Theor Chem Acc (2000) 103:483–494
DOI 10.1007/s002149900082

Theoretical
Chemistry Accounts

Regular article

Computation of the influence of chemical substitution on the pK_a of pyridine using semiempirical and ab initio methods

I-Jen Chen, Alexander D. MacKerell Jr

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 N. Pine Street, Baltimore, MD 21201, USA

Received: 27 April 1999 / Accepted: 27 July 1999 / Published online: 2 November 1999

Abstract. The physical properties of chemicals are strongly influenced by their protonation state, affecting, for example, solubility or hydrogen-bonding characteristics. The ability to accurately calculate protonation states in the form of pK_as is, therefore, desirable. Calculations of pK_a changes in a series of substituted pyridines are presented. Computations were performed using both ab initio and semiempirical approaches, including free energies of solvation via reaction-field models. The selected methods are readily accessible with respect to both software and computational feasibility. Comparison of calculated and experimental pK_as shows the experimental trends to be reasonably reproduced by the computations with root-mean-square differences ranging from 1.22 to 4.14 pK_a units. Of the theoretical methods applied the best agreement occurred using the second-order Møller–Plesset/6-31G(d)/isodensity surface polarized continuum solvation model, while the more computationally accessible Austin model 1/Solvent model 2 (SM2) approach yielded results similar to the ab initio methods. Analysis of component contributions to the calculated pK_as indicates the largest source of error to be associated with the free energies of solvation of the protonated species followed by the gas-phase protonation energies; while the latter may be improved via the use of higher levels of theory, enhancements in the former require improvements in the solvation models. The inclusion of alternate minimum in the computation of pK_as is also indicated to contribute to differences between experimental and calculated pK_a values.

Key words: pK_a – Quantum mechanical – Pyridine – Semiempirical – Ab initio

Introduction

The exact state of protonation is a key factor dictating the function of many biological molecules [1–5]. Changes in the protonation state of molecules of biological and pharmacological relevance can have a drastic impact on their physicochemical properties, such as solubility and lipophilicity, thereby affecting their absorption and disposition characteristics. Recently, computer-based studies have been performed from which good correlations between calculated free energies of aqueous solvation and blood-brain partitioning [6] and atomic solvent accessibilities and drug absorption [7] were obtained. In those studies, however, functional groups with pK_a values close to physiological pH were not included. Studies in our laboratory have extended the works cited previously to include the influence of substituents on drug permeability [8]. The compounds used in this study were substituted pyridines. Since the pK_a of the ring nitrogen in pyridine (pK_a = 5.25) is close to physiological pH there is a possibility that its protonation state could be influenced by substituents, thereby affecting the calculated physical properties as well as the experimentally measured permeabilities. This problem motivated the present investigation into computational approaches for the determination of pK_a differences in substituted pyridines.

Computational determination of pK_as has been applied to both small molecules [9] and macromolecules, typically proteins [10–13]. In pK_a calculations changes in free energies of solvation of the neutral and charged species are required as are free energies of protonation in the gas phase (see Sect. 2). Determination of absolute pK_a values also requires the free energy of solvation of the proton, a value that is not exactly known [14]. Accordingly, most pK_a calculations involve the determination of pK_a differences, although small molecule studies have attempted to determine absolute pK_as [9]. Small molecule pK_a studies have used a variety of techniques to obtain the free energies of solvation, including free-energy perturbation methods [15–17], quantum mechanical (QM) molecular mechanical free-energy perturbations [18], continuum electrostatic approaches that involve solving the Poisson–Boltzmann or generalized Born equations [19–21], and reaction-field methods [14, 22]. The majority of pK_a shifts in proteins are calculated using continuum methods based on the Poisson–Boltzmann

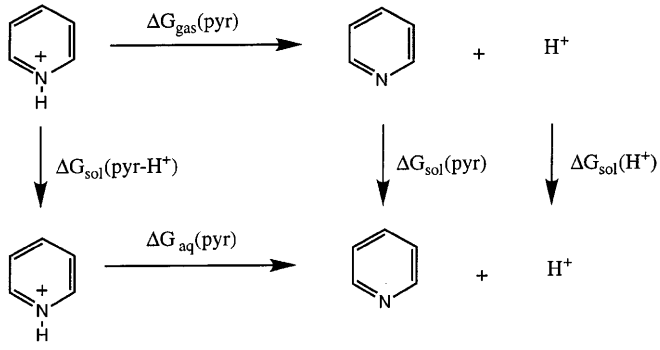
Correspondence to: A.D. Mackerell Jr
e-mail: alex@mmiris.ab.umd.edu

equation [23] due to computational considerations, although semimicroscopic approaches have been applied [24]. Of these approaches quantum-chemical-based methods combined with reaction-field models are attractive for small molecule studies due to their ability to determine gas-phase deprotonation energies and free energies of solvation in a computationally accessible manner.

Results on pK_as calculated using both ab initio and semiempirical approaches, including several reaction-field solvation models [25, 26], are presented. In Sect. 2 the methods of the computational approaches are described along with the equations for the calculation of pK_a differences. Section 3 includes data from a collection of results from the different levels of theory and includes direct comparison of pK_a, proton affinity (PA) and free energy of solvation data with experimental values. A brief conclusion is used to summarize and highlight the results.

2 Computational methods

pK_a differences in the present paper were calculated for substituted pyridines (xpyr) relative to unsubstituted pyridine (pyr) based on the thermodynamic cycle shown in Scheme 1 [9].



The absolute pK_a for unsubstituted pyridine, pK_a(pyr), is related to the free energies of solvation, $\Delta G_{sol}(pyrH^+)$ and $\Delta G_{sol}(pyr)$, for the protonated and neutral forms of pyridine, respectively, the free energies of solvation of the proton, $\Delta G_{sol}(H^+)$, and the deprotonation energy of pyridine in the gas phase, $\Delta G_{gas}(pyr)$, as shown in Eq. (1).

$$pK_a(pyr) = \frac{1}{2.3RT}[\Delta G_{sol}(pyr) + \Delta G_{sol}(H^+) - \Delta G_{sol}(pyrH^+) + \Delta G_{gas}(pyr)] \quad (1)$$

Similarly, Eq. (1) can be written for the pK_a of a substituted pyridine, pK_a(xpyr), as shown in Eq. (2).

$$pK_a(xpyr) = \frac{1}{2.3RT}[\Delta G_{sol}(xpyr) + \Delta G_{sol}(H^+) - \Delta G_{sol}(xpyrH^+) + \Delta G_{gas}(xpyr)] \quad (2)$$

Subtracting Eq. (1) from Eq. (2) yields the difference in pK_a, $\Delta pK_a(xpyr - pyr)$ as shown in Eq. (3),

$$\Delta pK_a(xpyr - pyr) = \frac{1}{2.3RT}[\Delta\Delta G_{sol}(xpyr - pyr) - \Delta\Delta G_{sol}(xpyrH^+ - pyrH^+) + \Delta\Delta G_{gas}(xpyr - pyr)] \quad , \quad (3)$$

where $\Delta\Delta G_{sol}(xpyr - pyr)$ and $\Delta\Delta G_{sol}(xpyrH^+ - pyrH^+)$ are the differences between the free energies of solvation of substituted pyridine and pyridine for the neutral and protonated species, respectively, and $\Delta\Delta G_{gas}(xpyr - pyr)$ is the difference in the gas-phase deprotonation energies of the substituted pyridine and pyridine. Importantly, the free energy of solvation of the proton cancels. Thus, only the difference in the gas-phase deprotonation energies and the solvation-free-energy differences for both the protonated (cationic) and neutral species of pyridine are needed to determine the pK_a differences. Calculation of the gas-phase deprotonation energies were performed using both ab initio and semiempirical methods and the free energies of solvation were obtained by using the different reaction-field models as presented later.

Gas-phase geometry optimizations were performed via both ab initio and semiempirical QM methods. Ab initio calculations were performed using Gaussian 94 [27] with the 6-31G(d) basis set. Geometry optimizations were performed at the Hartree–Fock (HF) level of theory. All Møller–Plesset perturbation theory to second-order (MP2) level calculations were performed using the HF/6-31G(d) optimized geometry. Semiempirical calculations were carried out using the Austin model 1 (AM1) Hamiltonian [28] via the program AMSOL 5.0 [29].

For several of the compounds studied rotation of the substituent leads to multiple minima. These include the 2-OH, 2-CHO, 2-C_2H_5, and all methyl-substituted pyridines. To identify their global minima these compounds were subjected to the following procedure. Potential-energy surfaces were obtained via AM1 calculations for both the neutral and protonated species by fixing the dihedral angle defining the rotation of the substituent (e.g. the O-C-C2-C3 dihedral angle of the 2-CHO pyridine analog) and allowing all other degrees of freedom to relax. The solvated surfaces (see later) were obtained at the AM1–SM2 level of theory using the structures from the gas-phase surfaces. From each surface the global minimum was identified and used for the determination of the AM1-SM2 pK_as and also as the starting geometry for the HF/6-31G(d) optimizations. Potential-energy surfaces were not performed at the HF/6-31G(d) level due to computational considerations. While individually such calculations are feasible, the goal was to identify a method that would allow pK_a calculations on a large number of compounds such that a significant number of HF/6-31G(d) potential-energy surfaces becomes prohibitive.

Free energies of solvation in aqueous solution were calculated using a variety of reaction-field methods. All solvation calculations were performed on the gas-phase optimized structures. Due to the pyridine ring, the presence of solvent is not expected to have a significant influence on the optimized geometries, as was verified in several tests (not described). Reaction-field models include AM1–SM2 [30, 31] implemented in AMSOL [29], and the isodensity surface polarized continuum model (IPCM) [32] and the self-consistent IPCM (SCIPCM) of Miertus et al. [33] and Cossi et al. [34]. The IPCM and SCIPCM are implemented in the Gaussian 94 package. For all ab initio reaction-field calculations the dielectric of water was set to 78.36 and the isodensity cutoff values were 0.001 a.u. for the IPCM and 0.0004 a.u. for the SCIPCM. Solvation free energies in aqueous solution were obtained as the difference between the total free energies in aqueous solution and the gas-phase heat of formation. pK_a values of the substituted pyridines were calculated via Eq. (3) assuming the pK_a of pyridine to be 5.25 [35] and a temperature of 298 K, with the exception of the 2-NH_2, 2-CHO, 2-OH, and 4-OH substituted compounds, where a temperature of 293 K was applied, consistent with experiment.

No free-energy (e.g. vibrational) corrections were taken into account in the present calculations, except for the calculated gas-phase PAs, based on the assumption that their contributions in the different substituted pyridines would be similar and, therefore, cancel. This was tested by including free-energy corrections into the gas-phase deprotonation energies based on HF/6-31G(d) frequency calculations [36]. Inclusion of these corrections yielded changes in the calculated pK_as ranging from 0.00 to 0.19 pK_a units supporting the above assumption.

The nomenclature describing the level of theory used throughout the manuscript is as follows. AM1/SM2 indicates the AM1 optimized structures and gas-phase deprotonation energies with the SM2 solvation models. HF/SM2 indicates the HF/6-31G(d) gas-

phase deprotonation energies with the AM1/SM2 solvation energies calculated using HF/6-31G(d) optimized geometries. MP2/SM2 indicates the HF/6-31G(d) optimized structure with the single-point MP2/6-31G(d) gas-phase deprotonation energies and the AM1/SM2 solvation energies. SM2 can be substituted with IPCM and SCIPCM to indicate use of the different solvation models. In those cases the HF/6-31G(d) gas-phase optimized structure was used for the reaction-field calculations.

3 Results and discussion

The compounds included in the present study are those for which experimental pK_a values are available [35] and were included in our study on drug permeability [8]. In addition, four compounds, three dimethyl-substituted pyridines and one bromide meta-substituted pyridine, were studied due to both pK_a [35] and gas-phase PAs and/or free energy of solvation experimental data [37] being available for these compounds. Both the position of the substituents and the types of functional groups vary significantly. This variability allows a more rigorous test of the computational approaches for the determination of pK_a differences. Emphasis was placed on both accuracy and computational feasibility. While the highest accuracy is always desirable, in order to apply a method to a large number of compounds as is often required in rational drug design studies, computational accessibility is necessary. Examples of recently published works using higher levels of theory are discussed in the concluding section. Preliminary studies using the Onsager model [38–41] yielded poor agreement with experiment (results not given).

3.1 Tautomers of 2- and 4-hydroxyl pyridines

Among the compounds studied, the 2- and 4-hydroxyl substituted pyridines are in tautomeric equilibria with the corresponding pyridones [40, 42–45], as shown in Fig. 1. It has been observed experimentally that the tautomerization is medium-dependent. In the aqueous phase the pyridone forms are predominant while in the gas phase the hydroxyl forms are preferred [42, 45, 46]. To verify these results calculations were performed to obtain the tautomerization energies, with the results being presented in Table 1. As may be seen the tautomerization energies in both the gas phase and with the solvation models are dependent on the theory applied, in agreement with previous studies [40, 43]. In the gas phase the HF/6-31G(d) calculation indicates the pyridone form of the 2-hydroxyl analog to be favored; however, both the

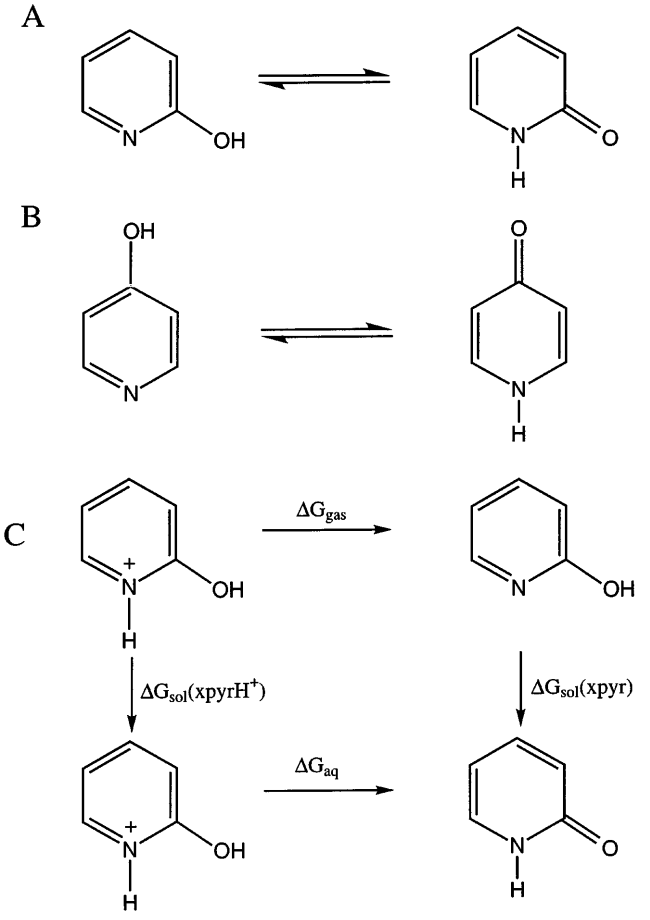


Fig. 1. Diagram of the tautomers of **A** the 2-hydroxyl-, **B** the 4-hydroxyl pyridines and **C** the species used in the pK_a calculations of the hydroxyl pyridines

Table 1. Tautomerization energies (kcal/mol) of 2-hydroxyl and 4-hydroxyl pyridines

Compound[a]	Level of theory[b,c]				
	HF/6-31G(d)	MP2/6-31G(d)	AM1	Exp.[d]	
Gas phase					
2-Hydroxyl	−0.09	1.61	0.47	0.4 ± 0.7	
4-Hydroxyl	2.71	4.54	8.17	> 1.36	
	HF/I	MP2/I	HF/S	AM1/SM2	Exp.[d]
Aqueous phase					
2-Hydroxyl	−5.8	−2.2	−4.3	−1.8	−4.1
4-Hydroxyl	−12.0	−6.1	−2.7	1.4	−4.5

[a] Energies represent the conversion from the hydroxyl to oxo forms
[b] All calculations were done using HF/6-31G(d) optimized geometries except for AM1
[c] I indicates the isodensity surface polarized continuum mode (*IPCM*) and S the self-consistent IPCM
[d] Beak et al. [12]. No errors were reported for the aqueous-phase data

MP2/6-31G(d) and AM1 levels indicate the hydroxyl form to predominate, consistent with the experimental data, although the quantitative agreement is poor. For 2-hydroxyl pyridine in the aqueous phase all calculations indicate the pyridone form to predominate, consistent with experiment although the quantitative agreement is again poor. In the case of 4-hydroxyl pyridine the gas-phase calculations all indicate the hydroxyl form to be favored, which is consistent with experiment. In aqueous solution, three of the four approaches applied predict the pyridone to be favored; only the AM1/SM2 level of theory predicts the 4-hydroxyl form to be favored, however, the magnitude is smaller than that for the other levels of theory. Based on both the experimental and calculated data further calculations on the 2-hydroxyl and 4-hydroxyl pyridines were performed using the hydroxyl forms in the gas phase and for the protonated form in the aqueous solution, while the pyridone forms were used for the neutral form in aqueous solution (Fig. 1C).

3.2 Gas phase

Experimental data on the gas-phase protonation energies are available for most of the compounds included in the present study. This data allows the levels of theory used in the present study to be quantified, such that discrepancies between experimental and calculated values of the pK_as may be understood in greater detail. The total energies of each compound in the protonated and neutral states were calculated at the HF/6-31G(d), MP2/6-31G(d), and AM1 levels of theory. The PAs based on these calculations are presented in Table 2 along with the available experimental data. Root-mean-square (rms) differences of the absolute PAs at the HF and MP2 levels were 5.8 and 1.0 kcal/mol, while the AM1 values are not directly comparable to experiment (Table 2). Inclusion of electron correlation clearly improved the predicted values of the PAs. The PA values predicted by HF, MP2, and AM1 follow the experimental trend in a satisfactory manner as indicated by correlation coefficients of 0.98, 0.99, and 0.90, respectively. Analysis of the data in Table 2 shows the HF PAs to be consistently too large, while the MP2 values are scattered around the experimental data, as emphasized by the average difference of 0.1. While both the HF and AM1 calculated gas-phase PA values are in relatively poor agreement with experiment, it should be emphasized that it is the differences in these values as a function of substituent that are required for determination of pK_a differences, as shown in Eq. (3). Additional details of the relative PA values are presented later.

Experimentally it has been observed [47–51] that electron-withdrawing substituents (e.g. Cl and CHO) lower the acidities of the substituted pyridinium ions by diminishing the electron density around the ring nitrogen and the acidic hydrogen, while electron-donating groups (e.g. NH_2) increase the pyridinium acidity by stabilizing the electron-deficient ring nitrogen [52, 53]. For the compounds studied for which experimental PAs are available, the amino substituent is the strongest electron-donor group and two aminopyridines in our calculations indeed have higher PAs than the parent pyridine. The position of substitution also seems to play a role in determining the corresponding PA values. The 4-amino pyridinium ion has a stronger PA than the 2-amino analog partly due to greater delocalization of the positive charge on the ring nitrogen. The halogen-substituted pyridines gave the lowest PAs in the calculations since they are the most electronegative atoms on the list. Their PA values correlate well with their electronegativity. There are also several alkyl pyridines in the series. As expected, they all have higher PAs than pyridine itself due to their electron-donating nature. The preferential charge-induced dipole stabilization of the alkyl-substituted ion and, to a lesser extent, the electron-donating inductive effects contribute to their elevated PAs [54]. The polarizability effects of the alkyl groups have been suggested to be more important in the gas phase than in the aqueous phase, which leads to a solvent attenuation factor greater than 1.0 (see later) [51, 55].

Table 2. Gas-phase proton affinities (*PA*) of pyridines. The PA values were calculated as the sum of enthalpies at 0 K and the unscaled thermal corrections (vibrational) under standard conditions of $T = 298.15$ K and $P = 1$ atm. The average difference and the root mean-square (rms) difference are the differences between the calculated and experimental values for the compounds for which experimental data is available

	Exp.[a]	HF/6-31G(d)	MP2/6-31G(d)	AM1
PA				
2-C_2H_5	224.9	230.3	224.5	−148.3
2-CH_3	223.7	230.0	224.3	−148.3
2-CHO		214.1	210.5	−158.5
2-NH_2	223.8	231.6	224.3	−144.5
2-OH		222.6	216.2	−148.3
3-CH_3	222.8	228.1	223.2	−148.3
3-COO^-		316.7	314.0	−73.9
3-Cl	215.7	218.1	214.0	−156.6
4-CH_3	223.7	230.1	224.2	−149.2
4-NH_2	231.0	239.3	232.7	−140.1
4-OH		230.0	224.1	−150.0
2-3-DiCH_3	226.2	232.2	227.0	−146.7
2-4-DiCH_3	227.1	233.5	227.2	−145.5
3-5-DiCH_3	225.6	230.1	225.7	−149.0
3-Br	216.3	218.8	214.2	−156.9
Pyridine	220.4	225.8	220.5	−152.1
Average difference		5.56	0.05	
Rms difference		5.82	1.00	

[a] Aue et al. [37]

3.3 Free energy of aqueous solvation

Analysis of Eq. (3) shows that the relative free energies of solvation of the different substituted species in both the neutral and protonated states are required for the determination of the pK_as. Accordingly, the free energies of aqueous solvation were calculated, with the results being presented in Table 3 along with the available experimental data. For the neutral compounds,

Table 3. Absolute aqueous solvation free energies (kcal/mol) of pyridines

	Exp.[a]	HF/I	MP2/I	HF/S	AMI/SM2
Neutral					
2-C_2H_5	−2.4	−3.9	−3.4	−4.5	−2.6
2-CH_3	−2.7	−4.5	−4.0	−4.6	−3.4
2-CHO	−4.3[b]	−6.1	−4.6	−7.7	−6.8
2-NH_2		−7.0	−6.4	−7.9	−8.8
2-OH		−14.3	−8.5	−9.2	−10.4
3-CH_3	−2.9	−4.7	−4.3	−4.9	−4.1
3-COO^-		−71.6	−66.1	−56.5	−72.0
3-Cl	−2.1	−4.2	−3.6	−5.1	−4.6
4-CH_3	−3.0	−4.7	−4.1	−5.1	−4.2
4-NH_2		−5.8	−5.3	−9.8	−10.1
4-OH		−16.5	−9.7	−10.0	−8.2
2-3-DiCH_3	−2.9	−5.0	−4.0	−3.1	−3.2
2-4-DiCH_3	−3.0	−5.4	−4.4	−3.1	−3.2
3-5-DiCH_3	−3.0	−5.6	−4.8	−3.4	−3.7
3-Br		−4.3	−3.3	−3.4	−5.7
Pyridine	−2.8	−4.5	−3.9	−5.0	−4.4
Average difference		−1.95	−1.20	−1.74	−1.11
Rms difference		1.98	1.20	2.05	1.38
Protonated					
2-C_2H_5	−53.9	−52.3	−52.5	−67.5	−49.1
2-CH_3	−55.5	−54.0	−54.2	−68.2	−50.8
2-CHO		−60.6	−58.7	−73.7	56.6
2-NH_2		−56.9	−57.0	−71.2	−57.5
2-OH		−61.6	−61.5	−73.6	−59.2
3-CH_3	−56.1	−54.9	−54.8	−68.6	−54.6
3-COO^-		−44.9	−35.9	−41.4	−42.6
3-Cl	−58.6	−58.8	−57.9	−74.0	−59.0
4-CH_3	−55.9	−55.3	−55.5	−68.2	−54.6
4-NH_2		−55.1	−54.6	−69.3	−57.7
4-OH		−57.2	−56.8	−74.1	−61.8
2-3-DiCH_3	−54.0	−52.7	−52.0	−54.5	−48.6
2-4-DiCH_3	−53.2	−52.7	−52.4	−54.0	−48.4
3-5-DiCH_3	−54.0	−53.1	−52.0	−55.3	−52.1
3-Br		−61.2	−58.8	−60.8	−58.7
Pyridine	−57.8	−57.8	−57.8	−70.4	−57.1
Average difference		0.82	1.10	−9.08	2.74
Rms difference		1.02	1.27	10.81	3.42

[a] Aue et al. [17]
[b] The hydration free energy of 2-CHO pyridine is from Cabani et al. [56]

the rms differences ranged from 1.2 to 2.1 kcal/mol for the MP2/IPCM and the HF/SCIPCM, respectively. For the protonated species, the rms differences varied from 1.0 to 10.8 kcal/mol, with the best and poorest agreement occurring with the HF/IPCM and the HF/SCIPCM, respectively. Thus, the absolute values of the free energies of solvation for the neutral compounds are reasonably well predicted, while those of the protonated species are generally in poorer agreement with experiment. With the neutral species the calculated free energies of solvation are generally too favorable, as indicated by the negative average differences. This is due to all the solvation models predicting the free energy of solvation of pyridine to be too favorable. As with the gas-phase data, however, it is the relative values of the free energies of solvation (see Eq. 3) that are required for the calculation of the change in pK_a; the quality of the relative values will be discussed later.

3.4 Calculated versus experimental pK_a*s*

Based on Eq. (3) and a pK_a of 5.25 for pyridine, pK_as were determined for the substituted pyridines listed in Table 4. Included in Table 4 along with the calculated pK_as are the corresponding experimental values. The same data is presented in graphical form in Fig. 2, with the compounds ordered along the *x*-axis according to their experimental pK_a values. For the majority of compounds the agreement with experiment is reasonable. Only with the 2-CHO and 4-OH analogs are systematic deviations from experiment calculated for the majority of the methods applied.

To assess the quality of the methods applied for the determination of pK_a differences in a series of substituted compounds the average and rms differences between the calculated and experimental numbers were determined and are included in Table 4. Overall, the quality of the predictions can be considered to be satisfactory. The best method is the MP2/IPCM method and the worst is the HF/SCIPCM method. The poor agreement of the HF/SCIPCM method was somewhat surprising. The self-consistent determination of the solvent accessible surface is expected to yield improvements over the ICPM method. Analysis of the individual values in Table 4 shows some of the HF/SCIPCM values to differ significantly from experiment, including the three disubstituted compounds and the 3-Br substituted compound; additional analysis of these compounds indicates problems associated with the free energy of solvation calculation (see later).

The remaining approaches all yielded similar results with rms differences from 1.9 to 2.6. The computationally least expensive AM1/SM2 method yielded a rms difference of 2.1 with respect to experiment. This value is better than the HF/SM2 and MP2/SM2 results, where improved values are expected due to more accurate determination of the PAs, while the solvation energies are identical. To better understand this result, as well as to investigate the contributions to the pK_a values determined via all the methods applied, component analysis of the data was performed.

3.5 Component contributions to the calculated pK_a*s*

Inspection of Eq. (3) shows the relative gas-phase protonation energies and the relative free energies of solvation of both the neutral and protonated species with respect to pyridine contribute to the calculated changes in pK_as. Detailed analysis of these results was performed by determining these relative energies from both the calculations and from experiment. This data is presented in Tables 5 and 6 for the gas-phase deprotonation energies and free energies of solvation, respectively. In addition, the overall contribution of the free energy of solvation based on the difference between the relative free energies of solvation of the neutral and protonated species is presented in Table 7. The data in Table 7 will be referred to as $\Delta\Delta\Delta G_{sol}$. Comparison of the available experimental values in Tables 5, 6, and 7 with the various calculated data shows the agreement to be

Table 4. Experimental and calculated pK_as for the pyridines studied

	Exp.[a]	HF/I	MP2/I	HF/S	AM1/SM2	HF/SM2	MP2/SM2
2-C_2H_5	5.89	4.91	4.59	6.69	3.54	4.00	3.58
2-CH_3	5.97	5.55	5.32	6.96	4.20	4.45	4.13
2-CHO	3.80	1.15	1.41	0.64	0.92	−1.90	−0.89
2-NH_2	6.82	6.71	5.42	7.64	7.95	6.31	4.83
2-OH	0.75	−2.01	1.02	1.76	1.80	−0.33	−1.15
3-CH_3	5.68	5.05	5.19	6.04	4.80	5.72	6.07
3-COO^-	4.77[b]	6.22	5.19	5.88	2.39	4.70	6.60
3-Cl	2.84	0.33	0.62	1.93	3.18	0.66	1.55
4-CH_3	6.02	6.36	6.08	6.62	5.70	6.61	6.21
4-NH_2	9.11	12.12	10.83	10.77	10.31	11.33	10.39
4-OH	3.20	−1.03	2.83	7.36	7.50	8.99	8.53
2-3-DiCH_3	6.57	5.87	5.67	−0.45	3.78	4.52	4.62
2-4-DiCH_3	6.99	6.87	6.20	0.52	4.62	5.80	5.03
3-5-DiCH_3	6.15	4.99	5.10	−0.73	4.27	6.07	6.76
3-Br	2.84	2.52	1.66	−6.10	1.93	0.12	0.64
Average difference		−0.77	−0.67	−1.45	−0.69	−0.68	−0.69
Rms difference		1.86	1.22	4.14	2.09	2.56	2.43

[a] Ref. [35]
[b] Experimental pK_a of 3-COO^- is from Jaffé and Doak [72]

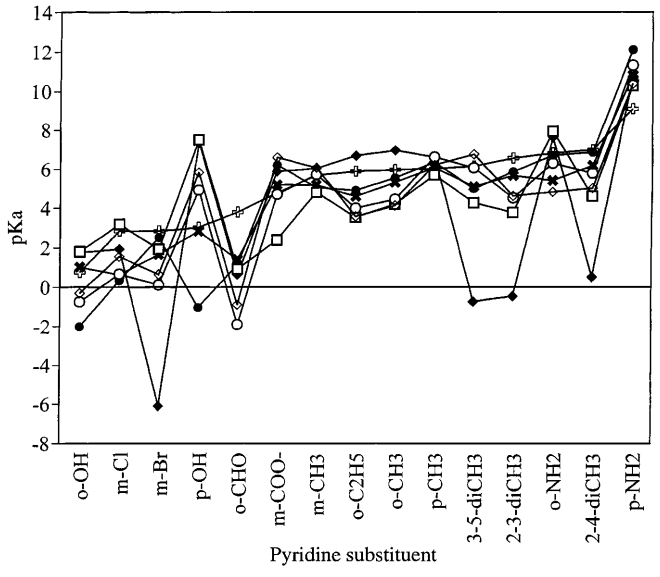


Fig. 2. pK_a values as a function of the pyridine substituent. Compounds are ordered based on increasing experimental pK_a values [35]. Data presented include experimental pK_as (✠) and those calculated at the following levels of theory: HF/IPCM (●), MP2/IPCM (✖), HF/SCIPCM (◆), AM1/SM2 (□), HF/SM2 (○). See Sect. 2 for a description of the levels of theory

reasonable in the majority of cases, consistent with the pK_a data in Table 4. Thus, while the absolute values of the energies (see Tables 2, 3) often differ significantly from experiment, relative energies are typically in much better agreement with experiment, allowing reasonable estimates of changes in pK_a values to be made via the selected methods.

The data in Tables 5, 6, and 7 will first be used to investigate the problems with the calculated pK_as of the 4-OH and 2-CHO substituted species. These compounds show more systematic deviations from experiment compared to the remaining compounds (Fig. 2). With the 4-OH analog, the variation in the calculated pK_as is large. Analysis of the results in Tables 5, 6, and 7 indicates that the larger pK_a variation of the 4-OH analog could be due to both the gas-phase deprotonation energy and the solvation energy of the protonated species. The relative gas-phase deprotonation energies are systematically too large, compared to the negative values for the 2-OH analog. Positive relative protonation energies will yield pK_a differences that are large and positive, consistent with the overestimation of the 4-OH pK_as. Problems in the gas-phase deprotonation energy calculations may also contribute to the large calculated tautomerization energies in the gas phase and the wide variation of the tautomerization energies in solution (Table 1). Also, the solvation energies of the protonated species by the SCIPCM and SM2 methods are too favorable, which may also contribute to the high calculated pK_a values. Moreover, it is suggested that para substitution by the hydroxyl group may contribute to resonance effects that are poorly modeled, especially at the HF level. This is consistent with the MP2 calculated pK_a (IPCM and SM2) being in better agreement with experiment than the HF counterparts, although the absolute MP2/SM2 values are still too high (Table 4).

With the 2-CHO analog the calculated pK_as are consistently lower than the experimental value. Examination of Table 7 shows the overall solvation contributions to be relatively small, ranging from −2.9 to 1.2 kcal/mol, while the gas-phase deprotonation energies are larger in magnitude, ranging from −6.7 to −2.9 kcal/mol (Table 5). The only available experimental value besides the pK_a for 2-CHO is the hydration free energy [56] of the neutral species. The relative free energies of solvation of the 2-CHO analog (Table 6) are satisfactorily treated by the IPCM method, with the SCIPCM and SM2 approaches yielding values that are too negative, which may contribute to the corresponding pK_as being too small. While the lack of gas-phase experimental data prevents rigorous comparison, comparison can be made for 4-CHO, which has an experimental PA of −5.2 kcal/mol relative to pyridine, which

Table 5. Relative deprotonation energies (kcal/mol) versus pyridine in the gas and aqueous phases

	Exp.[a]	HF/6-31G(d)	MP2/6-31G(d)	AMI	
Gas phase					
2-C_2H_5	4.5	4.5	3.9	3.9	
2-CH_3	3.3	4.2	3.8	3.9	
2-CHO		−6.7	−5.3	−2.9	
2-NH_2	3.4	5.4	3.4	7.6	
2-OH		−3.7	−3.2	−0.3	
3-CH_3	2.4	2.8	3.3	1.6	
3-COO^-		81.4	83.9	78.2	
3-Cl	−4.7	−7.9	−6.7	−4.5	
4-CH_3	3.3	4.2	3.6	2.9	
4-NH_2	10.6	13.5	12.2	12.1	
4-OH		4.1	3.5	2.1	
2-3-DiCH_3	5.8	6.4	6.5	5.4	
2-4-DiCH_3	6.7	7.7	6.7	6.6	
3-5-DiCH_3	5.2	4.3	5.2	3.1	
3-Br	−4.1	−7.0	−6.3	−4.8	
	Exp.[b]	HF/I	MP2/I	HF/S	AMI/SM2
Aqueous phase					
2-C_2H_5	0.87	−0.46	−0.90	1.97	−2.33
2-CH_3	0.98	0.41	0.09	2.33	−1.43
2-CHO	−2.06	−5.62	−5.27	−6.31	−5.93
2-NH_2	1.98	1.84	0.11	3.08	3.50
2-OH	−6.15	−9.85	−5.79	−4.80	−4.74
3-CH_3	0.59	−0.28	−0.08	1.08	−0.62
3-Cl	−3.29	−6.71	−6.32	−4.52	−2.82
3-COO^-	−0.65[c]	1.33	−0.08	0.85	−3.90
4-CH_3	1.05	1.52	1.14	1.87	0.61
4-NH_2	5.26	9.37	7.61	7.53	6.90
4-OH	−3.11	−8.53	−3.36	2.71	2.90
2-3-DiCH_3	1.80	0.85	0.57	−7.77	−2.00
2-4-DiCH_3	2.37	2.21	1.30	−6.44	−0.85
3-5-DiCH_3	1.23	−0.35	−0.21	−8.15	−1.33
3-Br	−3.29	−3.72	−4.90	−15.48	−4.53

[a] Aue et al. [37]
[b] Experimental free energies of deprotonation were converted from the corresponding pK_a values in Ref. [35]
[c] The experimental free energy of 3-COO^- pyridine was converted from the pK_a value reported by Jaffé and Doak [72]

compares with HF/6-31G(d) and MP2 single-point values of −7.2 and −6.1 kcal/mol, respectively. This suggests that the error in the pK_a values with the 2-CHO analog may also be partly due to overestimation of the PA. While the level of theory used in the present calculations may be the problem, the intramolecular hydrogen bond between the formyl hydrogen and the pyridine nitrogen may also contribute. Additional details of the possible effects associated with this interaction are presented later.

It should be noted that the 2-CHO analog is in equilibrium with its hydrated tautomeric form at neutral pH [57]. Calculations, therefore, were performed at the HF/IPCM and MP2/IPCM levels using the zwitterionic tautomer of the hydrated species as the alternate final product in aqueous solution. From these calculations pK_as of 8.3 and 6.9 were determined, respectively, in poorer agreement with the experimental value of 3.8 compared to the calculated values of 1.2 and 1.4 for the nonhydrated species. Due to this poorer agreement additional analysis of the hydrated 2-CHO analog was not performed, although the possibility that this species may contribute to the experimental pK_a cannot be excluded.

The data in Tables 5, 6, and 7 may also be used to understand the overall quality of agreement of the calculated pK_as for the different methods applied. To aid in this comparison, average and rms differences of the relative energies in Tables 5 and 6 were determined and are presented in Table 8. For the gas-phase PAs, the MP2/6-31G(d) calculations yielded the best agreement with experiment, followed by the AMI and HF/6-31G(d) methods. This, along with the pK_a data, indicates that explicit treatment of electron correlation via MP2 is necessary for accurate calculation of PA values. Concerning the free energies of solvation, with the neutral species all approaches yield satisfactory agreement with experiment. The MP2/IPCM method is best, followed by the HF/IPCM, AM1/SM2, and HF/SCIPCM methods. For the protonated species the quality of the relative free energies of solvation is significantly worse than with the neutral compounds, with both the average and rms differences increasing by an order of magnitude. The HF/IPCM method is best, closely followed by the MP2/IPCM approach, with the AM1/SM2 and HF/SCIPCM methods being the worst. Thus, from the data in Table 8 it is evident that to improve calculations of pK_a differ-

ences, significant improvements in the determination of the free energies of solvation of charged species are required.

Another trend that may be explained via component analysis is the average underestimation of the pK_as by all the methods applied as seen in the average differences in Table 4. Analysis of Table 8 shows that the average differences of the relative free energies of solvation of the protonated species are all positive. Inspection of Eq. (3) indicates that the $\Delta\Delta G_{solv}$ of the protonated species being too positive leads to a decrease in the ΔpK_a, leading to the average differences in Table 4 being negative. Thus, limitation in the free energies of solvation of the protonated species leads to systematic variation in the calculated pK_as.

Of note based on the present study, are the reasonably high quality of the AM1/SM2 method and the poor quality of the HF/SCIPCM method for the calculation of pK_as (Table 4). The quality of the AM1/SM2 approach appears to be due to the reasonable quality of the method in reproducing experiment for the gas-phase and solvation data (Table 8). The quality of the gas-phase result appears to be due to the parameterized component of the semiempirical method, while the quality of the solvation data is attributable to the form of the SM2 model, including the large set of compounds used in the parameterization (150 neutral molecules and 28 ions) [30, 58]. It has previously been suggested that ab initio data be used for gas-phase energetics, with the SM2 model used for treatment of the free energies of solvation [59]. In the context of this suggestion it is interesting that the AM1/SM2 pK_as are in better agreement with experiment than the MP2/SM2 values (Table 2). This appears to be associated with the class and limited number of compounds used

Table 6. Relative free energies (kcal/mol) of solvation versus pyridine for the SM2, IPCM, and SCIPCM

Compound	Exp.[a]	HF/I	MP2/1	HF/S	AM1/SM2
Neutral					
2-C_2H_5	0.4	0.6	0.5	0.5	1.8
2-CH_3	0.1	0.1	−0.1	0.4	1.0
2-CHO	−1.5	−1.6	−0.7	−2.7	−2.4
2-NH_2		−2.5	−2.5	−2.9	−4.4
2-OH		−9.8	−4.6	−4.2	−6.0
3-CH_3	−0.1	−0.2	−0.4	0.1	0.3
3-COO^-		−67.1	−62.2	−51.5	−67.6
3-Cl	0.7	0.3	0.3	−0.1	0.2
4-CH_3	−0.2	−0.2	−0.2	−0.1	−0.2
4-NH_2		−1.3	−1.4	−4.8	−5.7
4-OH		−12.0	−5.8	−5.0	−3.8
2-3-$DiCH_3$	−0.1	−0.5	−0.1	−1.9	1.2
2-4-$DiCH_3$	−0.2	−0.9	−0.5	1.9	1.2
3-5-$DiCH_3$	−0.2	−1.2	−0.9	1.6	0.7
3-Br		0.1	0.6	1.6	−1.3
Protonated					
2-C_2H_5	3.9	5.5	5.3	2.9	8.0
2-CH_3	2.3	3.8	3.6	2.2	6.3
2-CHO		−2.8	−0.9	−3.3	0.5
2-NH_2		0.9	0.8	−0.8	−0.4
2-OH		−3.8	−3.7	−3.2	−2.1
3-CH_3	1.7	2.8	3.0	1.8	2.5
3-COO^-		12.9	21.9	29.0	14.5
3-Cl	−0.8	−1.0	−0.1	−3.6	−1.9
4-CH_3	1.9	2.5	2.3	2.2	2.5
4-NH_2		2.7	3.2	1.1	−0.6
4-OH		0.6	1.0	−3.7	−4.7
2-3-$DiCH_3$	3.8	5.1	5.8	15.9	8.5
2-4-$DiCH_3$	4.6	5.1	5.4	16.4	8.7
3-5-$DiCH_3$	3.8	4.7	5.8	15.1	5.0
3-Br		−3.4	−1.0	9.6	−1.6

[a] Aue et al. [37]

Table 7. Differences between the neutral and protonated relative free energies (kcal/mol) of solvation. The difference ($\Delta\Delta\Delta G_{sol}$) = neutral $\Delta\Delta G_{sol}$−protonated $\Delta\Delta G_{sol}$ is based on the values in Table 6

Compound	Exp.[a]	HF/I	MP2/I	HF/S	AM1/SM2
2-C_2H_5	−3.5	−4.9	−4.8	−2.5	−6.2
2-CH_3	−2.2	−3.8	−3.7	−1.9	−5.3
2-CHO		1.2	0.2	0.5	−2.9
2-NH_2		−3.4	−3.2	−2.2	−4.0
2-OH		−6.0	−2.5	−0.9	−4.3
3-CH_3	−1.8	−3.1	−3.4	−1.8	−2.2
3-COO^-		−80.0	−84.0	−80.5	−82.1
3-Cl	1.5	1.2	0.4	3.4	1.7
4-CH_3	−2.1	−2.7	−2.5	−2.3	−2.3
4-NH_2		−4.1	−4.6	−5.9	−5.2
4-OH		−12.6	−6.8	−1.3	0.9
2-3-$DiCH_3$	−2.1	−5.5	−5.9	−14.0	−7.4
2-4-$DiCH_3$	3.9	−6.0	−5.9	−14.5	−7.4
3-5-$DiCH_3$	4.8	−5.8	−6.6	−13.5	−4.4
3-Br	4.0	3.6	−1.7	−8.0	0.3

[a] Aue et al. [37]

Table 8. Average and rms differences of the relative calculated and experimental gas-phase PAs and free energies (kcal/mol) of solvation of the neutral and protonated species with respect to pyridine. The original data for the gas-phase PA are in Table 5, while those for the solvation energies are in Table 6

	HF/6-31G(d)	MP2/6-31G(d)	AM1	
Gas-phase PA				
Average difference	0.15	−0.07	0.13	
Rms difference	1.78	1.10	1.55	
	HF/I	MP2/I	HF/S	AM1/SM2
ΔG_{aq}, neutral				
Average difference	0.21	0.17	1.54	1.02
Rms difference	0.46	0.41	1.24	1.01
ΔG_{aq}, protonated				
Average difference	1.14	1.83	52.8	9.42
Rms difference	1.07	1.35	7.27	3.07

in the present study. With the HF/SCIPCM method the poor quality of the calculated pK_as appears to be due SCIPCM significantly overestimating the solvation energies of the charged species (Table 8). It has been noted in several cases that SCIPCM tends to overestimate the solvation energies [60, 61], which may be attributable to the SCIPCM method being developed using a rather limited set of compounds [34, 62, 63]. Furthermore, the approach used in the SCIPCM to obtain the cavity size from a charge distribution isosurface has been suggested to be unlikely to give improved accuracy [64]. Detailed analysis of Table 6 shows the largest errors with the HF/SCIPCM method to be associated with the three disubstituted compounds and the 3-Br analog. This indicates that improvements in the treatment of disubstituted compounds is required, while the poor treatment of the bromine is not surprising since it was not included in the training set during the parameterization of the SCIPCM.

3.6 Solvent attenuation effect

Solvent attenuation effects, which relate the influence of solvation on deprotonation energies, offer an additional means to test the present calculations. Solvent attenuation factors may be determined from the relative deprotonation free energies of pyridinium ions in the gas and aqueous phases as presented in Table 5. Inspecting the data in Table 5 indicated that the substituent effect on the acidity of pyridinium ions is reduced upon aqueous solvation. The attenuated substituent effect by the solvent is partly attributed to the solvation at the ring nitrogen of the neutral and protonated pyridines as well as at the substituents, although it has been suggested that the attenuation lies more in the hydration of the protonated pyridines [65]. Electron-donating substituents have been shown to have larger attenuation effects than electron-withdrawing ones [51]. Substantial deviations from the linear correlation between gas-phase PA and aqueous-phase acidity have been noted [37, 51, 66, 67] for strong π-electron donors (e.g. NH_2 and OCH_3) or electron-withdrawing groups (e.g. CHO and NO_2) due to specific solute–solvent interactions (e.g. hydrogen boding) involving the substituents. Studies have shown that 2-substituted pyridines have smaller solvent attenuation factors than 3- and 4-substituted ones [37, 48, 55] due to limited solvent accessibility of the ring nitrogen and stronger polarization in the gas phase due to the position of the 2-substituents relative to the 3- and 4-substituents [68, 69]. The experimentally observed solvent attenuation factor, calculated as the slope of the gas-phase versus the aqueous-phase deprotonation energies, for 3- and 4-substituted pyridines is approximately 2.5 [37], while that for 2-substituted analogs is 1.7 [37]. In the present study, solvent attenuation factors of 3.1 and 4.0 were determined for 3- and 4-substituted and 2-substituted species, respectively, at the MP2/IPCM level. Rigorous comparison between calculated and experimental values, however, is limited by there only being five compounds common to both studies and the omission of 2-NH_2, 4-NH_2, and 4-CHO substituted pyridines from determination of the solvent attenuation factor in the experimental study.

3.7 Influence of conformation on calculated pK_as

For several of the substituted pyridines rotation of the substituent yields multiple minima, which may influence the calculated pK_a. This problem may be especially acute with 2-CHO, where direct interactions of the formyl and the pyridine nitrogen can occur and which will be highly dependent on the protonation state of the ring. The potential-energy surfaces associated with the 2-CHO and 2-OH substituents are presented in Fig. 3; both protonation states are presented for both compounds. Only the protonated form is shown for the 2-OH analog in the aqueous phase (Fig. 3C) due to the pyridone being the dominate species for the neutral form in solution. Energy differences for the 2-CHO substituent are 1.0 and 1.8 kcal/mol for the gas and solvated neutral species, respectively, and 1.0 and 3.5 kcal/mol for the gas and solvated protonated species, respectively. For the 2-OH substituent the gas-phase energy difference between two minima is approximately 4.7 kcal/mol for the neutral state, and 2.0 and 2.5 kcal/mol for the gas

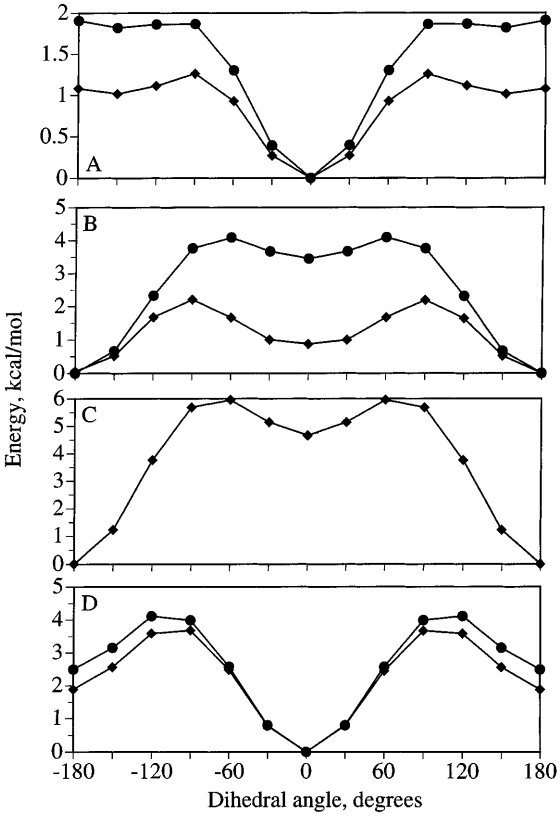


Fig. 3. Gas-phase potential-energy (●) and solvation free-energy (■) surfaces in kcalorie per mole as a function of substituent dihedral angle for **A** the neutral and **B** the protonated forms of the 2-CHO substituted compound and of **C** the neutral form and **D** the protonated forms of the 2-OH substituted compound. The 2-CHO dihedral angle was defined as C3-C2-C=O and the 2-OH dihedral angle was defined as C3-C2-O-H

and solvated protonated species, respectively. In both compounds the location of the global minimum is switched upon going from the neutral to the protonated species. Given the limitations of the level of theory used in the calculations, energy differences between some of the minima being less than 2 *KT*, and the switch in the location of the minima upon protonation it is probable that alternate conformations, other than the global minima used in the present calculations, could contribute to the pK_as.

To test the influence of conformation on the calculated pK_as all possible pK_a values were determined for the 2-OH and 2-CHO substituted compounds based on all minimum-energy structures using the AM1/SM2 level of theory. Since three calculated values are required for each pK_a determination and two conformers exist for each species yielding four possible values for each term on the right-hand side of Eq. (3), 64 combinations exist for the 2-CHO analog. With the 2-OH analog, there are only 32 combinations to be considered due to the use of the pyridone for the neutral species in solution. Histograms of the 64 and 32 pK_as calculated for the 2-CHO and 2-OH substituted analogs, respectively are presented in Fig. 4A and B. Clearly, the variations in the pK_as are quite large. In the 2-CHO compound the variation is from −4 to 6 and in the 2-OH analog it is from −5 to 7. Note that the ranges of calculated pK_as for the 2-CHO and 2-OH analogs encompass the experimental values (see Table 4). Thus, the calculated pK_a is sensitive to the conformation of the substituent. While this effect may be assumed to be magnified for the 2-position substituents due to their being ortho to the site of protonation, it is clear that alternate conformations accessible at room temperature can significantly influence the pK_a of the molecule. The importance of taking into account alternate minima in pK_a calculations has previously been noted in studies using free-energy perturbation methods [17]. It is expected that accounting for these contributions will improve the calculation of pK_a differences via QM-based methods.

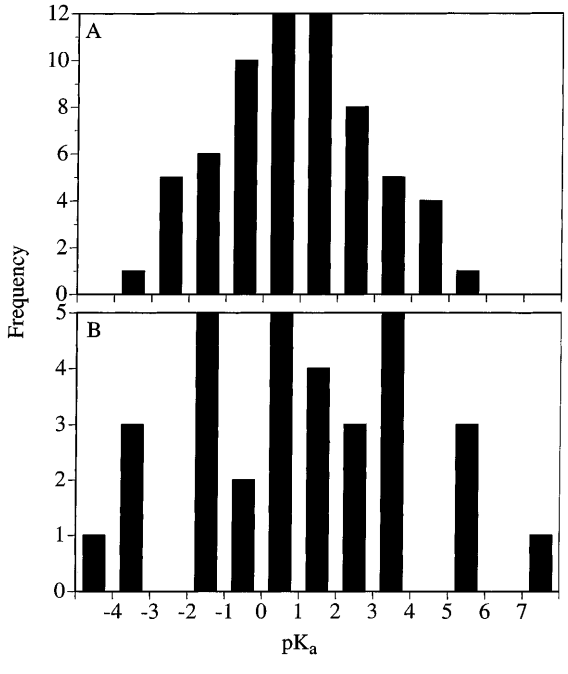


Fig. 4. Histograms of all possible pK_a values associated with all possible permutations of energy minima for **A** the 2-CHO and **B** the 2-OH substituted compounds at the AM1/SM2 level of theory

4 Conclusions

Calculations to determine changes in the pK_a of pyridines associated with substitution have been presented. Three methods were used for the determination of gas-phase deprotonation energies: the AM1 semiempirical and the HF/6-31G(d) and MP2/6-31G(d) ab initio approaches. Three different solvation models were applied, including the AM1/SM2 model, the IPCM and the SCIPCM. These models vary widely in computational requirements as well as assumptions made; however, all are implemented in available quantum chemistry programs allowing them to be readily applied. Emphasis was placed on using computationally accessible methods as required to study a collection of compounds, some of which include conformational flexibility.

Of the methods tested the MP2/IPCM approach yielded the best agreement with experiment, HF/SCIPCM was significantly worse than the other approaches, with the remaining approaches being of similar quality (Table 8). Accordingly, the MP2/IPCM method is suggested for use in determination of pK_a changes in a series of compounds. This approach is computationally feasible due to the use of the HF/6-31G(d) level of theory for the geometry optimizations and the semiempirical AM1 and AM1/SM2 to identify the location of global energy minima via energy surfaces. Alternatively, to minimize computational requirements, the AM1/SM2 method is suggested as it yields results of a quality similar to more computationally expensive ab initio methods, although the level of agreement obtained may be fortuitous. The present results, however, indicate that further improvements in methods for the computation of pK_a differences are required. It should be noted that the present conclusions are based on a relatively small number of compounds of the same structural class and should be interpreted in this context.

Analysis of the component contributions allowed the strengths and weaknesses of the present methods in determining pK_a differences to be better understood. This involved comparing the calculated relative gas-phase PAs and free energies of solvation with experimental data (Table 8). The quality of the MP2/IPCM approach is due to this method being in best or near-best agreement with experiment for the three relative energies required for the pK_a difference calculations (Eq. 3). The explicit treatment of electron correlation via MP2 clearly improves the gas-phase protonation energies over the HF and AM1 calculations. MP2 versus HF does not yield significant improvement in the ab initio based free energies of solvation, although these methods are better than AM1/SM2 for the present series of compounds. The AM1/SM2 method,

however, yields pK_a differences which are better than the MP2/SM2 results; considering the better quality of the MP2 gas-phase PAs, it appears that the improved AM1/SM2 results are due to the relatively small number of compounds studied. Comparison of all the approaches for the free energy of solvation calculations shows relative free energies to be of poorer quality for the protonated versus the neutral species. This is especially evident with the HF/SCIPCM model, although the limitation of the method is dominated by poor treatment of the disubstituted and 3-Br analogs included in the present work.

Of the three relative energies that contribute to the pK_a difference calculations (Eq. 3), the gas-phase protonation energies can currently be improved via the use of larger basis sets and more rigorous treatment of electron correlation in the ab initio calculations [70]. The use of larger basis sets has been shown to improve calculated pK_a differences for a series of benzoic acids [14], however, other studies using density functional methods with extended basis sets [71] did not yield significant improvement over MP2/6-31G(d). Concerning the free energy of solvation calculations, it is clear that improvements in the QM reaction-field models are required, especially for the protonated species. Alternatively, different methods for calculation of solvation energies, as presented in the Introduction, could be employed.

The majority of molecules of biochemical and pharmacological interest include rotatable bonds, making it necessary to take into account all accessible minima when determining pK_a values. In the present study the global minima at the AM1 and AM1/SM2 levels were identified and were used for the pK_a computations. To obtain a better idea of how the use of alternate conformations could influence the calculated pK_a difference, all possible pK_a values were determined for all permutations of the minimum-energy structures at the AM1/SM2 level for the 2-OH and 2-CHO substituted compounds. As shown in Fig. 4, the variation in pK_a values is quite large for both compounds. Clearly, such variability associated with alternate conformation is expected to contribute to the pK_a in the experimental regime and, accordingly, must be taken into account in pK_a calculations.

Acknowledgements. This work was supported financially by a grant from the Procter & Gamble International Program for Animal Alternatives. Computational support was supplied by NCI's Frederick Biomedical Supercomputing Center. Appreciation is also extended to the referee of this manuscript for the high-quality and insightful comments.

References

1. Fruton JS (1976) Adv Enzymol Relat Areas Mol Biol 44: 1–36
2. Mackenzie NE, Malthouse JP, Scott AI (1984) Science 225: 883–889
3. Lin XL, Lin YZ, Tang J (1994) Methods Enzymol 241: 195–224
4. Lubkowski J, Yang F, Alexandratos J, Merkel G, Katz RA, Gravuer K, Skalka AM, Wlodawer A (1998) J Biol Chem 273: 32685–32689
5. Lubkowski J, Yang F, Alexandratos J, Wlodawer A, Zhao H, Burke TR, Neamati N, Pommier Y, Merkel G, Skalka AM (1998) Proc Natl Aacd Sci USA 95: 4831–4836
6. Lombardo F, Blake JF, Curatolo WJ (1996) J Med Chem 39: 4750–4755
7. Palm K, Luthman K, Ungell A-L, Strandlund G, Artursson P (1996) J Pharm Sci 85: 32–39
8. MacKerell AD Jr, Chen I-J, Taneja R, Yin D, Young D, Polli J Manuscript in Preparation (1999)
9. Jorgensen WL, Briggs JM, Gao J (1987) J Am Chem Soc 109: 6857–6858
10. Bashford D, Karplus M (1991) J Phys Chem 95: 9556–9561
11. Yang A-S, Honig B (1993) J Mol Biol 231: 459–474
12. Demchuk E, Wade RC (1996) J Phys Chem 100: 17373–17387
13. Schaefer M, Sommer M, Karplus M (1997) J Phys Chem B 101: 1663–1683
14. Shapley WA, Bacskay GB, Warr GG (1998) J Phys Chem B 102: 1938–1944
15. Tanaka Y, Shiratori Y, Nakagawa S (1990) Chem Phys Lett 169: 513–518
16. Balbuena PB, Johnston KP, Rossky PJ (1996) J Phys Chem B 100: 2716–2722
17. Figueirido F, Del Buono GS, Levy RM (1996) J Phys Chem 100: 6389–6392
18. Gao J, Li N, Freindorf M (1996) J Am Chem Soc 118: 4912–4913
19. Potter MJ, Gilson MK, McCammon JA (1994) J Am Chem Soc 116: 10298–10299
20. Honig B, Nicholls A (1995) Science 268: 1144–1149
21. Qiu D, Shenkin PS, Hollinger FP, Still WC (1997) J Phys Chem A 101: 3005–3014
22. Sakurai M, Furuki T, Inoue Y (1995) J Phys Chem 99: 17789–17794
23. Davis ME, McCammon JA (1990) Chem Rev 90: 509–521
24. Sham YY, Chu ZT, Warshel A (1997) J Phys Chem B 101: 4458–4472
25. Cramer CJ, Truhlar DG (1991) J Am Chem Soc 113: 8552–8554
26. Tomasi J, Persico M (1994) Chem Rev 94: 2027–2094
27. Frisch MJ, Trucks GW, Schiegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Cioslowski J, Stefanov BB, Nanayakkara A, Challocombe M, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox DJ, Binkley JS, Defrees DJ, Baker J, Stewart JJP, Head–Gordon M, Gonzalez C, Pople JA (1996) Gaussian 94, revision C.3. Gaussian, Pittsburgh, Pa
28. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 3902–3909
29. Hawkins GD, Lynch GC, Giesen DJ, Rossi J, Storer JW, Liotard DA, Cramer CJ, Truhlar DG (1996) QCPE Bull 16: 11
30. Cramer CJ, Truhlar DG (1992) J Comput-Aid Mol Des 6: 629–666
31. Hawkins GD, Cramer CJ, Truhlar DG (1996) J Phys Chem 100: 19824–19839
32. Foresman JB, Keith TA, Wiberg KB, Snoonian J, Frisch MJ (1996) J Phys Chem 100: 16098–16104
33. Miertus S, Scrocco E, Tomasi J (1981) Chem Phys 55: 117–129
34. Cossi M, Barone V, Camni R, Tomasi L (1996) Chem Phys Lett 255: 327–335
35. Weast RC (ed) (1988) Handbook of chemistry and physics. 57th edition CRC Press, Boca Raton
36. McQuarrie DA (1976) Statistical mechanics. Harper & Row, New York
37. Aue DH, Webb HM, Davidson WR, Toure P, Hopkins HP Jr, Moulik SP, Jahagirdar DV (1991) J Am Chem Soc 113: 1770–1780
38. Wong MW, Frisch ML, Wiberg KB (1991) J Am Chem Soc 113: 4776–4782
39. Wong MW, Wiberg KB, Frisch MJ (1992) J Am Chem Soc 114: 523
40. Wong MW, Wiberg KB, Frisch MJ (1992) J Am Chem Soc 114: 1645–1652
41. Wong MW, Wiberg KB, Frisch MJ (1991) J Chem Phys 95: 8991

42. Beak P, Fry FS Jr, Lee J, Steele F (1976) J Am Chem Soc 98: 171–179
43. Scanlan MJ, Hillier IH, MacDowell AA (1983) J Am Chem Soc 105: 3568–3571
44. Kwizikowski JS, Bartlett RJ, Person WB (1988) J Am Chem Soc 110: 2353–2358
45. Karelson MM, Katritzky AR, Szafran M, Zerner MC (1989) J Org Chem 54: 6030–6034
46. Cieplak P, Bash P, Singh C, Kollman PA (1987) J Am Chem Soc 109: 6283–6289
47. Liotta CL, Perdue EM, Hopkins HP Jr (1974) J Am Chem Soc 96: 7308–7311
48. Valentini JJ, Coggiola MJ, Lee YT (1976) J Am Chem Soc 98: 854–856
49. Arnett EM, Chawla B, Bell L, Taagepera M, Hehre WJ, Taft RW (1977) J Am Chem Soc 99: 5729–5738
50. Davidson WR, Sunner J, Kebarle P (1979) J Am Chem Soc 101: 1675–1680
51. Taagepera M, Summerhays KD, Hehre WL, Topsom RD, Pross A, Radom L, Taft RW (1981) J Org Chem 46: 891–903
52. Morrison RT (1998) Organic chemistry, 7th edn. Prentice Hall, New York
53. Wade LG (1998) Organic chemistry, 4th edn. Prentice Hall, New York
54. Taft RW (1983) Prog Phys Org Chem 14: 248–350
55. Abboud JM, Catalan J, Elguero J, Taft RW (1988) J Org Chem 53: 1137–1140
56. Cabani S, Gianni P, Matteoli E (1972) J Phys Chem 76: 2959–2966
57. Nakamoto K, Martell AE (1959) J Am Chem Soc 81: 5857–5863
58. Cramer CJ, Truhlar DG (1992) Science 256: 213–217
59. Cramer CJ, Truhlar DG (1993) J Am Chem Soc 115: 8810–8817
60. Truong TN, Stefanovich EV (1995) J Phys Chem 99: 14700–14706
61. Chen J, McAllister MA, Lee JK, Houk KN (1998) J Org Chem 63: 4611–4619
62. Wiberg KB, Keith TA, Frisch MJ, Murcko M (1995) J Phys Chem 99: 9072–9079
63. Wiberg KB, Castejon H, Keith TA (1996) J Comput Chem 17: 185–190
64. Marten B, Kim K, Cortis C, Friesner RA, Murphy RB, Ringnalda MN, Sitkoff D, Hong B (1996) J Phys Chem 100: 11775–11788
65. Arnett EM, Chawla B (1979) J Am Chem Soc 101: 7141–7146
66. Fuijo M, McIver RT Jr, Taft RW (1981) J Am Chem Soc 103: 4017–4029
67. Mishima M, McIver RR Jr, Taft RW, Bordwell FG, Olmstead WN (1984) J Am Chem Soc 106: 2717–2718
68. King EJ (1965) Acid-base equilibria, 1st edn. Pergamon, New York
69. Gal J-F, Maria P-C (1990) Prog Phys Org Chem 21: 159–238
70. Bliznyuk AA, Schaefer HF III, Amster IJ (1993) J Am Chem Soc 115: 5149–5154
71. Kallies B, Mitzner R (1997) J Phys Chem B 101: 2959–2967
72. Jaffe HH, Doak GO (1955) J Am Chem Soc 77: 4441–4444