Biophysical Chemistry, 43 (1992) 139–147
Elsevier Science Publishers B.V., Amsterdam

BIOCHE 01657

Theoretical calculation of tautomer equilibria of 4-substituted imidazoles in the gas phase and in solution

Alex H. de Vries and Piet Th. van Duijnen

Department of Chemistry, State University of Groningen, Nijenborgh 4, 9747 AG Groningen (Netherlands)

(Received 8 November 1991; accepted in revised form 20 January 1992)

Abstract

The tautomer equilibrium of a number of 4-substituted imidazoles in the gas phase and in aqueous solution was calculated by combining quantum chemical *ab initio* calculations on the tautomers in the gas phase with classical electrostatics calculations (a continuum model and a Monte Carlo method with a limited number of discrete solvent molecules) to evaluate the solvation energy. The results were in good agreement with experimental data from ^{15}N-NMR studies. It was found to be important to include counter ions in the calculations for imidazoles with a charged side chain. The methods for evaluation of the solvation energy were compared.

Keywords: Imidazole, substituted; Tautomer equilibrium; Molecular orbital calculations; Solvation energy; Poisson equation; Monte Carlo simulation

1. Introduction

The imidazole ring is a very important functional group in many chemical processes, and has therefore received a great deal of attention over the years [1]. In biochemical processes imidazole rings play an important part in many active centres of enzymes, e.g. those of serine proteases [2] and blue Copper proteins [3]. The state and site of protonation may be crucial to the activity of the enzyme and is subject to the influences of the environment. A neutral imidazole ring has one of its two N-atoms protonated, as in Fig. 1.

A substituent R ≠ H results in preferential protonation of $N_{\delta 1}$ or $N_{\varepsilon 2}$ (this nomenclature of the imidazole N-atoms derives from that of histidine side-chains in proteins), as was shown by ^{15}N-NMR spectroscopy in aqueous solution for a number of substituents [4]. From the NMR shifts the position of the equilibrium may be calculated in terms of the free energy difference between the two tautomeric forms. It is our aim to reproduce the energy differences found for various substituents from equilibrium studies, by combining quantum chemical *ab initio* calculations on the imidazole species (to obtain energy differ-

Fig. 1. $N_{\varepsilon 2}$–$N_{\delta 1}$ protonation equilibrium.

Correspondence to: A.H. de Vries, Department of Chemistry, State University of Groningen, Nijenborgh 4, 9747 AG Groningen (Netherlands). Fax: (int) - 31 - 50634800, E-mail: ALEXXX@RUGR86.RUG.NL.

Table 1

4-Substituted imidazoles under investigation, experimental data

Species	R	K_{aq} [a]	$-\Delta\Delta G^{\theta}_{aq}$ (kJ/mol) [b]
4-Ethylimidazole	CH_2CH_3	–	≈ -2 [c]
Z-Urocanic acid	*Z*-CHCHCOO$^-$	5.2 ± 1.6	4.0 ± 0.8
E-Urocanic acid	*E*-CHCHCOO$^-$	0.37 ± 0.04	-2.5 ± 0.3

[a] $K_{aq} = [H\text{–}N_{\delta 1}]/[H\text{–}N_{\varepsilon 2}]$, from reference [4].
[b] $\Delta\Delta G = -RT \ln(K_{aq})$.
[c] Based on 4-methyl imidazole [8].

ences *in vacuo*) with a discrete classical and a continuum model method (to obtain solvation energy terms). This study is part of our efforts to develop and apply approximate, (semi-)classical methods to evaluate the interactions between a quantum mechanical system and its surroundings [5–7].

Table 1 shows the imidazole species investigated. The 4-ethylimidazole was chosen as a reference point to study the substituent effect.

Experimentally, the $N_{\varepsilon 2}$ tautomer is favoured, whatever the nature of the substituent (R) is [4], with the exception of the *Z*-isomer of urocanic acid, supposedly because of intramolecular hydrogen bonding, involving $H\text{–}N_{\delta 1}$ (Fig. 2).

2. Theory

It is our aim to connect quantum mechanical calculations with thermodynamic observables via the framework of statistical mechanics. In this section the connection is made explicit and the assumptions and approximations made in the process are briefly outlined.

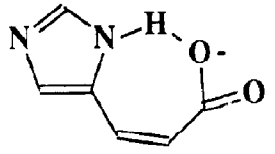

Fig. 2. Intramolecular hydrogen bond in the *Z*-urocanic acid anion.

Consider the protonation equilibrium to be studied in aqueous solution shown in Fig. 1. The equilibrium constant K_{aq} associated with this equilibrium may be calculated from the Gibbs free energy difference $\Delta\Delta G^{\theta}_{aq}$ between the species in solution by eq. (1).

$$K_{aq} = \exp\left(-\Delta\Delta G^{\theta}_{aq}/RT\right) \tag{1}$$

with $\Delta\Delta G^{\theta}_{aq} = \Delta G^{\theta}_{f}(H\text{–}N_{\delta 1(aq)}) - \Delta G^{\theta}_{f}(H\text{–}N_{\varepsilon 2(aq)})$, R the universal gas constant, and T the absolute temperature.

The state function G may be calculated from the energy levels E_i of the system under study via the partition function $Z = \sum_i \exp(-E_i/kT)$ [9]. The energy levels of a system can be found by solving the Schrödinger equation for stationary states [10] $\hat{H}\Psi_i = E_i\Psi_i$, in which $\hat{H}$ is the time-independent Hamiltonian and Ψ_i the wave function associated with energy level E_i. Solving this Schrödinger equation is a formidable task, if at all possible, and therefore approximations are necessary.

The system is partitioned into solute and solvent, and only the solute is treated quantum mechanically. Quantum chemical *ab-initio* calculations on the solute, *in vacuo*, in the Born–Oppenheimer (BO) approximation [11,12] yield the equilibrium conformation and its energy at 0 K, D_0, the energy at the bottom of the BO potential energy surface. The true ground state energy is found by adding the zero-point vibrational energy E_0 to D_0. Looking at chemical problems, at room temperature, the higher energy levels may be calculated approximately by treating the different motions of the molecule as *independent*. Then, the contributions of these motions to E_i are additive.

In the approximation of treating a molecule as a rigid rotor for the rotational motions and as a collection of harmonic oscillators for the vibrational motions [13], the zero-point vibrational energy is given by:

$$E_0 = \tfrac{1}{2} \sum_{i=1}^{N_{vib}} h\nu_i \tag{2}$$

in which h is Planck's constant and ν_i is the

frequency of the ith normal mode. The vibrational and rotational partition functions in these approximations are given by [14]:

$$z_{\text{vib}} = \frac{1}{\prod_i [1 - \exp(-h\nu_i/kT)]} \tag{3}$$

$$z_{\text{rot}} = \sqrt{\pi}\left(\frac{8\pi^2 kT}{h^2}\right)^{3/2} (I_x I_y I_z)^{1/2} \tag{4}$$

where I_x is the moment of inertia with respect to the x-axis, etc., and k Boltzmann's constant.

Because of the independence of the modes of motion the expression for the partition function is greatly simplified into a product of the partition functions for the independent modes. Consequently their contribution to the Helmholtz free energy A is additive:

$$Z = \exp(-\varepsilon_0/kT)\, z_{\text{elec}} z_{\text{vib}} z_{\text{rot}} z_{\text{trans}};$$
$$A = \varepsilon_0 - kT(\ln z_{\text{elec}} + \ln z_{\text{vib}} + \ln z_{\text{rot}} + \ln z_{\text{trans}}) \tag{5}$$

in which ε_0 is the reference energy level from which the independent contributions are counted. It is now possible to compute the free energy difference of the tautomers in the gas phase, at a certain temperature T, from quantum mechanical calculations on the separate tautomers, introducing the approximations discussed. Considering only the electronic ground states and the equality of mass for the tautomers, ΔA_{elec} and ΔA_{trans} become zero, so that:

$$\Delta A_{\text{gas;T}} = \Delta D_0 + \Delta E_0 - kT \ln\{z_{\text{vib}}(N_{\delta 1})/z_{\text{vib}}(N_{\varepsilon 2})\} - kT \ln\{z_{\text{rot}}(N_{\delta 1})/z_{\text{rot}}(N_{\varepsilon 2})\} \tag{6}$$

Ignoring pressure–volume work differences, we make the approximation $\Delta G \approx \Delta A$, to arrive at the thermodynamic quantity of interest.

The next step is to take the isomers from the gas phase into solution. As a crude approximation the solvent may be treated as a dielectric surrounding the solute. The solvation free energy may then be approximated by the electrostatic work done on bringing the charge distribution representing the solute into a cavity in the dielectric from infinity. A first approximation to this problem is to reduce the charge distribution of the solute to its overall charge Q and overall dipole moment M, and place these at the centre of a spherical cavity with radius r, the size of which reflects the size of the solute. This model, combining the Born [15] and Onsager [16] formulae, gives the solvation energy ΔW as:

$$\Delta W = -\frac{1}{2}\left\{\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon}\right\}\frac{Q^2}{r} - \frac{(\varepsilon - 1)}{(2\varepsilon + 1)}\frac{M^2}{r^3} \tag{7}$$

in which ε_i is the dielectric constant in the interior of the sphere. Both terms have been derived using the reaction potentials of the interior charge and dipole. They are analytical solutions to the Poisson equations for the point charge and point dipole in a spherical cavity. This simple model may be refined by a more detailed description of the charge distribution of the solute and of the surface separating it from the dielectric. Several such methods have been developed, e.g. [17,18]. These are all numerical approaches to the solution of the Poisson equation, which cannot be solved analytically if the cavity surface is not of a special geometry. The method by Juffer et al., used in this study [19], consists of a solute represented by point charges at the positions of the nuclei, separated from a dielectric medium (with possibly non-zero ionic strength) by a surface that follows the shape of the solute. A pair of coupled equations for the potential and the normal component of the electric field (the latter only with non-zero ionic strength) at the dielectric surface is solved by a straightforward application of boundary element techniques.

More elaborate models explicitly consider discrete solvent molecules, and the configurations they may adopt. The configuration space is sampled using Molecular Dynamics (MD) or Monte Carlo (MC) techniques [20], with simple interaction potentials between the atoms that make up the system to evaluate the energy of the system. These approaches to the solute–solvent interaction give more detailed information about the nature of the interactions, but they are very time-consuming computationally.

A model that combines continuum and Monte Carlo approaches has been developed in our laboratory by Rullmann and Van Duijnen [6]. In this model the solute is again represented by point charges and is surrounded by a number of discrete solvent molecules, also represented by point charges. The number of solvent molecules is limited to two to three solvation layers and this entire collection is restricted to a sphere, which is the boundary between the discrete part and a dielectric representing the bulk solvent. The distribution of the discrete solvent molecules is sampled with Monte Carlo techniques to obtain an average over solvent configurations. Short contacts between solvent molecules and the dielectric are prevented to circumvent over-polarization effects. The solvation energy may then be calculated as the difference between the energies of the solvent ensemble and the solute plus solvent ensemble. Pressure–volume terms are accounted for if the radius of the sphere is being enlarged as the solute is entered. This model has been called the RFE (Reaction Field plus Exclusion) model.

Note that for both models, the solvation energy is, at present, evaluated for *one solute configuration only*. For this reason we write ΔW for the solvation energies calculated by the Born–Onsager (eq. 7) and continuum (Juffer et al.) approach, and $\Delta H_{\mathrm{solv;T}}$ for those by the Monte Carlo method. Finally, Δ, the approximation to $\Delta\Delta G^{\theta}_{\mathrm{aq}}$ for the tautomer equilibrium is calculated as:

$$\Delta = \Delta A_{\mathrm{gas;T}} + \Delta\Delta W \quad \text{(continuum models)} \quad (8a)$$

and

$$\Delta = \Delta A_{\mathrm{gas;T}} + \Delta\Delta H_{\mathrm{solv;T}} \quad \text{(Monte Carlo model)} \quad (8b)$$

3. Methods

Quantum mechanical energies and geometries of the 4-substituted imidazoles *in vacuo* were obtained by using the routines of the HONDO [21] programs package for MO-LCAO (Molecular Orbital–Linear Combination of Atomic Orbitals) single determinant energies and geometry optimizations at the SCF (Self Consistent Field) level. All calculations were at the STO4-31G basis set level for second row atoms and STO3-21G for hydrogen. The results at this level concur with the results at higher levels for the cases checked. The minimal STO-3G level does not give consistent results. The charge distribution of the final vacuum structures was represented by Dipole Preserving (DP) [22] charges at the positions of the nuclei.

Solvation energies were calculated with the MOLPOT [23] programs package for the continuum approach by solving the Poisson equation as de-

Table 2

Results of gas phase calculations [a]

Species	ΔD_0 [b]	$E_{0;\mathrm{vib}}$	$\Delta E_{0;\mathrm{tot}}$	A_{vib} [c]	A_{rot} [c]	ΔA_{gas} [c]
Eth-$N_{\delta 1}$	0	364.5	0	−6.72	−24.1	0
Eth-$N_{\epsilon 2}$	−2.4	364.2	−2.7	−6.84	−24.1	−2.8
Z-$N_{\delta 1}$	0	307.5	0	−9.42	−26.9	0
Z-$N_{\epsilon 2}$	67.2	307.0	66.7	−10.7	−27.0	65.3
Z-acid	79.5	304.9	76.9	−10.9	−27.0	75.3
E-$N_{\delta 1}$	0	306.0	0	−12.2	−27.3	0
E-$N_{\epsilon 2}$	33.9	305.4	33.3	−11.6	−27.3	34
E-acid	−5.2	304.1	−7.1	−10.4	−27.4	−5.4

[a] All quantities in kJ/mol, and relative to $N_{\delta 1}$ species.

[b] relative to $N_{\delta 1}$ species. D_0 for these species in a.u.: *Eth*: −302.41945; *Z*: −488.03457; *E*: −488.01469.

[c] The Helmholtz free energies have been evaluated at T = 298.15 K, using eqs. (3) and (4) for the partition functions.

scribed earlier. (The package can also solve the linearized form of the Poisson–Boltzmann equation to take the ionic strength of the solution into account.) A triangulated surface surrounding the molecule was determined from the dotted Connolly surface [24].

The program MCDIP [25] was used for the solvation energy calculation via the Monte Carlo (RFE) approach. In this case, the discrete part consisted of 72 SPC (Simple Point Charge) [26] water molecules, which roughly corresponds to three solvation layers. For each species the solvation energy for the equilibrium geometry only was calculated.

Calculations were performed on CYBER 760/962 and CRAY-YMP (HONDO and MCDIP) and CONVEX C220 (MOLPOT) mainframes.

4. Results

The gas phase results are shown in Table 2. For the ethylimidazole, the proton was put in two positions: near $N_{\delta 1}$ and near $N_{\varepsilon 2}$ denoted as *Eth*-$N_{\delta 1}$ and *Eth*-$N_{\varepsilon 2}$. For the urocanic acid anions, the proton was put in three positions: near $N_{\delta 1}$, near $N_{\varepsilon 2}$ and near the acid group denoted as *Z*/*E*-$N_{\delta 1}$/$N_{\varepsilon 2}$/acid, respectively.

The energies reflect the optimized geometries of the species. The ethyl group of 4-ethylimidazole was found to be rotated out of the imidazole plane by about 72°. The optimized geometries of the urocanic acid anions were all planar, as expected.

The solvation energies calculated are presented in Table 3, together with the total energy differences between the tautomers in aqueous solution. These differences have been denoted Δ, as explained above (eq. 8).

In the case of the urocanic acids, the NMR-experiment is done at high ionic strength: 0.2–0.4 *M* in Na^+ and urocanic acid anion. This means that approximately one Na^+-ion is present to only two hundred water molecules, so there will generally be such an ion in the vicinity of the imidazole. In order to investigate the influence of the Na^+-ion, gas phase and solvation energies were calculated for the acid anions, with Na^+-ion near the acid group. The position of the Na^+-ion was determined by gas phase geometry optimization, starting at approximately equal distances to both

Table 3

Solvation energies (kJ/mol); total energy differences (kJ/mol) in aqueous solution

Species	ΔA_{gas}	$\mu(D)$ [a]	ΔW(BO) [b]	ΔW(solv, MP) [c]	ΔH(solv, MC) [d]	Δ(tot, MP)	Δ(tot, MC)
Eth-$N_{\delta 1}$	0	4.33	−4.9	−47.4	−80	0	0
Eth-$N_{\varepsilon 2}$	−2.8	3.70	−3.6	−50.8	−81	−6.2	−4
Z-$N_{\delta 1}$	0	6.24	−142	−302	−314	0	0
Z-$N_{\varepsilon 2}$	65	11.4	−134	−362	−387	5	−8
Z-acid	75	2.73	−128	−293	−301	84	88
E-$N_{\delta 1}$	0	11.7	−137	−372	−387	0	0
E-$N_{\varepsilon 2}$	34	17.7	−137	−401	−412	5	9
E-acid	−5.4	4.65	−123	−295	−301	72	81

[a] Vacuum dipole moment with respect to centre of mass.

[b] Born–Onsager formula (eq. 7); cavity radius is calculated as the largest distance of an atom to the centre of mass, augmented by that atom's associated Van der Waals radius.
For the charged species, the charge was placed in the centre of charge, so as to let the dipole moment vanish, with the centre of mass as the centre of the cavity.

[c] MOLPOT (MP) calculation. The Van der Waals radii for C, H, O, N and Na was used to generate Connolly's surface with probe radius of 1.4 Å were 1.7, 1.2, 1.52, 1.55 and 1.54 Å, respectively. The triangulated surface followed Connolly's surface closely.

[d] Monte Carlo (MC) method; RFE model.

carboxylic oxygens. The Na^+-ion remained at such a position, with a Na–O distance of approximately 2.2 Å, with the exception of the Z-$N_{\delta 1}$ form of urocanic acid, for which optimization resulted in an asymmetric geometry.

The sensitivity of the solvation energy term to the position of the Na^+-ion was investigated by placing the counter ion near the unprotonated N of the imidazole ring. Although some such arrangements led to appreciably higher solvation energies, the total energies always exceeded those of the corresponding geometries with the Na^+-ion near the acid group, due to the higher vacuum energies. The results for the lowest total energy geometries are shown in Table 4.

5. Discussion

5.1 Gas phase calculations

5.1.1 Substituent effect

The calculated total energies in the gas phase of the species investigated (Table 2) indicate that introducing a substituent at the 4-position of an imidazole ring results in a preferred protonation site. The aliphatic ethyl group favours $N_{\varepsilon 2}$-protonation, in accord with experimental observation for the majority of 4-substituted imidazoles [4], while the charged Z- and E-propenoic acid anion groups both favour $N_{\delta 1}$-protonation. In the case of Z-urocanic acid, the reason for this preferred $N_{\delta 1}$-protonation is quite obvious: intramolecular hydrogen bonding between H–$N_{\delta 1}$ and a carboxyl oxygen, as shown in Fig. 1, was found in the optimized gas phase geometry. Bond-order analysis [27] showed that, apart from H-bond formation, pseudo-aromaticity [28] may contribute to the stability of the Z-$N_{\delta 1}$ species.

In the gas phase the proton in the E-urocanic acid anion prefers to attach itself to the acid group. We are not aware of experimental data concerning the acidity of the groups in the gas phase, but an explanation for the calculated relative acidities is easily found in terms of delocalization of the negative charge. If the proton resides on the acid group, more resonance structures can be written down, compared to the proton residing on one of the imidazole N-atoms.

The large difference in energy between the $N_{\delta 1}$ and $N_{\varepsilon 2}$ species of the E-urocanic acid anion (34 kJ/mol), compared to the 4-ethylmidazole (−2 kJ/mol) remains unexpected, however, because a pseudo-aromatic bicyclic system such as in the H–$N_{\delta 1}$ Z-urocanic acid anion is not possible in either of the E-urocanic acid anions. It must find its explanation in the nature of the side chain. Adding a Na^+-ion to the structures in the neighbourhood of the acid group resulted in a considerable decrease of the energy differences, as can be seen from Table 4. The difference is now of the order found in 4-ethylimidazole for the E-urocanic acid. A similar effect has been observed for the histamine tautomerism [29], on

Table 4

Gas phase and solvation energies (kJ/mol) for the urocanic acid anions plus a Na^+-ion near the acid group; total energy differences (kJ/mol) in aqueous solution

Species	ΔD_0 [a]	$\mu(D)$ [b]	ΔW(solv, MP) [c]	ΔH(solv, MC) [d]	Δ(tot, MP)	Δ(tot, MC)	Δ(exp)
Z-$N_{\delta 1}$ + Na	0	8.84	−200	−155	0	0	0
Z-$N_{\varepsilon 2}$ + Na	28.7	8.88	−223	−174	6	10	4.0
E-$N_{\delta 1}$ + Na	0	7.65	−224	−168	0	0	0
E-$N_{\varepsilon 2}$ + Na	2.4	4.34	−228	−168	−2	2	−2.5

[a] Relative to $N_{\delta 1}$ species. D_0 for these species in a.u.: Z: −649.92131; E: −649.90646.

[b] Vacuum dipole moment with respect to centre of mass.

[c] MOLPOT calculation.

[d] Monte Carlo (MC) method; RFE model.

going from histamine itself to its cation (with an extra proton on the amino group).

Apparently the charge has a pronounced effect on the imidazole ring, even though it is somewhat remote. The *Z*-urocanic acid with counter ion still displays a strong preference for the hydrogen-bonded conformation, as may be expected.

5.1.2 Vibrational and rotational corrections

In general, addition of zero-point vibrational energy, rotational and vibrational Helmholtz free energy only slightly modifies the results at the usual 0 K level (Table 2). Only the more tightly bound proton in the $N_{\delta 1}$-species of the *Z*-urocanic acid causes a strong decrease in Helmholtz free energy, because the vibrational frequencies are higher and contribute less to the vibrational partition function. A similar effect is seen in the $N_{\varepsilon 2}$-species of the *Z*-urocanic acid, where there is some restriction of motion as a consequence of the folding back of the side chain as well.

Rotational contributions to the Helmholtz free energy cancel already in the gas phase. In solution, rotational freedom of the species can be neglected altogether because of the friction of the solvent, and we have not given it any more consideration here.

5.2 Solvation energy calculations

5.2.1 Born–Onsager approximation

The solvation energies calculated by different methods may be conveniently compared by inspection of Table 3. The first observation is the inaccuracy of the Born–Onsager approach. For neutral solutes, the solvation energies are too small by one order of magnitude. For the charged solutes, the solvation energy calculated is still much too small. This observation is not at all surprising, since the Born–Onsager approximation is very sensitive to the cavity radius (with a dependence of r^{-3} for neutral solutes as in eq. 7). The approach may be expected to be reasonable for small, spherical species, such as ions or very small molecules. However, the molecules under study here are far from spherical and the cavity is too large for the molecule.

The solvation energy calculated by the Born–Onsager formula is of course quadratic in the overall dipole moment, but the more accurate (MOLPOT and MC) calculations show that the charge *distribution* may give rise to deviations from what would be expected if only the solute's total dipole moment is considered. Thus a species with the smaller overall dipole moment may still have the larger solvent stabilization.

5.2.2 Comparison between Monte Carlo and continuum approaches

The discrete solvent model and the continuum model give very similar results, as concerns both the magnitude of the solvation energy and the differences between the species. This can be seen from the last four columns of Table 3, where the solvation energies and the final differences Δ are tabulated. Note that the acid-protonated species are not at all important anymore in solution, as expected. This is due to the relatively small solvation energy for this species, mainly because of the relatively small dipole moment.

The differences as concerns the differential solvation term between the two approaches seldomly exceed 5 kJ/mol, except in the case of *Z*-urocanic acid, where the MC approach seems to overestimate the solvation energy of the $H–N_{\varepsilon 2}$ form, so that the equilibrium shifts rather drastically to the wrong side compared to experiment (Δ should be +4, Table 1). This may be due to the limited size of our solvent sphere, so that the outer water molecules do not yet show bulk behaviour.

As was mentioned in the previous section, the equilibrium under study is measured at high ionic strength, whereas the calculations are valid only at infinite dilution. Including a Na^+ counter ion in the quantum mechanical calculations reduces the differences in energy between the species in the gas phase and shifts the equilibria toward their experimental values, as can be seen from Table 4 (note that zero-point and free energy corrections have been neglected). For the *E*-urocanic acid the situation becomes very similar to that of 4-ethylimidazole, which re-emphasizes the major influence of the side chain charge. The solvation energies of the tautomers are very close, and the equilibrium is subtle, as was observed before for neutral imidazole solutes by Worth et

al., using a free energy perturbation approach [29].

The continuum approach gives results that are in accord with experiment, whereas the position of the equilibrium is reversed with the Monte Carlo approach.

The continuum approach has the advantage that it is much faster computationally (about a hundred times), but it is rather sensitive to a number of important parameters, which are more or less hidden. Among these are the radii of the atoms that determine the size of the solute cavity, via the construction of Connolly's surface, and the accuracy of the description of the surface. In this study all calculations were performed with the same set of parameters; the radii used were the Van der Waals radii [30] and the surface was described up to the level where improvement of the surface's description did not alter the solvation energies significantly.

The Monte Carlo method is much slower than the continuum method computationally, as was already mentioned. Generally, at least 10^6 configurations were necessary to give statistically good results, but for the uncharged species twice this number was required, because convergence was much slower.

Recently, a discrete free energy perturbation (FEP) method [31] and continuum calculations [18] were compared [32] for a number of solutes. Good agreement was found between the two approaches, as in this study. It thus seems, that if the interest is only in solvation (free) energies, a continuum model is accurate, while discrete models should be employed to obtain detailed information about solvent orientation. The usefulness and accuracy of the continuum approach has been demonstrated in this case, where experimental data were available. This supports the application of this approach in cases where experimental data are not available, although the hidden parameters discussed need to kept in mind and checked critically.

6. Conclusions

The protonation equilibrium of 4-substituted imidazoles was calculated to be sensitive to substituent and solvation effects, confirming experimental data. In the case of *E*-urocanic acid, further calculations are necessary if the Monte Carlo approach is employed, including counter ions in the calculations in a more realistic way, by allowing them to move around in the solvent. Solvation energies (or at least differences in solvation energies) may be reliably calculated via the continuum approach, as implemented by Juffer et al., saving considerable computer time with respect to Monte Carlo sampling methods.

The gas phase calculations indicate large differences in behaviour of the urocanic acids compared to the solvated species. Experimental work on these molecules may be both interesting and stimulating to the study of the biologically important [33] urocanic acids.

Acknowledgement

The authors are grateful to A.H. Juffer for helpful discussions and to A.A. van Dijk and G.T. Robillard for instigating this particular application. This work was supported by a grant to A.H. de Vries by the Dutch Organisation for Scientific Research (NWO), and a grant for computing time on a CRAY-YMP by the (Dutch) National Computer Facilities (NCF).

References

1 Extensive reviews by M.R. Grimmett can be found in: Advances in heterocyclic chemistry, eds. A.R. Katritzky and A.J. Boulton, vols. 12, p. 104 and 27, p. 255 (Academic Press, New York, NY, 1970 and 1980).
2 W.W. Bachovchin and J.D. Roberts, J. Am. Chem. Soc. 100 (1978) 8041.
3 E.T. Adman, Structure and function of small blue copper proteins, in: Topics in molecular and structural biology: Metalloproteins, ed. P.M. Harrison (MacMillan, New York, NY, 1986) Vol. 6 (Part 1), pp. 1–42.
4 J.D. Roberts, C.Yu, C. Flanagan and T.R. Birdseye, J. Am. Chem. Soc. 104 (1982) 3945.
5 B.T. Thole and P.Th. van Duijnen, Theor. Chim. Acta 55 (1980) 307.
6 J.A.C. Rullmann and P.Th. van Duijnen, Mol. Phys. 61 (1987) 293.
7 P.Th. van Duijnen, A.H. Juffer and J.P. Dijkman, J. Mol. Struct. (Theochem) 1992, in press.
8 I.I. Schuster and J.D. Roberts, J. Org. Chem. 44 (1979) 3864.

9 D.A. McQuarrie, Statistical mechanics (Harper and Row, New York, NY, 1976).
10 R. McWeeny and B.T. Sutcliffe, Methods of molecular quantum mechanics (Academic Press, London, 1976).
11 M. Born and R. Oppenheimer, Ann. Phys. 84 (1927) 457.
12 M. Born and K. Huang, Dynamical theory of crystal lattices (Clarendon Press, Oxford, 1966).
13 E.B. Wilson, Jr., J.C. Decius and P.C. Cross, Molecular vibrations (McGraw-Hill, New York, NY, 1955).
14 H. Eyring, J. Walter and G.E. Kimball, Quantum Chemistry (Wiley, Inc., New York, NY, 1944).
15 M. Born, Z. Phys. 1 (1920) 45.
16 L. Onsager, J. Am. Chem. Soc. 58 (1936) 1486.
17 R.J. Zauhar and R.S. Morgan, J. Comput. Chem. 9 (1988) 171.
18 M.K. Gilson, K.A. Sharp and B.H. Honig, J. Comput. Chem. 9 (1988) 327.
19 A.H. Juffer, E.F.F. Botta, B.A.M. van Keulen, A. van der Ploeg and H.J.C. Berendsen, J. Comput. Phys. 97 (1991) 144.
20 M.P. Allen and D.J. Tildesley, Computer simulation of liquids (Clarendon Press, Oxford, 1987).
21 M. Dupuis, J. Rys and H.F. King, J. Chem. Phys. 65 (1976) 111.
22 B.T. Thole and P.Th. van Duijnen, Theor. Chim. Acta (Berlin) 63 (1983) 209.
23 MOLPOT programs package written and kindly supplied by: A.H. Juffer, Dept. of Biophysical Chemistry, State University of Groningen, Nijenborgh 4, 9747 AG Groningen (Netherlands).
24 M.L. Connolly, Science 221 (1983) 709.
25 MCDIF programs package was written by J.A.C. Rullmann and developed at our laboratory.
26 H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren and J. Hermans, in: Intermolecular forces, ed. B. Pullman (Reidel, Dordrecht, 1981) p. 331.
27 H.O. Villar and M. Dupuis, Chem. Phys. Lett. 142 (1987) 59.
28 Aromaticity, Pseudo-aromaticity, anti-aromaticity, eds. E.D. Bergmann and B. Pullmann, The Jerusalem symposia on quantum chemistry and biochemistry III, Proc., Jerusalem, 1970.
29 G.A. Worth, P.M. King and W.G. Richards, Biochim. Biophys. Acta 1036 (1990) 158.
30 A. Bondi, J. Phys. Chem. 68 (1964) 441.
31 W.L. Jorgensen, and J. Gao, J. Am. Chem. Soc. 110 (1988) 4212
32 A. Jean-Charles, A. Nicholls, K. Sharp, B. Honig, A. Tempczyk, T.F. Hendrickson and W.C. Still, J. Amer. Chem. Soc. 113 (1991) 1454.
33 M. Norval, T.J. Simpson and A. Ross, Photochem. Photobiol. 50 (1989) 267 (Review).