed at the MP2/6-31G* level (A) and MP2/6-311++G**//MP2/6-31G* level (B). D_0 is the association energy after including ZPE corrections. Distances in Å and energies in kcal/mol.

molecule is 0.969 Å. Thus, the NH bonds are lengthened by ~0.01 Å when the proton is coordinated to a water molecule. Similarly, the OH bonds are lengthened by a comparable amount when the water molecule coordinates MeIm. The hydrogen bonds are not far from linear in the case of N−H···O arrangements but are substantially bent for N···H−O ones, especially in 5-MeIm−water complexes. Owing to that, the NO distances in 5-MeIm−water complexes (imidazole being the proton acceptor) are notably shorter than in other cases.

Association energies are very similar for 4-MeIm and 5-MeIm, the differences being smaller than ±0.2 kcal/mol in all cases. These quantities are quite sensitive to the computational level, as expected. The largest association energy for monohydrated complexes is obtained when the imidazole ring behaves as a proton acceptor group, although at the highest level of theory the energy difference is slight. Note however that zero-point energy (ZPE) corrections are larger when MeIm behaves as proton acceptor. Hence, when $D_0 = E_{assoc.} + \Delta ZPE$ is considered, the monohydrated complexes in which MeIm behaves as proton acceptor are less stable at level B. ZPE corrections for the dihydrated complexes are very close, and therefore they do not modify the relative stability of the tautomers.

One must note that in Tables 1 and 2 we have not considered BSSE's, and it is interesting to estimate their effect on association energies. Let us first consider the monohydrated systems. Our computations (at level A) show that BSSE is larger for the complexes in which methylimidazole behaves as proton acceptor. For instance, for 4(5)-MeIm, BSSE is 1.9 (1.9) kcal/mol for **I** (**IV**) and only 0.8 (0.7) kcal/mol for **II** (**V**). Therefore, after BSSE correction, the complexes with MeIm being the proton-donating group become slightly more stable at level A. BSSE is obviously larger for dihydrated complexes but the differences between 4- and 5-MeIm are negligible (2.7 kcal/mol for both **III** and **VI**).

The cooperativity phenomenon is manifested by the variation of the N−H bond distances going from isolated MeIm to mono- and dihydrated complexes. This quantity increases by 0.009 Å from 5-MeIm to **II** and by 0.001 Å from **II** to **III**. It increases by 0.006 Å from 4-MeIm to **V** and by 0.003 Å from **V** to **VI**. The cooperative effect may be also illustrated through the variation of the corresponding Mayer bond orders, which decrease in going from the monohydrated to the dihydrated complex showing that the intermolecular interaction is enhanced. For instance, the NH bond order changes from 0.805 in isolated 4-MeIm to 0.696 in **V** and 0.687 in **VI**. Comparable results are found for 5-MeIm.

As said above, cooperativity may be illustrated also through the variation of the frequency corresponding to the donating HY stretching mode after H-bonding. We take the example of 4-MeIm. The NH frequencies computed for 4-MeIm and for the complexes **V** and **VI** are 3671.7, 3552.7, and 3515.4 cm^{-1} respectively. Therefore, the formation of a second hydrogen bond in **VI** enhances the effect of the first one and weakens the NH bond, the cooperativity factor being $\Delta\nu(\mathbf{VI})/\Delta\nu(\mathbf{V}) = 156.3/119.0 = 1.31$.

The analysis of the electronic densities at the bond critical points, ρ_c, is interesting. This property is related to the strength of a bond[42] and has been used to investigate the cooperative effect in water trimers[46] and in solvated water dimers,[41] for instance. The magnitude of ρ_c for the N···H hydrogen bond in **IV** and **VI** increases from 0.0270 to 0.0281 e/au^3, whereas the corresponding quantities for the donating OH bond are 0.3290 and 0.3271 e/au^3. Thus, proton donation from water to imidazole is favored through formation of a second hydrogen bond of imidazole with water.

Effects of a Dielectric Surrounding. We discuss now the effect of a dielectric surrounding on the structure of a methylimidazole−water H-bonded complex. For this study we have selected the complex formed by 4-MeIm with a water molecule, the imidazole ring being the proton−acceptor. The gas-phase optimized geometry is that given in Figure 1 (structure **IV**). A geometry optimization in the presence of a dielectric polarizable continuum with relative dielectric permittivity equal to that of water (78.4) is now carried out, and the final geometry is represented in Figure 2 (in the geometry optimization calculation, development (1) was limited to the dipole moment).

The electrostatic interactions of the complex with the polarized continuum induce changes in the structure of the solute that may be associated to a cooperative effect. Such an effect is quite remarkable in the present case. The N····H bond is substantially shorter and the O−H bond longer in the solvated complex. This change is even larger than that presented above when a second hydrogen bond is formed by the imidazole ring with another water molecule. The Mayer OH bond order analysis leads to the same conclusions. This quantity changes from 0.680 in the isolated monohydrated complex (**IV**) to 0.664 in the electrostatically solvated complex. The hydrogen bond appears to be strengthened through the effect of a polarizable dielectric environment. This is also confirmed by the analysis of vibration modes and bond critical point densities. Taking the symmetric OH stretching mode of water, the frequencies vary from 3774.8 cm^{-1} in free water to 3629.6 cm^{-1} in the gas-phase complex **IV** and 3529.6 cm^{-1} in the electrostatically solvated complex. We can compute a cooperative factor equivalent to that derived above by considering the frequency shifts free $H_2O \rightarrow$ gas-phase complex and free $H_2O \rightarrow$ solvated complex. This gives 245.2/145.2 = 1.69, which is higher than the factor computed for NH after hydrogen-bond formation of methylimidazole with one or two water molecules. The density of the N···HOH hydrogen-bond critical point increases from 0.0270 to 0.0294 e/au^3 through the effect of solvation whereas for the donating OH bond this property decreases from 0.3290 to 0.3209 e/au^3.

3.2. Solvation Energy. The effect of a dielectric solvent on the geometries of 4- and 5-methylimidazole is summarized in Figure 3. Small modifications of the gas-phase geometry through the effect of solvation are predicted. In particular, the N−H bond lengths change very little compared to the effect produced by a water molecule bonded directly to N−H (compare Figures 1 and 3).

In Table 3, we present the computed solvation energies and we compare them with available experimental data. The electrostatic solvation energy of 5-MeIm is slightly larger than

TABLE 2: Computed Properties of the Hydrogen-Bonded Complexes of 4(5)-MeIm with Two Water Molecules[a]

	d_{N-O}		E_{assoc}		ΔZPE	D_0	
system	in NH···O	in N···HO	A	B	A	A	B
5-MeIm, **III**	2.939	2.875	−18.79	−15.69	3.76	−15.03	−11.93
4-MeIm, **IV**	2.933	2.928	−18.85	−15.85	3.75	−15.10	−12.10

[a] The geometries have been optimized at the MP2/6-31G* level (see also Figure 1). The association energies are computed at the MP2/6-31G* level (A) and MP2/6-311++G**//MP2/6-31G* Level (B). D_0 is the association energy after including ZPE corrections. Distances in Å and energies in kcal/mol.

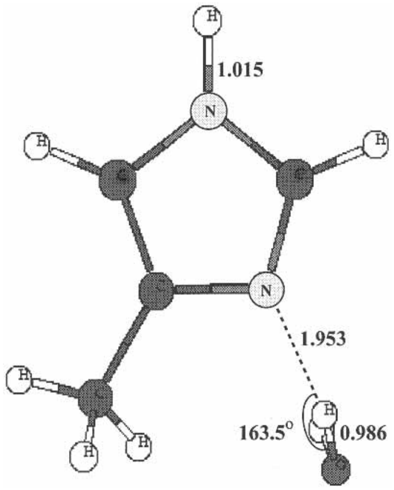


Figure 2. Optimized geometry of the 4-methylimidazole:H_2O complex in aqueous solution. Distances in angstroms and angles in degrees. Values are obtained at the MP2/6-31G* level.

that of 4-MeIm which is due in part to its larger dipole moment, as can be seen from the values presented in Table 4. Computations at the MP2/6-311++G**//MP2/6-31G* level lead to solvation energies a little larger due mainly to the fact that the polarizability of the molecules increases by increasing the basis set although this is not the only important factor. In Table 4, one may see that there is a strong polarization of the molecules when going from gas phase to solution. This table also contains the total volume of the cavities used. The polarization is essentially due to electronic effects since the geometries of the species change very little through the electrostatic interactions with the continuum. The nonelectrostatic contribution for these isomers is a positive quantity which means that the hydrophobic contribution, due essentially to the cavitation term, predominates over stabilizing factors due mainly to the dispersion energy. The values are slightly larger for 5-MeIm. As a result, the total solvation energies predicted for 4-MeIm and 5-MeIm are very close, the difference of −0.1 kcal/mol obtained at the largest computational level being below the accuracy of the theoretical model.

Our results may be compared to those reported previously in the literature. Recently, Luque et al.[30] used a similar continuum approach and predicted solvation energies of −6.6 and −7.4 kcal/mol for 4- and 5- methylimidazole, respectively. These values increased to −7.3 and −8.1 kcal/mol respectively when the geometries of the molecules were relaxed in the continuum. These computations were carried out using optimized geometries at the RHF/6-31G** level, and the difference with ours in solvation energy may be due to difference in computational level and cavity definition. Other theoretical evaluations have been done.[8,22−23] The difference between the solvation energy of 4- and 5-MeIm reported by Luque et al.[30] is −0.8 kcal/mol, and molecular dynamics simulations reported by Worth et al.[22] gave −0.2 kcal/mol.

The computed solvation energy for these molecules is in all cases notably smaller than the reported experimental value, which is −10.27 kcal/mol.[4a] It must be emphasized that in our computations, the nonelectrostatic term is obtained with the use of parameters[34] derived from calculations of the electrostatic solvation energy at the RHF/6-31G* level, which are not necessarily good for computations at the MP2/6-311++G** level. One could also argue that part of the difference between theoretical and experimental solvation energies lies in the omission of explicit hydrogen bonds. To study this hypothesis, we have carried out computations using a mixed discrete-continuum model. The results are presented below, but one should keep in mind that the main effect of hydrogen-bond formation is due to electrostatic interactions, the average value of which is accounted for in the dielectric continuum models.

In the discrete-continuum model, the solvation process may be described by considering the following steps. First, a limited number of solvent molecules are extracted from the liquid, which requires a positive work equal to the vaporization energy. Then, these solvent molecules and the solute interact to yield a gas-phase complex. Finally, the complex is transferred from the gas phase to the dielectric continuum. Accordingly, the hydration free energy of a species M within the discrete-continuum model may be written as[47]

$$\Delta G_{sol}(M) = n\Delta G_{vap}(H_2O) + \Delta G_{assoc.}(M(H_2O)_n) + \Delta G_{sol}(M(H_2O)_n) \quad (5)$$

The vaporization energy of water $\Delta G_{vap}(H_2O)$ may be taken from the experimental value, 6.3 kcal/mol.[48] The remaining terms in eq 5 are the free energy change for the process in which the complex $M(H_2O)_n$ is formed from the separated molecules (association free energy, $\Delta G_{assoc.}(M(H_2O)_n)$) and the solvation free energy of the same complex, $\Delta G_{sol}(M(H_2O)_n)$, i.e., the free energy for the transfer of the complex from gas phase to the aqueous solution. The free energy of association may be computed from the association energies in Tables 1 and 2 by adding the corresponding zero-point energy, thermal corrections and entropic contributions. Such computations predict the complexes to be unstable with respect to the separated molecules at normal temperature and pressure conditions in the gas phase. Indeed, the free energy of association is close to +1 kcal/mol for monohydrated complexes and +3 kcal/mol for the dihydrated ones. Thus, the sum of the two first terms on the right-hand side of eq 4 amounts to about 7 kcal/mol for monohydrated complexes and 15 kcal/mol for the dihydrated systems.

The third term in this equation can be computed using the continuum model. Results are presented in Table 5. Only the most stable monohydrated complex and the dihydrated complex are included. The values in Table 5 do not include nonelectrostatic effects because the parametrized model is not expected to work accurately for multimolecular systems especially when water molecules are considered. One can do a rough estimate of an upper limit by adding the nonelectrostatic solvation energies of each molecule species considered separately. For the water molecule, the nonelectrostatic solvation term is computed as proposed in a previous work,[34] i.e., by making the difference between the experimental solvation energy (−6.3

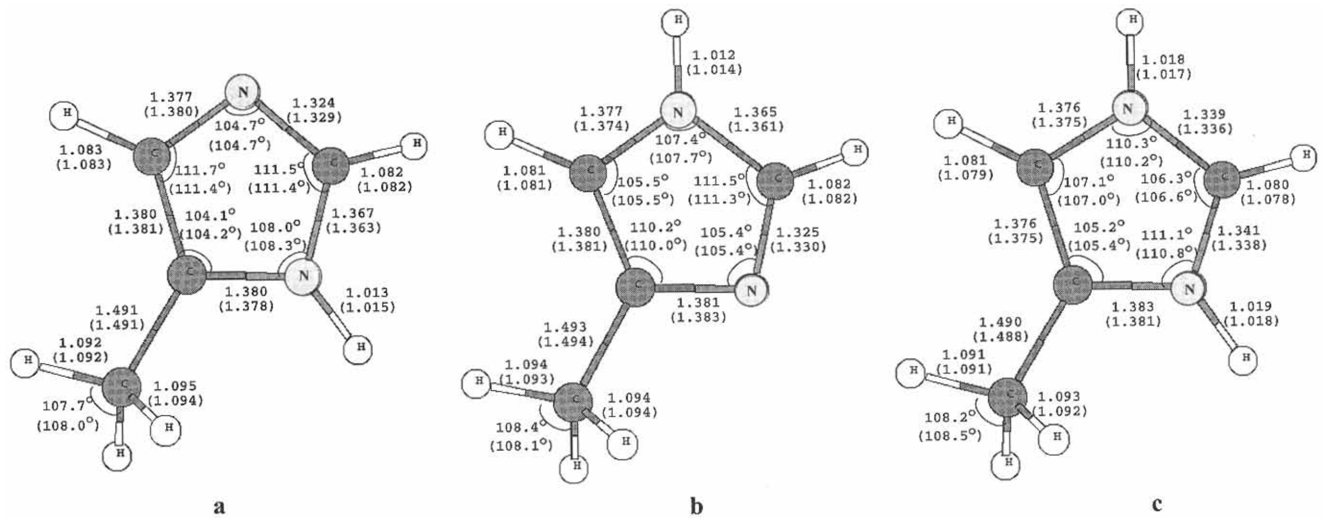


Figure 3. Structures and geometry parameters of (a) 5-methylimidazole, (b) 4-methylimidazole, and (c) protonated methylimidazole in gas phase and in aqueous solution (in parentheses). Distances in angstroms and angles in degrees. Values are obtained at the MP2/6-31G* level.

TABLE 3: Solvation Free Energies Calculated Using the Continuum Model at the MP2/6-31G* Level (A) and MP2/6-311++G//MP2/6-31G* Level (B) (Values in kcal/mol)**

	A			B			
	ΔG_{sol}^{EP}	ΔG_{sol}^{NE}	ΔG_{sol}	ΔG_{sol}^{EP}	ΔG_{sol}^{NE}	ΔG_{sol}	exptl.[4a]
5-MeIm	−6.64	1.04	−5.60	−7.54	1.04	−6.50	−10.27
4-MeIm	−6.39	0.79	−5.60	−7.18	0.79	−6.39	

TABLE 4: Cavity Volumes (in Å^3) and Dipole Moments (in D) Computed at the MP2/6-31G* Level

		dipole moment	
	volume	gas phase	solution
5-MeIm	138.13	4.117	5.214
4-MeIm	138.16	3.646	4.690

TABLE 5: Electrostatic Solvation Free Energies (kcal/mol) Calculated for the Hydrated Complexes at the MP2/6-31G* Level (A) and MP2/6-311++G//MP2/6-31G* Level (B)**

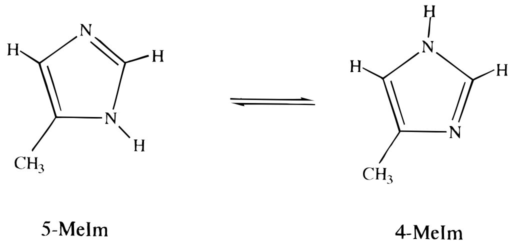
			$\Delta G_{sol}^{EP}(M(H_2O)_n)$	
system	structure	nH$_2$O	A	B
5-MeIm	**I**	1	−6.24	−6.68
	III	2	−9.49	−10.02
4-MeIm	**IV**	1	−7.01	−7.56
	VI	2	−11.75	−11.97

kcal/mol) and the electrostatic energy (−5.2 kcal/mol) computed here at the MP2/6-311++G**//MP2/6-31G* level. This gives −1.1 kcal/mol for such a term. The nonelectrostatic solvation energy for 4- and 5- MeIm is close to 1 kcal/mol (see Table 3). Therefore, the nonelectrostatic terms for the complexes are expected to be negligible for the monohydrated systems or slightly negative for the dihydrated systems. In summary, the electrostatic solvation energies of the complexes presented in Table 5 are presumably not far from the total solvation energies for these complexes.

Using eq 5 and the values presented above, we find that the mixed discrete-continuum model predicts small negative or even positive values of $\Delta G_{sol}(M)$. In other words, it predicts low-stability or unstable methylimidazole−water complexes in liquid water.

There are several crude approximations in this evaluation of the solvation term with the mixed model that would require a deeper analysis. For instance, it must be stressed that the complex geometry has not been optimized in the dielectric medium. This may introduce a nonnegligible error in the solvation energy since the cooperative effect is important, as shown above. The geometry relaxation in solution would also influence the complexation entropic contribution, which has a large magnitude in vacuum. Moreover, as pointed out by one of the referees, large errors may be introduced in the computation of the entropic term in hydrogen-bonded clusters due to very anharmonic low-frequency modes. Nevertheless, a fine analysis of the free energy contributions in the hydration process of imidazole derivatives deserves further investigation using more elaborated techniques such as molecular dynamics. Simulations using combined quantum mechanics and molecular mechanics potentials are being considered.[49]

SCHEME 1



3.3. Tautomerization Equilibrium in Aqueous Solution. Let us now consider the taumerization equilibrium in Scheme 1. Two environments have been considered to analyze the role of hydration. In the first model, we assume the species to be placed in a polarizable continuum. Second, the molecules are assumed to interact with one or two water molecules through hydrogen bonds.

In Table 6, we summarize the results. Values in the gas phase were reported in ref 29. The calculations at the highest level used here predicted the isomer 4-MeIm to be more stable than 5-MeIm by about 0.7 kcal/mol at normal conditions, in good agreement with other calculations.[30]

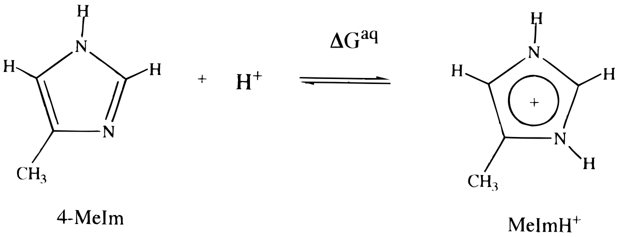
All the models predict a small modification of the tautomerization free energy in solution (±0.1 kcal/mol). At the highest computational level, the electrostatic continuum model leads to a free-energy decrease (in absolute value), i.e., to an increase of the equilibrium constant. At the same level, the formation of hydrogen bonds leads to the opposite effect. As said in the Introduction section, experimentally, the data is quite dependent on the measurement conditions leading to either positive or negative free energy differences and therefore to tautomeric

TABLE 6: Tautomerization Energetics for 4-Methylimidazole (Energies in kcal/mol, Values for 4-MeIm Minus 5-MeIm) Computed at the MP2/6-31G* Level (A) and MP2/6-311++G//MP2/6-31G* Level (B)**

medium	level	ΔE	ΔH	ΔG	K_T
gas-phase	A	−0.21	−0.30	−0.38	1.92
	B	−0.51	−0.61	−0.69	3.20
aqueous solution					
cavity model	A	0.04	−0.05	−0.39	2.07
	B	−0.15	−0.25	−0.58	2.99
discrete model					
monohydrated (**IV−I**)	A	−0.27	−0.37	−0.36	1.84
	B	−0.65	−0.75	−0.74	3.46
dihydrated (**VI−III**)	A	−0.27	−0.38	−0.31	1.69
	B	−0.68	−0.79	−0.72	3.39
experiment[a]				−0.24	1.5

[a] Measured in water at pH = 11 using NMR techniques (see ref 7c). ΔG is calculated from K_T by $\Delta G = -RT \ln(K_T)$.

SCHEME 2



constants that may be smaller or larger than unity. The value $K_T = 1.5$ has been reported by Wasylishen et al.[7c] in water at pH = 11, which corresponds to $\Delta G = -0.24$ kcal/mol. This free energy change is not far from our calculation, although it is slightly smaller.

3.4. Computation of pK_a's. The accurate computation of pK_a's in aqueous solution is, as said before, a difficult task. Consider the process in aqueous solution shown in Scheme 2. One defines the pK_a as

$$\mathrm{p}K_a = -\frac{1}{2.303RT}\Delta G^{aq} \qquad (6)$$

The difficulties in obtaining absolute values of pK_a's from theoretical calculations may be understood if one writes

$$\Delta G^{aq} = \Delta G^{gas} - \Delta G_{sol}(\mathrm{H}^+) + \delta\Delta G_{sol,AH^+/A} \qquad (7)$$

where

$$\delta\Delta G_{sol,AH^+/A} = \Delta G_{sol}(\mathrm{AH}^+) - \Delta G_{sol}(\mathrm{A}) \qquad (8)$$

Here, ΔG^{gas} is the gas-phase free energy for the process and ΔG_{sol} are solvation energies. A first problem is due to the presence of the proton solvation energy in eq 7. Experimentally,[18,50] this quantity varies in a wide range, −254 to −261 kcal/mol, and is difficult to evaluate theoretically with high accuracy. Moreover, to get reliable pK_a values it is also necessary to compute accurate protonation free energies in gas-phase and free energies of solvation, both for neutral and charged species. In practice, only the estimation of relative pK_a's may be presently done with confidence. A detailed discussion on the calculation of pK_a's using dielectric continuum models may be found in ref 18.

In Table 7, we present the results obtained for the pK_a in water using the continuum model to represent the liquid. Because of the large sensitivity of the protonation energy to the basis set,[29] we have used here the highest computational level only.

The calculated solvation energy for the protonated species deserves some comments. The total value (−52.2 and −55.6 kcal/mol for protonated 4-methylimidazole and imidazole respectively) is the sum of an electrostatic contribution (−51.6 and −54.8 kcal/mol, respectively) and a nonelectrostatic term (−0.6 and −0.9 kcal/mol, respectively). Contrary to the case of the neutral molecules, the latter contributions are negative, which can be explained by the presence of a supplementary hydrogen atom bonded to N, for which the surface parameter γ_i (see eq 3) is large and negative (−40 cal/mol·Å^2).[34] The total solvation energy for MeImH$^+$ is in very good agreement with estimations made using eqs 6−8 and experimental data for pK_a (7.52),[51] ΔG^{gas} (226.2 kcal/mol)[17] and ΔG_{sol}(A) (−10.27 kcal/mol),[4a] which lead to solvation energies for MeImH$^+$ in the range −48.3 to −55.3 kcal/mol, depending on the value used for $\Delta G_{sol}(\mathrm{H}^+)$.

The agreement between the computed protonation free energies and pK_a values for methyimidazole and imidazole and the experimental quantities is reasonable. Indeed, the experimental data is within the limits of the values predicted using different experimental values for the proton solvation energy. The agreement between theory and experiment is very good for the relative variation of the protonation energy and pK_a on going from imidazole to 4-methylimidazole.

4. Conclusions

The results presented above show that the simple continuum model is able to describe the main solvent effects on the tautomeric equilibrium and allows a good estimation of relative pK_a's for imidazole derivatives. Conversely, the computation of solvation free energies and absolute pK_a is not very accurate. The use of a discrete-continuum model does not allow improvement of the results, since the entropy decrease accompanying the methylimidazole−water complex formation is quite large. We have shown, however, that the association energy of 4- and 5-methylimidazole with water molecule is also substantial.

Cooperativity appears to be an important aspect of the interaction of these bifunctional molecules with their surroundings. For instance, the strength of a hydrogen bond between 4- or 5-methylimidazole and water molecule is enhanced through either the formation of a second hydrogen bond or the electrostatic solvation of the complex. These results show that nonadditive effects are important when considering the role of the first solvation shell of hydrated imidazole derivatives. A realistic description of solvated molecules requires the use of molecular dynamics simulations, but our results here suggest

TABLE 7: Absolute and Relative pK_a Values in Water Obtained Using the Continuum Model and MP2/6-311++G//MP2/6-31G* Calculations**

	calculations					experiment[a]	
	ΔG^{gas}	ΔGsol(A)	ΔGsol(AH$^+$)	ΔG^{aq}	pK_a	ΔG^{aq}	pK_a
4-MeIm	−220.07	−6.39	−52.23	−4.92/−11.92	3.60/8.73	−10.26	7.52
imidazole	−216.47	−6.94	−55.70	−4.23/−11.23	3.10/8.23	−9.48	6.95
δ	−3.6	0.55	3.47	−0.69	0.50	−0.78	0.57

[a] pK_a is taken from ref 51, ΔG^{aq} is calculated from experimental pK_a according to eq 6.

that caution should be paid to the nonadditivity terms arising from hydrogen-bond formation by using for instance flexible models of the solute and solvent molecules and explicit electronic polarizabilities.

Acknowledgment. The authors thank Drs. Cramer, Luque, and Orozco for fruitful discussions. G.-S.L. thanks the Ministère des Affaires Etrangères of the French government for a doctoral fellowship.

References and Notes

(1) (a) Christoffersen, R. E.; Maggiora, C. M. *Chem. Phys. Lett.* **1969**, *3*, 419. (b) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. (c) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640. (d) Warshel, A.; Werss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218. (e) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (f) Field, M. J.; Bash, P. A.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092. (g) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. (h) Gao, J.; Xia, X. *Science* **1992**, *258*, 631. (i) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537. (j) Stanton, R. V.; Hartsough, D. V.; Merz, K. M., Jr. *J. Phys. Chem.* **1993**, *97*, 11868. (k) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161. (l) Thery, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. G. *J. Comput. Chem.* **1994**, *15*, 269. (m) Stanton, R. V.; Little, L. R.; Merz, K. M. Jr. *J. Phys. Chem.* **1995**, *99*, 17344. (n) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F. *Chem. Phys. Lett.* **1995**, *241*, 450. (o) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F. *J. Mol. Model.* **1995**, *1*, 196. (p) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170. (q) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F.; Rivail, J.-L. *J. Comput. Chem.* **1996**, *17*, 19. (r) Monard, G.; Loos, M.; Thery, V.; Baka, K.; Rivail, J.-L. *Int. J. Quantum Chem.* **1996**, *58*, 153. (s) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580. (t) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B. Eds.; VCH Publishers: New York, 1996; Vol 7. (u) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F. *J. Chem. Phys.* **1997**, *106*, 3633. (v) Strnad, M.; Martins-Costa, M. T. C.; Millot, C.; Tuñón, I.; Ruiz-López, M. F.; Rivail, J.-L. *J. Chem. Phys.* **1997**, *106*, 3643.

(2) (a) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887. (b) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927. (c) Tobin, J. B.; Whitt, S. A.; Cassidy, C. S.; Frey, P. A. *Biochemistry* **1995**, *34*, 6919. (d) Warshel, A.; Papazyan, A.; Kollman, P. A. *Science* **1995**, *269*, 102. (e) Cleland, W. W.; Kreevoy, M. M. *Science* **1995**, *269*, 104. (f) Frey, P. A. *Science* **1995**, *269*, 104.

(3) Grimmett, M. R. In *Advances in heterocyclic chemistry*; Katritzky, A. R., Boulton, A. J., Eds.; Academic Press: New York, 1970; Vol. 12 p 104; *Ibid*, 1980; Vol. 27, p 255.

(4) (a) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849. (b) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664. (c) Meot-Ner, M.; Liebman, J. F.; Del Bene, J. E. *J. Org. Chem.* **1986**, *51*, 1105. (d) Gorman, G. S.; Amster, I. *J. Org. Mass Spectrom.* **1993**, *28*, 1602.

(5) (a) Weinstein, H.; Chou, D.; Johnson, C. L.; Kang, S.; Green, J. P. *Mol. Pharmacol.* **1976**, *12*, 738. (b) Weinstein, H.; Muzarek, A. P.; Osman, R.; Topiol, S. *Mol. Pharmacol.* **1986**, *29*, 28.

(6) Pyman, F. L. *J. Chem. Soc.* **1922**, *121*, 2616.

(7) (a) Charton, M. *J. Org. Chem.* **1965**, *30*, 3346. (b) Ganellin, C. R. *Jerusalem Symp. Quantum Chem. Biochem.* **1974**, 43. (c) Wasylishen, R. E.; Tomlinson, G. *Can. J. Biochem.* **1977**, *55*, 579. (d) Schuster, I. I.; Roberts, J. D. *J. Org. Chem.* **1979**, *44*, 3864.

(8) Worth, G. A.; King, P. M.; Richards, W. G. *Biochim. Biophys. Acta* **1989**, *993*, 134.

(9) Shan, S.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97.

(10) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1993**, *231*, 59.

(11) (a) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1994**, *237*, 602. (b) Takahashi, T.; Nakamura, H.; Wada, A. *Biopolymers* **1992**, *32*, 897.

(12) Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, *111*, 4190.

(13) Richards, W. G.; Wallis, J.; Ganellin, C. R. *Eur. J. Med. Chem.* **1979**, *14*, 9.

(14) Mó, O.; de Paz, J. L. G.; Yáñez, M. *J. Phys. Chem.* **1986**, *90*, 5597.

(15) Meot-Ner, M. *J. Am. Chem. Soc.* **1988**, *110*, 3071.

(16) Alagona, G.; Ghio, C.; Nagy, P.; Simon, K.; Náray-Szabo, G. *J. Comput. Chem.* **1990**, *11*, 1038.

(17) Catalán, J.; de Paz, J. L. G.; Yáñez, M.; Claramunt, R. M.; López, L.; Elguero, J.; Anvia, F.; Quian, J. H.; Taagepera, M.; Taft, R. W. *J. Am. Chem. Soc.* **1990**, *112*, 1303.

(18) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.

(19) de Vries, A. H.; van Duijnem, P. Th. *Biophys. Chem.* **1992**, *43*, 139.

(20) Bliznyuk, A. A.; Schaefer, H. F.; Amster, I. J. *J. Am. Chem. Soc.* **1993**, *115*, 5149.

(21) Ostrovskii, V. A.; Erusalimskii, G. B.; Shcherbina, M. B. *Zh. Org. Khim.* **1993**, *29*, 1297.

(22) Worth, G. A.; Richards, W. G. *J. Am. Chem. Soc.*, **1994**, *116*, 239.

(23) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978.

(24) Lee, H. M.; Lee, G. Y. J. *Korean Chem. Soc.* **1994**, *38*, 21.

(25) Mátyus, P.; Fuji, K.; Tanaka, K. *Tetrahedron* **1994**, *50*, 2405.

(26) Meyer, M. *THEOCHEM* **1994**, *304*, 45.

(27) Mulholland, A. J. Ph.D. Thesis, Oxford University, **1995** and personal communication.

(28) Raczynska, E. D. *Pol. J. Chem.* **1996**, *70*, 795.

(29) Li, G.-S.; Ruiz-López, M. F.; Zhang, M.-S.; Maigret, B. *THEOCHEM*, in press.

(30) Luque, F. J.; López-Bes, J. M.; Cemeli, J.; Aroztegui, M.; Orozco, M., submitted.

(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision B.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(32) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986; Chapter 6.

(33) (a) Rivail, J.-L.; Rinaldi, D.; Ruiz-López, M. F. In *Theoretical and Computational Models for Organic Chemistry*; Formosinho, S. J., Csizmadia, I. G., Arnaut, L., Eds.; NATO ASI Series C; Kluwer Academic: Dordrecht, 1991; Vol. 339, p 79−92. (b) Dillet, V.; Rinaldi, D.; Angyán, J. G.; Rivail, J.-L. *Chem. Phys. Lett.* **1993**, *202*, 18.

(34) Tuñón, I.; Ruiz-López, M. F.; Rinaldi, D.; Bertrán, J. *J. Comput. Chem.* **1996**, *17*, 148.

(35) Pappalardo, R. R.; Rinaldi, D. SCRFPAC, QCPE No. 622, Indiana University: Bloomington, IN, 1992.

(36) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *59*, 4557.

(37) Frank, H. S.; Wen, W. Y. *Discuss. Faraday Soc.*, **1957**, *24*, 133.

(38) Kleeberg, H.; Luck, W. A. P. *Z. Phys. Chem. (Leipzig)* **1989**, 613.

(39) Del Bene, J. E.; Pople, J. A. *Chem. Phys. Lett.*, **1969**, *4*, 426.

(40) (a) Peeters, D.; Leroy, G. *THEOCHEM* **1994**, *314*, 39. (b) Mó, O.; Yáñez, M.; Elguero, J. *THEOCHEM* **1994**, *314*, 73. (c) Kelterbaum, N.; Turki, N.; Rahmouni, A.; Kochanski, E. *THEOCHEM* **1994**, *314*, 191. (d) Kelterbaum, N.; Turki, N.; Rahmouni, A.; Kochanski, E. *J. Chem. Phys.* **1994**, *100*, 1589. (e) Karpfen, A.; Yanovitskii O. *THEOCHEM* **1994**, *314*, 211. (f) Contador, J. C.; Aguilar, M. A.; Sánchez, M. L.; Olivares del Valle, F. J. *THEOCHEM* **1994**, *314*, 229. (g) González, L.; Mó, O.; Yáñez, M.; Elguero, J. *THEOCHEM* **1994**, *371*, 1. (h) Suhai, S. *J. Phys. Chem.* **1994**, 101, 9766. (i) Ojamae, L.; Hermansson, K. *J. Phys. Chem.* **1994**, *98*, 4271. (j) Packer, M. J.; Clary, D. C. *J. Phys. Chem.* **1995**, *99*, 14323. (k) Karpfen, A. *J. Phys. Chem.* **1996**, *100*, 13474. (l) Estrin, D. A.; Paglieri, L.; Corongiu, G.; Clementi, E. *J. Phys. Chem.* **1996**, *100*, 8701. (m) Tschumper, G. S.; Yamaguchi, Y.; Schaefer III, H. F. *J. Chem. Phys.* **1997**, *106*, 9627. (n) Novoa, J. J.; Mota, F.; Perez del Valle, C.; Planas, Marc, *J. Phys. Chem*, in press.

(41) Bertrán, J.; Ruiz-López, M. F.; Rinaldi, D.; Rivail, J.-L. *Theor. Chim. Acta* **1992**, *84*, 181.

(42) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Oxford University Press: New York, 1990.

(43) Mayer I. *Chem. Phys. Lett.* **1983**, *97*, 270; *Ibid.* **1985**, *117*, 396. Mayer, I. *Int. J. Quantum Chem.* **1986**, *29*, 73, 477.

(44) Kleeberg, H.; Klein, D.; Luck, W. A. P. *J. Phys. Chem.* **1987**, *91*, 3200.

(45) Huyskens, P. L. *J. Am. Chem. Soc.* **1977**, *99*, 2578.

(46) Mó, O.; Yáñez, M.; Elguero, J. *J. Chem. Phys.* **1992**, *97*, 6628.

(47) Sánchez-Marcos, E.; Pappalardo, R. R.; Rinaldi, D. *J. Phys. Chem.* **1991**, *95*, 8928.

(48) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629. See also ref 25 of that paper.

(49) Li, G.-S.; Martins-Costa, M. T. C.; Millot, C.; Maigret, B.; Ruiz-López, M. F. Work in progress.

(50) Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207.

(51) Hofmann, K. *The chemistry of heterocyclic compounds: Imidazole and its derivatives*; Weissberger, A., Ed.; Interscience Publishers Ltd.: London, 1953; Part I, Vol. 6, p 15.