

Protonation Equilibria of *Hoechst 33258* in Aqueous Solution

by Manfred Ladinig^a), Werner Leupin^{*b}), Markus Meuwly^{*a}), Michal Respondek^a), Jakob Wirz^{*a}),
and Vincent Zoete^a)^c)

^a) Departement Chemie der Universität Basel, Klingelbergstrasse 80, CH-4056 Basel
(e-mail: M.Meuwly@unibas.ch, J.Wirz@unibas.ch)

^b) Gymnasium Liestal, Abteilung Chemie, Friedensstrasse 20, CH-4410 Liestal
(e-mail: Leupin.Werner@gymliestal.ch)

^c) Swiss Institute of Bioinformatics, Molecular Modeling Group., Ch. des Boveresses 155,
CH-1066 Epalinges s/Lausanne

The bis-benzimidazole derivative *Hoechst 33258* (2'-(4-hydroxyphenyl)-5-(4-methylpiperazin-1-yl)-2,5'-bi-1*H*-benzimidazole) binds to the minor groove of DNA duplexes and is widely used as fluorescent cytological stain for DNA. The neutral compound, **1**, is amphiphilic with four basic and three acidic sites. We have determined all seven acidity constants by spectrophotometric titration to define the pH-dependent distribution of species, from the fully protonated tetracation $\mathbf{1}^{4+}$ to the fully deprotonated trianion $\mathbf{1}^{3-}$, in aqueous solution. The structures of the intermediate protonation states were assigned with the aid of density-functional calculations. Electrostatic interaction free energies were calculated to adjust the acidity constants of the molecular subunits of **1** to their environment in the species $\mathbf{1}^{4+}$ to $\mathbf{1}^{3-}$. The experimental and theoretical pK_a values agree well, but they differ substantially from previous estimates given in the literature.

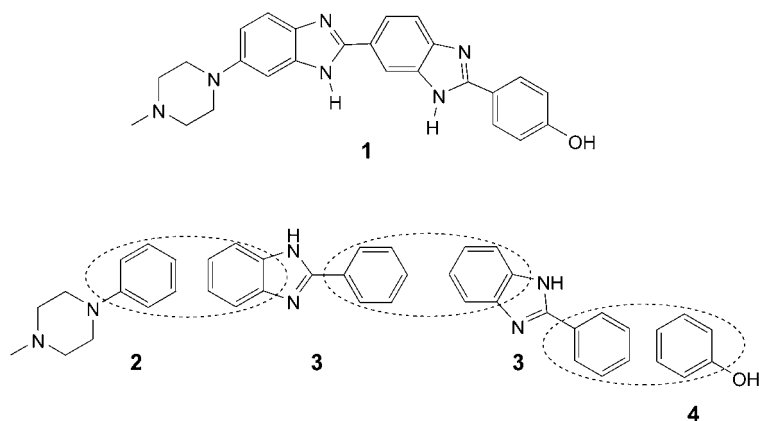
Introduction. – *Hoechst 33258* (2'-(4-Hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-1*H*-benzimidazole) is an anthelmintic compound, *i.e.*, it is active against infections by parasitic worms. Bis-benzimidazoles represent one of the most extensively investigated classes of compounds binding to the minor groove of DNA, and *Hoechst 33258* is a well-known DNA-binding ligand [1][2] with high affinity for DNA sequences containing solely AT base pairs. It is often used as a chromosomal stain [3][4]. Numerous studies have been reported on its interactions with DNA, the effects of this binding on transcription and translation [5], and on the inhibition of radiation-induced strand breakage [6] employing an array of analytical methods, *e.g.*, optical spectroscopy [2–9], footprinting studies [10–13], kinetics studies of the DNA-binding process [8][9], X-ray-diffraction studies on single crystals [14–20], and ¹H-NMR spectroscopy [21–29]. These studies have shown that *Hoechst 33258* binds to the minor groove of DNA, covering about four AT-base pairs and forming intermolecular H-bonds between thymidine-O-atoms and the imino H-atoms of both benzimidazole moieties. In purely GC-containing DNA sequences, the DNA-binding affinity of *Hoechst 33258* decreases at least 100-fold and changes the mode of binding to an intercalating one [30]. The change in binding mode in these GC sequences might be caused by the 2-amino group of guanine that protrudes into the minor groove and, thus, destabilizes minor-groove binding by bis-benzimidazoles. *Hoechst 33258* (**1**) has been reported to inhibit the action of both topoisomerase I [31][32] and topoisomerase II [33] in cell-free systems.

Remarkably, the protonation states of *Hoechst 33258* in aqueous solution are still poorly defined. It is not known whether the protonation state of *Hoechst 33258* changes

upon binding to DNA, or how the association constants depend on pH. The determination of these equilibria is not a trivial task. The fully protonated tetracation, $\mathbf{1}^{4+}$, contains seven acidic sites which give rise to $2^7 = 128$ different protonation states.

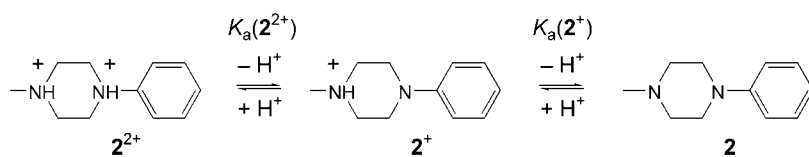
Here, we report a combined experimental and theoretical study to assess the protonation equilibria of $\mathbf{1}$ in aqueous solution. The seven acidity constants of $\mathbf{1}^{4+}$ as well as those of the model compounds $\mathbf{2}$ – $\mathbf{4}$ were determined by spectrophotometric titration. Compounds $\mathbf{2}$ – $\mathbf{4}$, each representing a sub-moiety of $\mathbf{1}$ (Scheme 1), were used as a benchmark to identify the structures of the intermediate protonation states $\mathbf{1}^{3+}$ to $\mathbf{1}^{2-}$ with the aid of a theoretical model. Density-functional theory (DFT) provided atomic charges for $\mathbf{2}$ – $\mathbf{4}$ and their ions, from which electrostatic interaction free energies were calculated for incorporation of the subunits into the 128 ‘supermolecules’ $\mathbf{1}^{4+}$ to $\mathbf{1}^{3-}$. The acidity constants of $\mathbf{1}^{4+}$ so obtained agree well with those determined experimentally and serve to identify the site of deprotonation in each step.

Scheme 1. Fusion of the Submoieties $\mathbf{2}$ – $\mathbf{4}$ to the ‘Supermolecule’ $\mathbf{1}$



Results. – *Protonation Equilibria of the Model Compounds $\mathbf{2}$ – $\mathbf{4}$.* – Protonation of 1-methyl-4-phenylpiperazine ($\mathbf{2}$) occurs first on the methyl-substituted N-atom (Scheme 2), as expected from the relative basicity of aniline and methyl-substituted amines. The second protonation is achieved only in strongly acidic solutions. It is inhibited by the ensuing Coulombic repulsion between the two positive charges in the diprotonated piperazine ring of $\mathbf{2}^{2+}$.

Scheme 2. Protonation Equilibria of 1-Methyl-4-phenylpiperazine ($\mathbf{2}$)



Protonation of the methyl-substituted N-atom of $\mathbf{2}$ hardly affects the absorption spectrum of the aniline chromophore (Fig. 1). The small changes observed upon

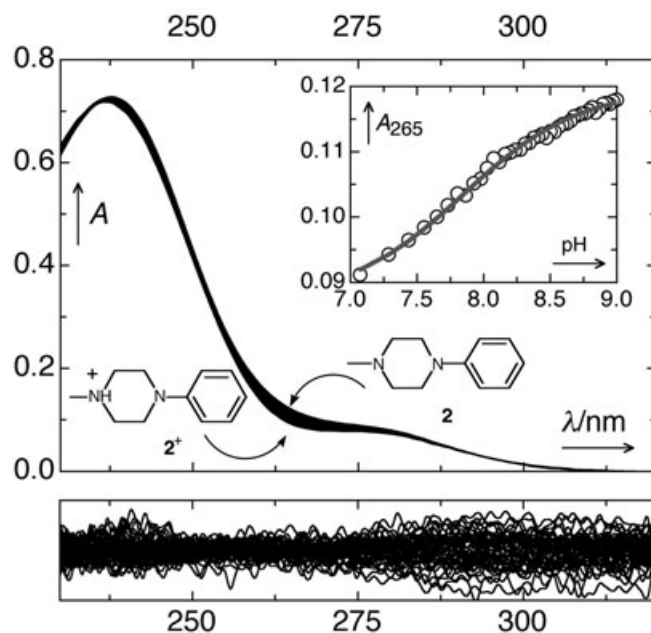


Fig. 1. Electronic spectra of **2** in borate buffer (pH 7.0–9.0). Inset: Absorbances at 265 nm (○) and the globally fitted titration function. Bottom traces: Residuals between the measured absorbances and those reconstructed with two spectral components (scale: $\pm 2 \times 10^{-3}$ absorbance units).

titration of **2** in borate buffer (buffer concentration was adjusted to ionic strength 0.1M), pH 7.0–9.0, were nevertheless sufficient to determine the acidity constant of 2^+ by global fitting of a monoacid titration model. The spectra were reproduced within experimental error (standard deviation of 0.0015 absorbance units) using the two dominant spectral components obtained by factor analysis. Four independent runs gave an average $pK_a(2^+)$ value of 7.82 ± 0.07 ($I=0.1M$, 25°). The error represents the standard deviation of the mean.

Potentiometric titrations were performed with the same setup by addition of a solution containing 0.1M KOH and 0.1M KCl to a 4×10^{-5} M solution of **2** in 0.1M KCl that was acidified with 5×10^{-5} M of HCl. Three independent runs gave $pK_a(2^+) = 7.82 \pm 0.01$ ($I=0.1M$, 25°). The agreement between the two methods is excellent. The precision of the spectrophotometric titration is lower, because the absorbance changes associated with protonation of the methyl-substituted N-atom are small.

The second protonation of 2^+ to 2^{2+} occurs only in concentrated acid solutions. A spectrophotometric titration was carried out by addition of 1N KOH to a solution of 2^{2+} in 1N HCl containing 1M KCl (glass electrode readings pH 0.03–1.5). The absorption by the aniline chromophore largely vanishes upon protonation. The resulting estimate is $pK_a(2^{2+}) \approx 0.7 \pm 0.1$ ($I=1M$).

The first absorption band of 2-phenyl-1H-benzimidazole (**3**) shifts slightly to the blue upon protonation (Scheme 3), whereas deprotonation is accompanied by a substantial red shift. Protonation of **3** was measured by addition of 0.1M HCl

containing 0.1M KCl to a solution of **3** in 0.1M AcONa, pH 1.9–7.3 (Fig. 2, left). Nonlinear fitting of four independent runs gave $\text{p}K_{\text{a}}(\mathbf{3}^+) = 5.24 \pm 0.03$ ($I = 0.1\text{M}$, 25°). Protonation of $\mathbf{3}^-$ in 0.1M KOH was observed by neutralization with a solution containing 0.1M HCl and 0.1M KCl, pH 10.4–12.8 (Fig. 2, right). Four independent runs provided $\text{p}K_{\text{a}}(\mathbf{3}) = 11.78 \pm 0.03$ ($I = 0.1\text{M}$, 25°).

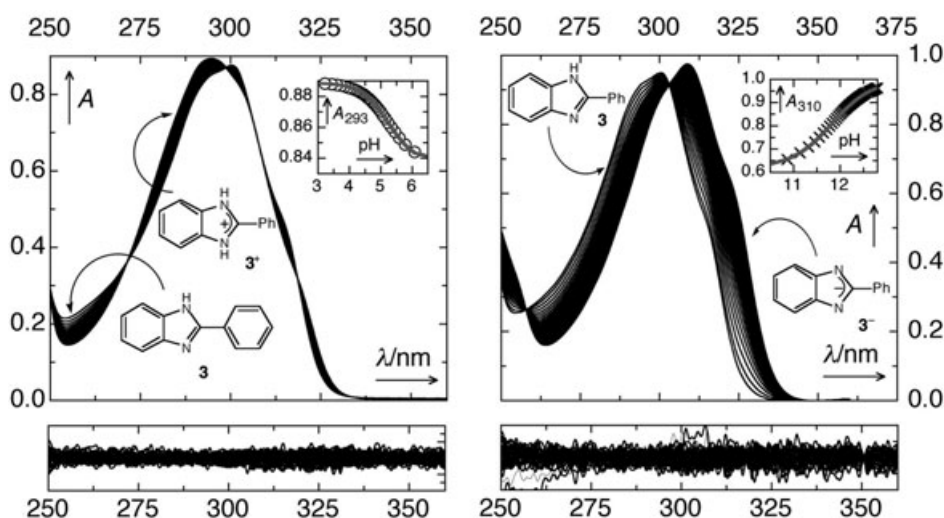
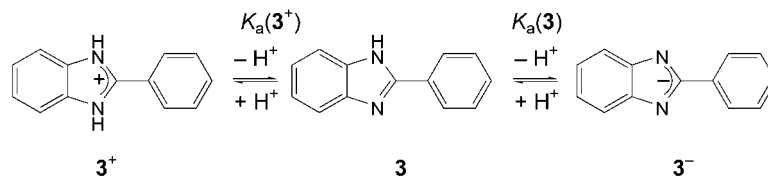
Scheme 3. Protonation Equilibria of **3**

Fig. 2. Electronic spectra of **3** at pH 1.9–7.3 (left) and pH 10.4–12.8 (right). Insets: Absorbances at 293 nm (left ○) and at 310.6 nm (right ×) with the globally fitted titration functions. Bottom traces: Residuals within $\pm 2 \times 10^{-3}$ absorbance units.

Walba and Isensee [34] have reported thermodynamic acidity constants of **3** at 25° , $\text{p}K_{\text{a}}(\mathbf{3}^+) = 5.23 \pm 0.03$ and $\text{p}K_{\text{a}}(\mathbf{3}) = 11.91 \pm 0.04$. The first value is almost identical with our concentration quotient $\text{p}K_{\text{a,c}}(\mathbf{3}^+) = 5.24 \pm 0.03$ ($I = 0.1\text{M}$, 25°). Indeed, deprotonation of $\mathbf{3}^+$ being a charge-shift reaction, the concentration quotient $K_{\text{a,c}}$ should be rather insensitive to ionic strength. From the reported [34] dependence of $\text{p}K_{\text{a,c}}(\mathbf{3})$ on ionic strength, one obtains $\text{p}K_{\text{a,c}}(\mathbf{3}, I = 0.1\text{M}) = 11.74 \pm 0.04$, in good agreement with the value determined here.

For the ionization quotient of phenol, we adopt the value $\text{p}K_{\text{a}}(\mathbf{4}) = 9.84 \pm 0.02$ ($I = 0.1$, 25°) that was determined previously [35] by the same method as is used here.

Protonation Equilibria of Hoechst 33258 (1). – Spectrophotometric titration of a heptaprotic acid, where the mass relationships between eight charge states are governed by seven partly overlapping equilibria, is a demanding task. Also, preparation

of a multi-buffer solution that enables the titration to be carried out over the whole pH scale is impracticable. The experiment was thus divided into sections, each covering a pH range of 3–4 units, and one to three protonation equilibria.

In the first set of experiments, the equilibria between $\mathbf{1}^{3-}$, $\mathbf{1}^{2-}$, and $\mathbf{1}^-$ were determined by neutralization of 1M KOH with a solution containing 1M HCl and 1M KCl, pH 11.9–13.7. Factor analysis indicated the presence of three species in the system. To determine the dissociation quotients at ionic strength $I=0.1\text{M}$, the same titration was performed with tenfold diluted acid and base. The highest pH value in this experiment was 12.8, so that only the midpoint for the deprotonation of $\mathbf{1}^{2-}$ was reached (Fig. 3). Nevertheless, the dissociation quotients obtained from the two sets of measurements were consistent within the error limits. The best estimates for the dissociation quotients of $\mathbf{1}^-$ and $\mathbf{1}^{2-}$ are $\text{p}K_a(\mathbf{1}^-) = 11.69 \pm 0.06$ and $\text{p}K_a(\mathbf{1}^{2-}) = 12.91 \pm 0.04$ ($I = 0.1\text{M}$, 25°).

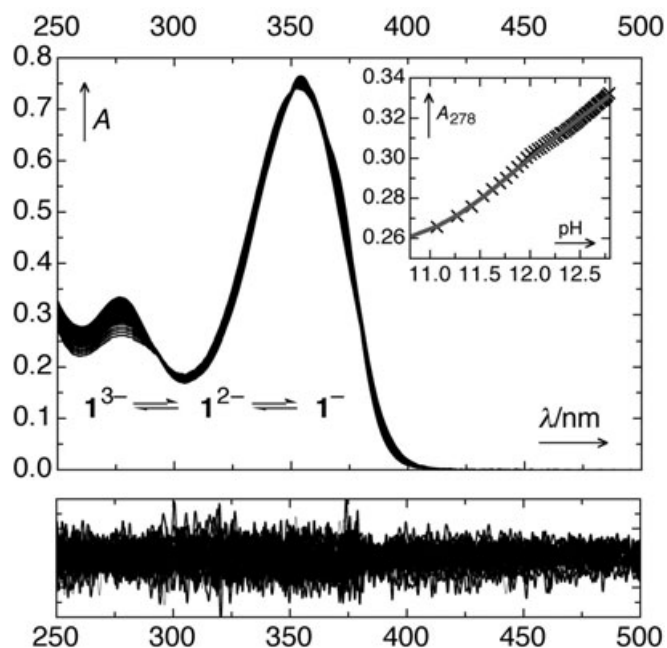


Fig. 3. Electronic spectra of $\mathbf{1}$ in 0.1M KOH titrated with 0.1M HCl containing 0.1M KCl, pH 10.8–12.8. Inset: Absorbances at 278 nm (\times) with the globally fitted titration function. Bottom traces: Residuals within $\pm 2 \times 10^{-3}$ absorbance units.

The protonation equilibria between $\mathbf{1}^+$, $\mathbf{1}$, and $\mathbf{1}^-$ were covered using *Tris* buffer, pH 6.8–9.7. The changes in the absorption spectrum (Fig. 4) are mostly attributed to protonation at the phenolate site of $\mathbf{1}^-$. Protonation of $\mathbf{1}$, which occurs on the methyl-substituted N-atom of the piperazine ring, is not expected to give a large change in the electronic spectrum (cf. Fig. 1). Factor analysis of the data matrices nevertheless gave three significant spectral components, indicating formation of $\mathbf{1}^+$ in the system. However, the third eigenvalue was four orders of magnitude smaller than the second one, not much above the noise level, so that the resulting acidity quotient of $\mathbf{1}^+$,

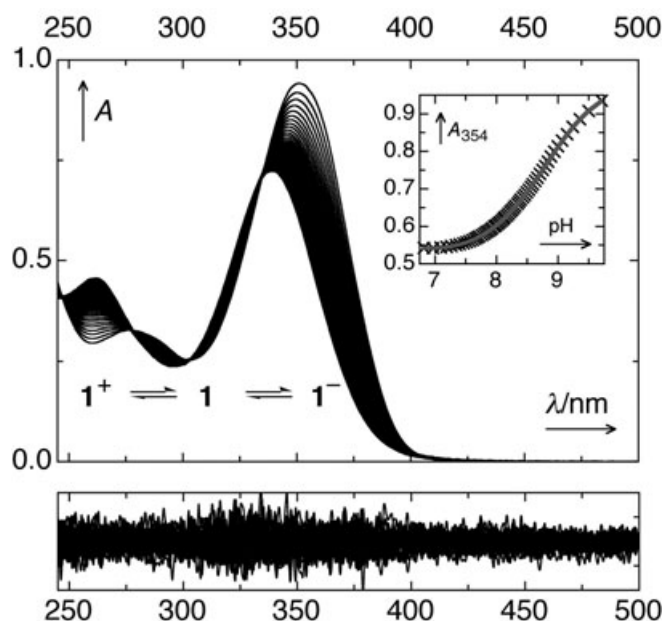


Fig. 4. Electronic spectra of **1** in Tris buffer, pH 6.8–9.7. Inset: Absorbances at 354 nm (×) and the globally fitted titration function. Bottom trace: Residuals within $\pm 2 \times 10^{-3}$ absorbance units.

although quite reproducible, must be viewed with some caution: $pK_a(\mathbf{1}^+) = 7.94 \pm 0.05$ and $pK_a(\mathbf{1}) = 8.90 \pm 0.03$ ($I = 0.1\text{M}$, 25°).

Two sets of measurements were performed covering the pH ranges 4.6–6.5 (AcONa/HCl) and 2.4–4.6 (HCOONa/HCl) to investigate the protonation equilibria between $\mathbf{1}^{3+}$, $\mathbf{1}^{2+}$, and $\mathbf{1}^+$. In the first set of measurements, the original spectra were reproduced with a standard error of 0.0015 absorbance units using four significant spectral components (Fig. 5, left). The fourth factor is attributed to a small contribution by species **1** in the spectra at highest pH. The contribution by the fourth factor is very small, three components were sufficient to reproduce the spectra recorded at pH < 6.2. These are attributed to species $\mathbf{1}^+$, $\mathbf{1}^{2+}$, and $\mathbf{1}^{3+}$. In the fitting procedure, the pK_a of $\mathbf{1}^+$ was fixed to the value 7.94 that had been determined in Tris buffer, that of $\mathbf{1}^{3+}$ to 3.40 from the measurements in formate buffer (*vide infra*).

Species $\mathbf{1}^{3+}$ had to be taken into account as it influenced the spectra, but the value was fixed to 3.40 in the fitting. Three spectral components were required for the spectra recorded with acetate buffer. The third factor was again attributed to $\mathbf{1}^{3+}$. The resulting best estimates are $pK_a(\mathbf{1}^{2+}) = 5.68 \pm 0.03$ and $pK_a(\mathbf{1}^+) = 7.94 \pm 0.05$.

In the second set of measurements covering the pH range of 2.4–4.6 (formate buffer), the original spectra were adequately reproduced with three spectral components (Fig. 5, right) with a standard error of 0.0015 absorbance units. The three factors are associated with the cations $\mathbf{1}^{3+}$, $\mathbf{1}^{2+}$, and $\mathbf{1}^+$. Here, the $pK_a(\mathbf{1}^{3+})$ was best defined. Least-squares fitting of the data with a diprotic acid model was done by fixing the value of $pK_a(\mathbf{1}^{2+})$ to 5.68 obtained above, which gave $pK_a(\mathbf{1}^{3+}) = 3.49 \pm 0.03$ ($I = 0.1\text{M}$, 25°).

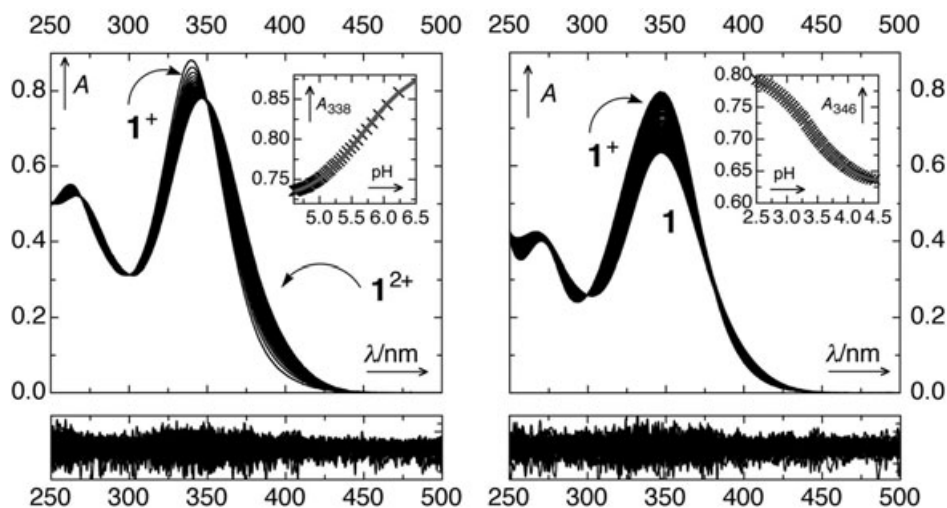


Fig. 5. Electronic spectra of **1** in acetate buffer, pH 4.6–6.5 (left), and formate buffer, pH 2.4–4.6 (right). Insets: Absorbances at 338.6 nm (left) and 346.4 nm (right) with the globally fitted titration function. Bottom traces: Residuals within $\pm 2 \times 10^{-3}$ absorbance units.

Finally, protonation of $\mathbf{1}^{3+}$ on the aromatic N-atom of the piperazine ring is accomplished only in strongly acidic solutions. Only small absorbance changes were observed upon increasing the concentration of HClO_4 up to 4M, namely a slight hypsochromic shift of the longest absorption band and reduction of the tailing around 400 nm, which are attributed to medium effects. A substantial change was seen in solutions of 6M HClO_4 . Therefore, the half-protonation occurs at around 6M HClO_4 , which corresponds to $H_0 = -2$ on the excess acidity scale [36], *i.e.*, $\text{p}K_a(\mathbf{3}^{4+}) \approx -2.0$.

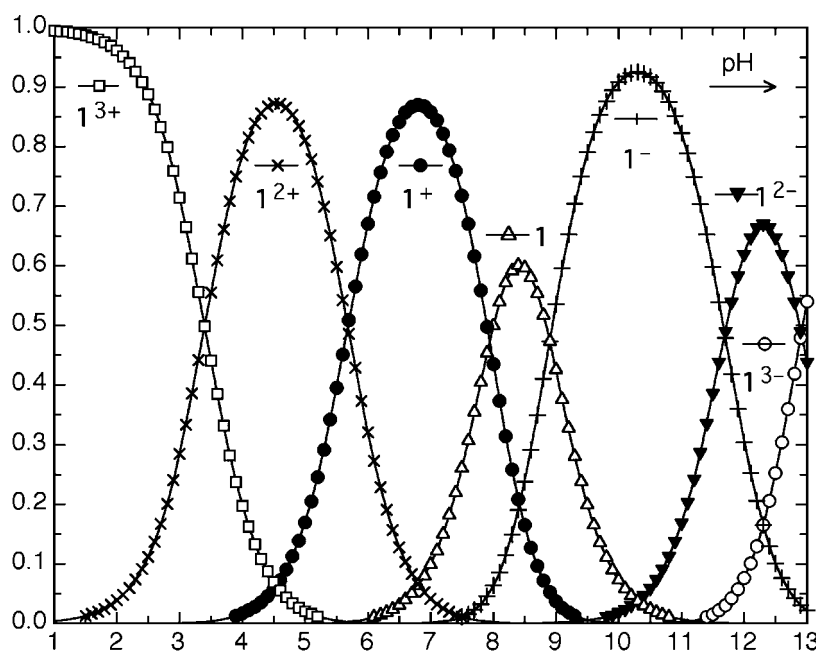
Discussion. – The best estimates for the acid dissociation quotients at ionic strength $I = 0.1\text{M}$ obtained in this work are collected in the *Table*. All except the crude estimate for $\text{p}K_a(\mathbf{1}^{4+})$ were obtained as an average of at least three independent runs.

In *Fig. 6*, the resulting distribution of **1** to its different protonation states is shown as a function of pH. The species concentrations are defined by *Eqn. 1*, where c_{tot} is the total concentration of the dye, arbitrarily set to 1M, c_L is the concentration of the fully deprotonated ligand $\mathbf{1}^{3-}$, K_1 is the first dissociation quotient of **1**, $K_a(\mathbf{1}^{4+})$, *etc.*

$$\begin{aligned}
 c_{\text{tot}} &= c_L \left[1 + c_{\text{H}^+}/K_n + (c_{\text{H}^+})^2/(K_n K_{n-1}) + \dots + (c_{\text{H}^+})^n/(K_n \dots K_1) \right] \\
 c_L &= c_{\text{tot}} / [\dots] \\
 c_{\text{LH}} &= c_L c_{\text{H}^+} / K_n \\
 c_{\text{LH}_2} &= c_{\text{LH}} c_{\text{H}^+} / K_{n-1} \\
 &\dots \\
 c_{\text{LH}_n} &= c_{\text{LH}_{n-1}} c_{\text{H}^+} / K_1
 \end{aligned} \tag{1}$$

Table. Ionization Quotients ($I=0.1\text{M}$) and Structural Assignments of the Acids

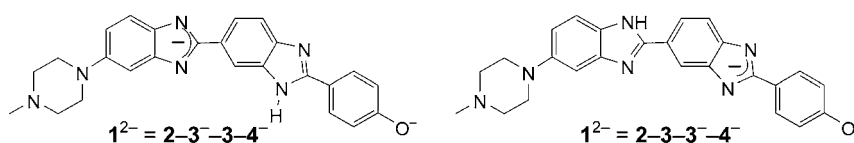
Protonation equilibrium	pK (25°, $I=0.1\text{M}$)	pK calculated	Comment	Structural assignment of acid
$2^{2+} \rightleftharpoons 2^+ + \text{H}^+$	-0.7 ± 0.1	0.53	$I=1\text{M}$	2^{2+}
$2^+ \rightleftharpoons 2 + \text{H}^+$	7.82 ± 0.01	7.79	potentiometric	2^+
$3^+ \rightleftharpoons 3 + \text{H}^+$	5.24 ± 0.03	5.26		3^+
$3 \rightleftharpoons 3^- + \text{H}^+$	11.78 ± 0.03	11.95		3
$4 \rightleftharpoons 4^- + \text{H}^+$	9.84 ± 0.02		from [35]	4
$1^{4+} \rightleftharpoons 1^{3+} + \text{H}^+$	~ -2	-0.25 ± 0.03		$2^{2+}-3^+-3^+-4$
$1^{3+} \rightleftharpoons 1^{2+} + \text{H}^+$	3.49 ± 0.03	4.16 ± 0.04		$2^+-3^+-3^+-4$
$1^{2+} \rightleftharpoons 1^+ + \text{H}^+$	5.68 ± 0.03	5.71 ± 0.04		2^+-3^+-3-4
$1^+ \rightleftharpoons 1 + \text{H}^+$	7.94 ± 0.05	7.86 ± 0.02		$2^+-3-3-4$
$1 \rightleftharpoons 1^- + \text{H}^+$	8.90 ± 0.03	8.77 ± 0.01		$2-3-3-4$
$1^- \rightleftharpoons 1^{2-} + \text{H}^+$	11.69 ± 0.06	12.45 ± 0.02		$2-3-3-4^-$
$1^{2-} \rightleftharpoons 1^{3-} + \text{H}^+$	12.91 ± 0.04	12.77 ± 0.03		$2-3-3^-4^-$

Fig. 6. Species distribution of **1**. The total concentration is arbitrarily set to 1M.

The plot does not include the species 1^{4+} . Its concentration becomes significant only at $\text{pH} < 0$. Note that several protonation states co-exist in appreciable concentrations at all pH values in the range of 1–13 and that the cation 1^+ dominates at $\text{pH} \approx 7$.

In the seven-step protonation sequence of **1** only the structures of the fully protonated species (1^{4+}) and of the fully deprotonated one (1^{3-}) are defined *a priori*. For the intermediate protonation stages, many sites of protonation are possible, and often several structures (or mixtures of these) are reasonable candidates.

Some structural assignments are obvious from the relative acidities of the model compounds **2–4**. Clearly, the benzimidazole N-atoms, $pK_a(\mathbf{3}) = 11.78$, offer the strongest basic sites of $\mathbf{1}^{3-}$. Protonation on the piperazine ring, $pK_a(\mathbf{2}^+) = 7.84$, or on phenolate, $pK_a(\mathbf{4}) = 9.84$, are unlikely to compete. It is more difficult to decide which benzimidazole ring will be protonated first. Electrostatic theory [37] would predict the two negative charges on $\mathbf{1}^{2-}$ to be separated as much as possible, favoring the structure **2–3⁻–3⁻–4⁻**. On the other hand, the donating effect of the amino substituent on the left-hand moiety tends to favor **2–3⁻–3⁻–4⁻**. For the same reason, protonated 5-aminobenzimidazole, $pK_a = 6.1$ [38], is a weaker acid than protonated **3**, $pK_a(\mathbf{3}^+) = 5.24$.

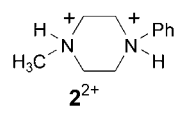
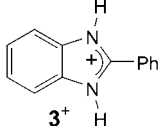


Protonation of the anion $\mathbf{1}^-$ should occur on the O-atom of phenolate, the methyl-substituted piperazine N-atom being less basic. That **1** is present mainly as a neutral molecule (**2–3–3–4**), and not as a zwitterion (**2⁺–3–3–4⁻**), is also indicated by the low solubility of **1** in aqueous solution around pH 8.4. In the neutral molecule, the piperazine ring, $pK_a(\mathbf{2}^+) = 7.84$, is expected to be more basic than the benzimidazoles, $pK_a(\mathbf{3}^+) = 5.24$, suggesting protonation of **1** to occur on the piperazine. Thus, $\mathbf{1}^+$, the most stable species at pH 7, is attributed to **2⁺–3–3–4**. Next, we expect protonation of $\mathbf{1}^+$ to take place sequentially on the two benzimidazole rings.

To obtain more-definite assignments of the protonation states, electrostatic free energies of solvation were calculated for all possible protonation states of **1**. This provides relative free energies of the different protonation states of **1** as a function of pH and, hence, the acidity constants. The approach employed here uses the fact that all protonation states of the ‘supermolecule’ **1** can be rebuilt by combining the four molecular sub-moieties **2–4** (*Scheme 1*) in their various protonation states. *Model compounds* are defined as the isolated sub-moieties (**2**, **3**, and **4**) in solution. Two types of pK_a can be defined for each sub-moiety, the *standard* and the *effective* pK_a . The standard pK_a is the experimental pK_a of the model compound, while the effective ones are those of the same sub-moiety within **1** or one of its ions. The difference between the standard and effective pK_a values is due to electrostatic interactions of the sub-moiety with the rest of the molecule, *i.e.*, the other sub-moieties in their different protonation states. Details of the procedure are given in the *Exper. Part*.

To test the computational approach, we first calculated effective pK_a values for the model compounds **2–4** (*Eqn. 2, Exper. Part*). Thus, each model compound is considered a ‘supermolecule’ composed only of one sub-moiety. Since, in this case, the model compound and the ‘supermolecule’ are identical, the calculated effective pK_a value should be ideally equal to the standard pK_a . An important difference between the two would arise from unrealistic atomic charges in the different protonation states. *Scheme 4* shows the pK_a values for **2** and **3**. The effective (calculated) and standard (experimental) pK_a values of **2** and **3** agree well, indicating that the calculated atomic partial charges for the different protonation states of the sub-moieties are meaningful.

Scheme 4. Calculated and Experimental pK_a Values for **2** and **3**

	Calculated	Experimental
	$pK_{a,1}$ 0.53 $pK_{a,2}$ 7.79	≈ 0.7 7.82
	$pK_{a,1}$ 5.26 $pK_{a,2}$ 11.95	5.24 11.78

The calculated pK_a values for **1** are given in the *Table*. These values were conformationally averaged as follows. First, a 10-ns molecular dynamics (MD) simulation of **1** was performed at 600 K with the CHARMM program [39]. Then 200 conformations were taken from the trajectory and minimized by 100 steps of the Steepest Descent minimization. Finally, the pK_a values were calculated for each of these 200 minimized conformations and averaged. The MD simulation was carried out at elevated temperature to allow for the crossing of energy barriers and, thus, to provide an exhaustive conformational sampling. The pK_a values themselves were calculated at 300 K. The standard deviations given in the *Table* correspond to variations that are due to the conformational sampling only. They do not include other sources of error such as the nuclear charges, or the finite difference solution of the *Poisson* equation.

The agreement between the calculated and experimental values is again satisfactory. The theoretical approach thus provides a reliable assignment of the experimentally determined acidity quotients to specific molecular groups and protonation states. Deprotonation is estimated to occur at lower pH for the right-hand benzimidazole structure ($pK_a = 4.2$ for the right-hand cycle and 5.7 for the other). The presence of the phenol group increases the pK_a value for protonation the adjacent benzimidazole ring. Thus, the theoretical calculations are able to predict the positive and negative shifts of the first pK_a values of the two benzimidazole moieties, as well as the positive shift of the second pK_a value of the left-hand one, relative to the model compounds. However, the predicted positive shift of the second pK_a value of the right-hand benzimidazole group is not observed experimentally. Also, the theoretical model does not allow to distinguish between the two sites within a given benzimidazole ring. The calculated pK_a values of the two conformers are nearly identical in all cases. It should be noted that the pK_a value of the phenol group in **1** was determined from a combination of electrostatic and quantum calculations (see *Exper. Part*).

Previous studies of the protonation states of *Hoechst 33258* in aqueous solution are controversial. *Umetskaya et al.* [40] have obtained four pK_a values for **1** by spectrophotometric titration in phosphate and carbonate buffers, but few details of the measurements are given. The numerical pK_a values and, more important, the postulated sequence of deprotonation differ from the present one as shown in *Scheme 5*. In particular, these authors implied that the zwitterion **2**⁺–**3**–**3**–**4**[–] is the most stable form of neutral **1**.

Scheme 5. Comparison of Postulated Deprotonation Sequences and pK_a Values

From [7]		This work
$2^+-3^+-3^+-4$	3.5	$2^+-3^+-3^+-4$ 3.49
2^+-3-3^+-4	5.5	$2^+-3-3-4$ 5.68
$2^+-3-3-4$	8.5	$2^+-3-3-4$ 7.94
$2^+-3-3-4^-$	9.8	$2-3-3-4$ 8.90
$2-3-3-4^-$		$2-3-3-4^-$

In a recent theoretical study, *Alemán et al.* [41] calculated gas-phase proton affinities for four conjugate bases of $\mathbf{1}^+$, and the relative stability of the four protomers of $\mathbf{1}$ in aqueous solution was predicted on the basis of a dielectric continuum model. Substantial discrepancies with the experimental data of *Umetskaya et al.* [40] were noted, but attributed to inadequacies of the theoretical model. It turns out that the predictions made by *Alemán et al.* are consistent with the present results. The pH-dependence of the fluorescence of *Hoechst 33258* was studied by several groups [42].

In conclusion, the present work defines the distribution of charge states of *Hoechst 33258* as function of pH (*Fig. 6* and *Table*). The cation $\mathbf{1}^+$ is the dominant species at pH values around 7. Its charge is on the methyl-substituted N-atom of the piperazine subunit ($2^+-3-3-4$). Deprotonation of this species, $pK_a(\mathbf{1}^+) = 7.98$, gives the neutral molecule $\mathbf{1} = (2-3-3-4)$, not the zwitterion ($2^+-3-3-4^-$). Protonation of $\mathbf{1}^+$ to the dication $\mathbf{1}^{2+}$, $pK_a(\mathbf{1}^{2+}) = 5.71$, occurs at the left-hand benzimidazole N-atom yielding (2^+-3^+-3-4), which carries two adjacent charges. The calculated pK_a values, which were obtained using a decomposition scheme combined with electrostatic free-energy calculations, agree well with the experimental ones. The theoretical method used here also holds promise to predict the protonation states of $\mathbf{1}$ and related dyes as they are bound to specific sites of DNA. Such information is difficult to obtain by any known experimental method.

Experimental Part

Materials. 2-Phenyl-1H-benzimidazole (**3**) and *Hoechst 33258* were supplied by *Fluka*. 1-Methyl-4-phenylpiperazine (**2**) was synthesized according to the literature [43] (colorless liquid, b.p. $95^\circ/0.075$ mm Hg). $^1\text{H-NMR}$ (400 MHz, CD_2Cl_2): 7.30–7.35 (*m*, 2 H); 6.97–7.01 (*m*, 2 H); 6.91 (*tt*, $J = 7.3, 0.5, 1$ H); 3.23–3.26 (*m*, 4 H), 2.58–2.62 (*m*, 4 H); 2.38 (*s*, 3 H). $^{13}\text{C-NMR}$ (400 MHz, CD_2Cl_2): 151.89; 129.34; 119.60; 116.15; 55.49; 49.30; 46.31. MS: 176 (M^+).

HCOONa, AcONa, and 2-amino-2-(hydroxymethyl)propane-1,3-diol (*Tris*) were obtained from *Fluka*, disodium tetraborate decahydrate from *Merck* and 1M and 0.1M HCl as volumetric solns. from *Riedel-de-Haën*. Volumetric solns. for 0.1M and 1M KOH were obtained from *Fluka* and *Aldrich*, resp. A 70 % soln. of HClO₄ was purchased from *Fluka*.

Data Collection and Analysis. Absorption spectra were recorded with a *Perkin-Elmer Lambda 9* spectrophotometer. The temp. was maintained within $25 \pm 0.3^\circ$ with a thermostated cell holder. A quartz cell of 1-cm path length and 2-cm width with a capacity of 6 ml was used. A baseline was recorded with H₂O prior to each set of experiments. The microprocessor-controlled *Metrohm LL* micro pH glass electrode (*Biotrode*) was thermostated at 25° for at least 24 h before use and calibrated with standard *Metrohm* buffer solns. (pH 4, 7, and 9) [44] before every set of measurements. Portions (50 μ l) of the titrant were added directly to the quartz cell inside the spectrophotometer using *Metrohm 725 Dosimat* equipped with an 806-exchange unit (10-cm³ cylinder). A magnetic stirrer inside the cell ensured rapid mixing of added titrant. Following each addition of titrant, a spectrum was acquired after the time required to keep the drift of the pH electrode within 0.01 pH units per minute. *Ca.* 50 spectra and pH readings were collected during each titration. Measurements with solns. of pH > 10 were performed under a N₂ atmosphere. Except where noted otherwise, titrations were conducted at constant ionic strength $I = 0.1\text{M}$. The resulting acidity constants are, therefore, dissociation quotients $K_{a,c}$ at $I = 0.1\text{M}$. To achieve protonation of the least basic site of **1**, the absorption spectra were measured in 1–6M aq. solns. of HClO₄.

AcONa, HCOONa, or 2-amino-2-(hydroxymethyl)propane-1,3-diol (*Tris*) were dissolved in doubly distilled, CO₂-free H₂O and adjusted to ionic strength $I = 0.1\text{M}$. The absorption by these buffer bases was negligible above 240 nm. Borate buffer with negligible absorption down to 220 nm was used for the titration of 1-methyl-4-phenylpiperazine (**2**), $\lambda_{\text{max}} = 236\text{ nm}$. Initial concentrations of the indicators were $7.6\text{--}10 \times 10^{-5}\text{ M}$ for **2**, $3.7\text{--}4.0 \times 10^{-5}\text{ M}$ for 2-phenyl-1*H*-benzimidazole (**3**), and $1.4\text{--}2.4 \times 10^{-5}\text{ M}$ for **1**. A portion of 50 μ l of 0.1M HCl, containing 0.1M NaCl or 0.1M KCl to keep ionic strength constant, was added before each measurement.

The spectra were corrected for dilution and analyzed with the help of the program SPECFIT [45], which performs a factor analysis and allows to calculate acidity constants as optimized model parameters, together with their standard errors, from multiwavelength spectrophotometric data by nonlinear least-squares fitting to the appropriate model function. Potentiometric data were analyzed with the program TITFIT [46].

Computational Methods. As shown in *Scheme 1*, molecule **1** can be decomposed into four sub-moieties, so that all protonation states of **1** can be rebuilt by combining the four molecular sub-moieties in their various protonation states. *Model compounds* are defined as the isolated sub-moieties, **2–4**, in solution. Four protonation states were considered for **2** (the three shown in *Scheme 2* plus the monocation protonated on the aromatic N-atom), four for **3**, and two for **4**. The total charge and the calculated partial atomic charges characterize the protonation states of the sub-moieties. Atomic charges were calculated with Gaussian 98 [47]. For this, the structure of each sub-moiety in each protonation state was optimized at the B3LYP/6-311++G** level. Atomic charges were then calculated from a natural bond-orbital analysis [48]. The decomposition of a molecular structure into sub-moieties has been successfully used for the calculation of the protonation states in polypeptide chains [49]. In the present case, it is important to note that electronic coupling between the fragments is more likely to occur than in the case of proteins. One example is the p*K*_a value of phenol, which is discussed below.

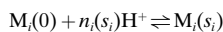
Two types of p*K*_a can be defined for each sub-moiety, the *standard* and the *effective* p*K*_a. The standard p*K*_a is the experimental p*K*_a of the model compound, while the effective one is that of the sub-moiety within **1**. Thus, the *standard* p*K*_a values are 0.7 and 7.82 for sub-moiety **2**, and 5.24 and 11.78 for sub-moiety **3** (see *Table*). A standard p*K*_a value of 8.88 was used for the phenol subgroup in **1** to account for electronic effects due to the coupling with the benzimidazole fragment (*vide infra*). The *effective* p*K*_a values are those of the sub-moieties within **1**. The difference between the standard p*K*_a and the effective p*K*_a for a given sub-moiety are due to the electrostatic interactions of this sub-moiety and the rest of the molecule, *i.e.*, the other sub-moieties in their different protonation states.

Here, we are interested in predicting the *effective* p*K*_a values of the titratable sites in **1** given the standard p*K*_a value of each site, the atomic charges in the different protonation states, a range of geometries and a method to calculate the electrostatic free energy for each possible protonation state (*vide infra*). The approach estimates the difference between the standard and the effective p*K*_a for each titratable sub-moiety. The effective p*K*_a values of interest are calculated as follows: given a molecule with N titratable sites, the protonation state is described by a vector \vec{s} with components s_i ($i = 1, \dots, N$). The s_i enumerate all possible protonation states of the site i and s_i runs from 0 to $P_i - 1$, where P_i is the number of protonation states of the site. It has been shown

[49] [50] that the pH-dependent free-energy difference, $\Delta G(\bar{s}, \text{pH})$, between a protein in a specific protonation state \bar{s} and the fully unprotonated state $\bar{0}$ is given by *Eqn. 2*:

$$\Delta G(\bar{s}, \text{pH}) = (\ln 10)k_B T \sum_{i=1}^N n_i(s_i) (\text{pH} - \text{p}K_i^{\text{std}}(s_i)) + E_{\text{el}}(\bar{s}) - E_{\text{el}}(\bar{0}) - \sum_{i=1}^N [E_{\text{el}}^{\text{M}_i}(s_i) - E_{\text{el}}^{\text{M}_i}(0)] \quad (2)$$

where $\text{p}K_i^{\text{std}}(s_i)$ is the standard $\text{p}K$ for the reaction corresponding to the protonation state s_i



with $n_i(s_i)$ the number of protons bound to site i in protonation state s_i , and M_i the model compound for site i . $E_{\text{el}}(\bar{s})$ and $E_{\text{el}}(\bar{0})$ are the electrostatic free energies of the molecule in protonation state \bar{s} and in the fully unprotonated reference state $\bar{0}$, resp. $E_{\text{el}}^{\text{M}_i}(s_i)$ and $E_{\text{el}}^{\text{M}_i}(0)$ are the electrostatic free energies of the model compound for site i in protonation state s_i and in the unprotonated state, resp. The $E_{\text{el}}(\bar{s})$ values were calculated by solving the finite-difference linearized *Poisson–Boltzmann* (LPB) equation using the program UHBD [51]. All calculations were carried out at 300 K, with an ionic strength of 100 mM, and a *Stern* (ion exclusion) layer of 2 Å. Several focusing grids were used to achieve a final grid spacing of 0.30 and a 3 Å border space. Dielectric constants of 80 and 20 were attributed to the solvent and to the volume occupied by the solute, resp. The atomic radii were taken from the CHARMM22 atom types [52], and the partial atomic charges were calculated as described above. Once the electrostatic free energies $E_{\text{el}}(\bar{s})$ and $E_{\text{el}}^{\text{M}_i}(s_i)$ have been determined for all the 2⁷ protonation states, the probability $X(s'_i)$ of finding the site i in the state s'_i for a given pH value can be obtained by evaluating the *Boltzmann*-weighted sum over all protonation states \bar{s} :

$$X(s'_i)(\text{pH}) = \frac{1}{Z(\text{pH})} \sum_{\bar{s}} \delta_{s,s'_i} \exp\left(\frac{-\Delta G(\bar{s}, \text{pH})}{k_B T}\right) \quad (3)$$

where $Z(\text{pH})$ is the protonation partition function of the molecule.

$$Z(\text{pH}) = \sum_{\bar{s}} \exp\left[\frac{-\Delta G(\bar{s}, \text{pH})}{k_B T}\right]$$

The *Kronecker* symbol δ_{s,s'_i} is equal to 1 if $s_i = s'_i$ and 0 otherwise. The effective $\text{p}K_a$ values for the different sites i are then determined by calculating the pH at which $X(s'_i)(\text{pH}) = 0.5$.

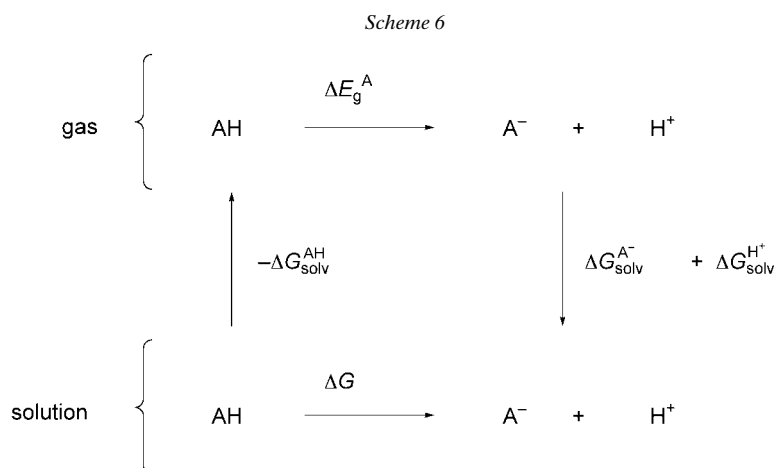
The advantage of decomposing the molecule into subgroups, over an approach based entirely on quantum calculations performed on the entire molecule in the different protonation states, is twofold. First, quantum calculations are only required for a small number of smaller molecules, and, second, conformational effects can be taken into account. Due to the considerable size of **1**, determining all data required from density-functional-theory calculations would already be prohibitive or even impossible. The approach used here requires only ten quantum calculations, corresponding to the four protonation states of **2**, the four protonation states of **3**, and the two protonation states of **4**. Treating **1** as an entity would have required 128 quantum calculations, each for the entire molecule. Also, it is possible to estimate the effective $\text{p}K_a$ value of analogues of **1** based on the same building blocks without the need for new demanding quantum calculations.

Standard $\text{p}K_a$ of Phenol (4). It is worthwhile to briefly comment on the estimation of the effective $\text{p}K_a$ of phenol, which serves as an example how electronic coupling influences the results. The measured $\text{p}K_a$ value of phenol in soln. is 9.84, while the experiments reported here show that the $\text{p}K_a$ value of the phenol moiety in **1** is 8.90. Thus, the presence of the benzimidazole lowers the $\text{p}K_a$ by about one unit. It was not possible to reproduce this $\text{p}K_a$ shift by consideration of electrostatic interactions only. To account for electronic coupling with the benzimidazole, we used the thermodynamic cycle shown in *Scheme 6*.

According to this cycle, the free energy of the acid ionization in solution, ΔG , can be written as *Eqn. 4*.

$$\Delta G = -\Delta G_{\text{solv}}^{\text{AH}} + \Delta E_g^{\text{A}} + \Delta G_{\text{solv}}^{\text{A}^-} + \Delta G_{\text{solv}}^{\text{H}^+} \quad (4)$$

Here, ΔE_g^{A} is the energy of deprotonation in the gas phase, which is obtained from quantum calculations. Entropy differences between AH and A[−] are neglected. $\Delta G_{\text{solv}}^{\text{AH}}$, $\Delta G_{\text{solv}}^{\text{A}^-}$, and $\Delta G_{\text{solv}}^{\text{H}^+}$ are the solvation free energies



of the acid AH, the base A⁻, and the proton, resp., which are calculated by solving the *Poisson–Boltzmann* equation. The difference in the free energies for the ionization of the phenol moiety in **1** and of phenol itself (**4**) can then be written as:

$$\Delta\Delta G = -(\Delta G_{\text{solv}}^1 - \Delta G_{\text{solv}}^4) + (\Delta E_g^1 - \Delta E_g^4) + (\Delta G_{\text{solv}}^{1-} - \Delta G_{\text{solv}}^{4-}) \quad (5)$$

We found $\Delta G_{\text{solv}}^1 - \Delta G_{\text{solv}}^4 = -10.41 \text{ kcal mol}^{-1}$, $\Delta E_g^1 - \Delta E_g^4 = -14.53 \text{ kcal mol}^{-1}$, and $\Delta G_{\text{solv}}^{1-} - \Delta G_{\text{solv}}^{4-} = 2.81 \text{ kcal mol}^{-1}$, leading to $\Delta\Delta G = -1.31 \text{ kcal mol}^{-1}$ and $\Delta\text{p}K_a = \Delta\Delta G/(2.3RT) = -0.96$. Thus, the $\text{p}K_a$ value of phenol (9.84) has to be corrected by -0.96 to account for the presence of benzimidazole, which gives a standard $\text{p}K_a = 8.88$ for the OH group in **1**, in good agreement with experiment (8.90).

REFERENCES

- [1] H. Loewe, J. Urbanietz, *Arzneim.-Forsch. (Drug Res.)* **1974**, *24*, 1927.
- [2] F. G. Loontjens, L. W. McLaughlin, S. Diekmann, R. M. Clegg, *Biochemistry* **1991**, *30*, 182.
- [3] T. Stokke, H. B. Steen, *J. Histochem. Cytochem.* **1985**, *33*, 333.
- [4] T. Araki, A. Yamamoto, M. Yamada, *Histochemistry* **1987**, *87*, 331.
- [5] K. Steinmetzer, K.-E. Reinert, *J. Biomol. Struct. Dyn.* **1998**, *15*, 779.
- [6] L. Denison, A. Haigh, G. D’Cunha, R. F. Martin, *Int. J. Radiat. Biol.* **1992**, *61*, 69.
- [7] F. G. Loontjens, P. Regenfus, A. Zechel, L. Dumortier, R. M. Clegg, *Biochemistry* **1990**, *29*, 9029.
- [8] S. Y. Breusegem, F. G. Loontjens, P. Regenfuss, R. M. Clegg, *Methods Enzymol.* **2001**, *340*, 212.
- [9] S. Y. Breusegem, R. M. Clegg, F. G. Loontjens, *J. Mol. Biol.* **2002**, *315*, 1049.
- [10] Y. Bathini, K. E. Rao, R. G. Shea, J. W. Lown, *Chem. Res. Toxicol.* **1990**, *3*, 268.
- [11] R. F. Martin, N. Holmes, *Nature* **1983**, *302*, 452.
- [12] K. D. Harshman, P. B. Dervan, *Nucleic Acids Res.* **1985**, *13*, 4825.
- [13] J. Portugal, M. J. Waring, *Biochim. Biophys. Acta* **1988**, *949*, 158.
- [14] P. E. Pjura, K. Grzeskowiak, R. E. Dickerson, *J. Mol. Biol.* **1987**, *197*, 257.
- [15] M. Teng, N. Usman, C. A. Frederick, A. H.-J. Wang, *Nucleic Acids Res.* **1988**, *16*, 2671.
- [16] J. R. Quintana, A. A. Lipanov, R. E. Dickerson, *Biochemistry* **1991**, *30*, 10294.
- [17] M. A. A. F. d. C. T. Carrondo, M. Coll, J. Aymami, A. H.-J. Wang, G. A. van der Marel, J. H. van Boom, A. Rich, *Biochemistry* **1989**, *28*, 7849.
- [18] M. Sriram, G. A. van der Marel, H. L. P. F. Roelen, J. H. van Boom, A. H.-J. Wang, *EMBO J.* **1992**, *11*, 225.
- [19] N. Spink, D. G. Brown, J. V. Skelly, S. Neidle, *Nucleic Acids Res.* **1994**, *22*, 1607.
- [20] M. C. Vega, I. Garcia Saez, J. Aymami, R. Eritja, G. A. Van der Marel, J. H. van Boom, A. Rich, M. Coll, *Eur. J. Biochem.* **1994**, *222*, 721.

- [21] A. Fede, M. Billeter, W. Leupin, K. Wüthrich, *Structure* **1993**, *1*, 177.
- [22] M. S. Searle, K. J. Embrey, *Nucl. Acids Res.* **1990**, *18*, 3753.
- [23] K. J. Embrey, M. S. Searle, D. J. Craik, *J. Chem. Soc., Chem. Commun.* **1991**, 1770.
- [24] K. J. Embrey, M. S. Searle, D. J. Craik, *Eur. J. Biochem.* **1993**, *211*, 437.
- [25] J. A. Parkinson, J. Barber, K. T. Douglas, J. Rosamund, D. Sharples, *J. Chem. Soc., Chem. Commun.* **1989**, 1023.
- [26] J. A. Parkinson, J. Barber, T. D. Kenneth, J. Rosamond, D. Sharples, *Biochemistry* **1990**, *29*, 10181.
- [27] J. A. Parkinson, J. Barber, B. A. Buckingham, K. T. Douglas, G. A. Morris, *Magn. Reson. Chem.* **1992**, *30*, 1064.
- [28] S. E. S. Ebrahimi, J. A. Parkinson, K. R. Fox, J. H. McKie, J. Barber, K. T. Douglas, *J. Chem. Soc., Chem. Commun.* **1992**, 1398.
- [29] A. Fede, A. Labhardt, W. Bannwarth, W. Leupin, *Biochemistry* **1991**, *30*, 11377.
- [30] C. Bailly, P. Colson, J.-P. Henichart, C. Houssier, *Nucleic Acids Res.* **1993**, *21*, 3705.
- [31] M. M. McHugh, J. M. Woynarowski, R. D. Sigmund, T. A. Beerman, *Biochem. Pharmacol.* **1989**, *38*, 2323.
- [32] A. Y. Chen, C. Yu, B. Gatto, L. F. Liu, *Proc. Natl. Acad. Sci., U.S.A.* **1993**, *90*, 8131.
- [33] J. M. Woynarowski, R. D. Sigmund, T. A. Beerman, *Biochemistry* **1989**, *28*, 3850.
- [34] H. Walba, R. W. Isensee, *J. Am. Chem. Soc.* **1955**, *77*, 5488. H. Walba, R. W. Isensee, *J. Org. Chem.* **1961**, *26*, 2789.
- [35] M. Capponi, I. G. Gut, B. Hellrung, G. Persy, J. Wirz, *Can. J. Chem.* **1999**, *77*, 605.
- [36] A. J. Kresge, H. J. Chen, G. L. Capen, M. F. Powell, *Can. J. Chem.* **1983**, *61*, 249.
- [37] N. Z. Bjerrum, *Z. Phys. Chem.* **1923**, *106*, 219. F. H. Westheimer, M. W. Shookhoff, *J. Chem. Phys.* **1939**, *6*, 555.
- [38] M. T. Davies, P. Mamalis, V. Petrow, B. Sturgeon, *J. Pharm. Pharmacol.* **1951**, *3*, 420.
- [39] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.* **1983**, *4*, 187.
- [40] V. N. Umetskaya, Y. M. Rozanov, *Biophysics (USSR)* **1990**, *35*, 404; *Biofizika (in Russian)* **1990**, *35*, 399. V. N. Umetskaya, A. A. Roslov, Y. M. Rozanov, A. V. Garabadzhiu, *Molek. Biol. (in Russian)*, **1989**, 379.
- [41] C. Alemán, A. Adhikary, D. Zanuy, J. Casanovas, *J. Biomol. Struct. Dyn.* **2002**, *20*, 301.
- [42] H. Görner, *Photochem. Photobiol.* **2001**, *73*, 339; G. Cosa, K.-S. Focsaneanu, J. R. N. McLean, J. P. McNamee, J. C. Scaiano, *Photochem. Photobiol.* **2001**, *73*, 585. V. N. Umetskaya, A. A. Roslov, Y. M. Rozanov, A. V. Garabadzhiu, *Molek. Biol. (in Russian)*, **1989**, *23*, 379.
- [43] W. den Hoeve, C. G. Kruse, J. M. Luteyn, J. R. G. Thiecke, H. Wynberg, *J. Org. Chem.* **1993**, *58*, 5101.
- [44] H. Sigel, A. D. Zuberbühler, O. Yamauchi, *Anal. Chim. Acta.* **1991**, *255*, 63.
- [45] H. Gampp, M. Maeder, C. J. Meyer, A. D. Zuberbühler, *Talanta* **1985**, *32*, 95, 257, and 1133; H. Gampp, M. Maeder, C. J. Meyer, A. D. Zuberbühler, *Talanta* **1986**, *33*, 943.
- [46] A. D. Zuberbühler, T. A. Kaden, *Talanta* **1982**, *29*, 206.
- [47] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. H. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople, Gaussian, Inc., Pittsburgh PA, Gaussian 98. Revision A.2.
- [48] A. E. Reed, R. B. Weinstock, F. Weinhold, *J. Chem. Phys.* **1985**, *83*, 735.
- [49] M. Schaefer, M. Sommer, M. Karplus, *J. Phys. Chem. B* **1997**, *101*, 1663.
- [50] H. W. T. van Vlijmen, M. Schaefer, M. Karplus, *Proteins: Struct., Funct. Genet.* **1998**, *33*, 145.
- [51] J. D. Madura, J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, J. A. McCammon, *Comp. Phys. Comm.* **1995**, *91*, 57.
- [52] A. D. MacKerell, D. Bashford, R. L. Bellot, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wierkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586.

Received September 29, 2004