Gas Phase Proton Affinity, Basicity, and pK_a Values for Nitrogen Containing Heterocyclic Aromatic Compounds

Sungu Hwang,[†,*] **Yun Hee Jang,**[‡,a] **and Doo Soo Chung**[‡]

[†]*School of Free Major, Miryang National University, Gyeongnam 627-706, Korea.* [*]*E-mail: sungu@mnu.ac.kr*
[‡]*School of Chemistry, NS60, Seoul National University, Seoul 151-747, Korea*

Received December 24, 2004

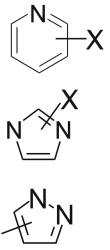
Bipyridine and its derivatives have been widely used as the ligands in transition metal complexes. The proton affinities of pyridine derivatives were calculated using an *ab initio* quantum mechanical method (B3LYP with various double zeta and triple zeta basis sets) in combination with the Poisson-Boltzmann continuum solvation model. Van der Waals radii of the atoms in the heterocyclic rings for the solvation energy calculation were set to values determined to reproduce the pK_a values of guanine and oxoguanine derivatives and that of chlorine was optimized to reproduce the experimental values of relating compounds. The pK_a values for the heterocyclic ring compounds were in agreement with the experimental values with a mean unsigned error of 0.45 pK_a units.

Key Words : p*K*a, DFT, Poisson-Boltzmann, Imidazole, Pyridine

Introduction

Bipyridine and its derivatives have widely been used as the ligands in transition metal complexes. Ruthenium polypyridine (bipyridine and terpyridine) complexes have received extensive attention due to their strong metal to ligand charge-transfer (MLCT) transitions, facile electron-transfer properties and long lived MLCT excited states, which makes them attractive for designing photo- and electrochemical devices.[1,2] Chiral ruthenium-containing terpyridyl-based compounds were synthesized and characterized due to their potential applications as nonlinear optical materials and as electrocatalysts for enantioselective transformations.[3] Iridium (III) bipyridine complexes were synthesized as multicolor luminescent cross-linkers for biological molecules.[4] Derivatives of polypyridine complexes of platinum (II) are receiving a great deal of attention for their photoliminescence.[5] In order to design a better ligand for the optical application it is required to obtain basic physical and chemical properties of the ligand and find their relations to the optical properties. One of the properties defining the characteristics of the ligand would be pK_a. This value can be regarded as one of the solution phase properties essential for the novel photosensitizer design.[6]

The heterocyclic molecules have lately attracted attention due to their "shifted pK_a values" upon complexation to metal ions because it can rationalize the existence of nucleobases of differing protonation states at physiological pH.[7] Polypyridine ligands also find its use in the nuclear fuels reprocessing by forming actinide/lanthanide complexes.[8] Studies on protonation of polypyridine molecules are also of great interest in pharmacological and biological research since the site of protonation is closely related to expression of biological activity.[9]



X = H (**1**), 2-Cl (**6**), 3-Cl (**7**), 4-Cl (**8**)
2-CH_3 (**10**), 3-CH_3 (**12**), 4-CH_3 (**15**)

X = H (**2**), 1-CH_3 (**14**), 2-CH_3 (**4**), 4-CH_3 (**5**)

X = H (**3**), 1-CH_3 (**16**), 3-CH_3 (**9**),
4-CH_3 (**11**), 4-Cl (**13**)

Scheme 1. pyridine, imidazole, and pyrazole derivatives.

In order to establish a method for calculating proton affinities in the gas phase and aqueous solution, we performed density functional theoretical study on 16 molecules shown in Scheme 1. This is also to check the vdw parameters of atoms used for the solvation energy calculation for the heterocyclic ring that was developed already and to optimize the value for chlorine attached to the ring.

Computational Details

The details of the pK_a calculation have been described elsewhere.[10] Here we summerize those features of the calculation relevant to the discussion of the data below.

The pK_a value of an acid HA in aqueous solution is related to the Gibbs free energy change for the deprotonation process expressed by[10]

$$\mathrm{p}K_a = \frac{1}{2.303RT}\Delta G^0_{deprot,aq} \qquad (1)$$

where R is the gas constant and T temperature in Kelvin. The standard free energy of deprotonation from one of the deprotonation sites of HA in water, $\Delta G^0_{deprot,aq}$, is defined as (See Scheme 2).

$$\Delta G^0_{deprot,aq} = \Delta G^0_{aq}(\mathrm{A}^-) + \Delta G^0_{aq}(\mathrm{H}^+) - \Delta G^0_{aq}(\mathrm{HA})\,. \qquad (2)$$

---

[a] Present address: Materials and Process Simulation Center, BI 139-74, California Institute of Technology, Pasadena 91125, CA, USA

$$\begin{tikzcd}
\mathrm{HA(g)} \arrow[r, "\Delta G^0_{\mathrm{deprot,\,g}}"] \arrow[d, "\Delta G^0_{\mathrm{solv}}\,(\mathrm{HA})"'] & \mathrm{A^-(g)} \arrow[d, "\Delta G^0_{\mathrm{solv}}\,(\mathrm{A^-})"'] & + & \mathrm{H^+(g)} \arrow[d, "\Delta G^0_{\mathrm{solv}}\,(\mathrm{H^+})"] \\
\mathrm{HA(aq)} \arrow[r, "\Delta G^0_{\mathrm{deprot,\,aq}}"'] & \mathrm{A^-(aq)} & + & \mathrm{H^+(aq)}
\end{tikzcd}$$

Scheme 2. Thermodynamic cycle used in the calculation of a pK_a value.

The standard free energy of each species (HA, A^-, and H^+) in water, ΔG^0_{aq}, can be written by the sum of the gas-phase standard free energy ΔG^0_g and the standard free energy of solvation in water ΔG^0_{solv}:

$$\Delta G^0_{aq} = \Delta G^0_g + \Delta G^0_{solv}. \tag{3}$$

The standard free energy of each species in the gas phase, ΔG^0_g, is obtained by

$$\Delta G^0_g = E_{0K} + \mathrm{ZPE} + \Delta\Delta G_{0\rightarrow 298K}. \tag{4}$$

The total energy of the molecule at 0 K (E_{0K}) is calculated at the optimum geometry from quantum mechanics (QM). Harmonic oscillator-rigid rotor approximation is applied for the calculation: The zero-point energy (ZPE) and the vibrational contribution to Gibbs free energy change from 0 K to 298 K ($\Delta\Delta G_{0\rightarrow 298K}$) are calculated from frequencies calculated using QM. The translational and rotational free energy contribution is also calculated in the ideal gas approximation. We used $\Delta G^0_g(H^+) = 2.5\ RT - T\Delta S^0 = 1.48\text{-}7.76 = -6.28$ kcal/mol from the literature.[11]

All QM calculations used the Jaguar v4.1 quantum chemistry software. To calculate the geometries and energies of the various molecules, we used the B3LYP flavor of density functional theory (DFT), which includes the generalized gradient approximation and a component of the exact Hartree-Fock (HF) exchange. Since calculations of vibration frequencies are generally quite time-consuming, the small basis set of 6-31G** basis set was first used to optimize the geometry and calculate the vibration frequencies. Recently, it is reported that B3LYP/6-31G** level of calculation is adequate to describe the conformational properties of bipyridines.[12] Then the other basis sets (6-31++G**, cc-pVDZ, aug-cc-pVDZ, cc-pVTZ(-f), and aug-cc-pVTZ) were used for the final geometry optimization started from the 6-31G** optimized geometry:

$$\Delta G^0_g = \mathrm{ZPE}^{\text{6-31G**}} + \Delta\Delta G^{\text{6-31G**}}_{0\rightarrow 298K} + E^{\text{basis set}}_{0K,g}. \tag{5}$$

Gas-phase proton affinity (PA) and gas-phase basicity (GB) at 298 K, which are defined as the enthalpy change and the free energy change, respectively, during the protonation process in gas-phase are calculated.

$$\mathrm{PA} = \Delta H^0_g(\mathrm{A^-}) + \Delta H^0_g(\mathrm{H^+}) - \Delta H^0_g(\mathrm{HA}), \tag{6}$$

Table 1. van der Waals (vdw) radii used for the solvation energy calculations

atom		vdw radii (Å)	scale factor[a]	references
N	(sp^2)	1.41	0.94	10,16
C	(sp^2)	1.88	0.94	10,16
	(sp^3)	1.90	1.00	14
H	(attached to sp^2 N)	1.08	0.94	10,16
	(attached to sp^2 C)	1.18	0.94	10,16
	(attached to sp^3 C)	1.15	1.00	14
Cl		1.66	0.84	present work

[a]Scaled from the valued reported in reference[14]

$$\mathrm{GB} = \Delta G^0_g(\mathrm{A^-}) + \Delta G^0_g(\mathrm{H^+}) - \Delta G^0_g(\mathrm{HA}), \tag{7}$$

where $\Delta H^0_g(H^+) = 2.5\ RT = 1.48$ kcal/mol and $\Delta G^0_g(H^+) = 2.5\ RT - T\Delta S^0 = 1.48 - 7.76 = -6.28$ kcal/mol at 298 K.

Details for the solvation energy calculation are given elsewhere,[10] and so we briefly described the overall procedure. Continuum solvent model was applied to the calculations. The solvation Gibbs free energy was given by the sum of the non-electrostatic contribution due to the creation of the solute cavity in the solvent and the electrostatic interaction between solute and solvent. The electrostatic part of the solvation free energy was evaluated by a self-consistent formalism[13,14] that cycles through quantum mechanical calculations in the solvent reaction field that involves numerical solution of the Poisson-Boltzmann (PB) equation.[15]

Non-electrostatic contribution such as creation of a cavity in the solvent by the solute are taken into consideration by a term proportional to the solvent-accessible surface (SAS) area of the solute defined by the surface traced out by the center of the sphere of probe radius (1.4 Å for water) as it is rolled around the solute. This term is based on the experimental solvation energy of linear and branched alkanes.[13,14]

Vdw radii of the atoms in the heterocyclic rings for the solvation energy calculation were set to according to guanine and oxyguanine.[10,16] For sp^3 carbon and hydrogen attached to it, vdw radii were taken from the original parameters of Marten. Vdw radius of chlorine was optimized to reproduce the pK_a of chlorine derivatives of ring compounds **6**, **7**, **8** and **13**.

Results and Discussion

Gas phase proton affinity (PA). The accuracy of pK_a calculations depends on two factors: one is the gas phase basicity and the other is the solvation energy. The former can be checked against the gas phase experiment. Therefore we checked the accuracy of the gas phase proton affinity (PA). In Figure 1, we depicted PA, the enthalpy change corresponding to the protonation reaction for the 16 molecules shown in Scheme 2. Regression Coefficient (R) is similar with each other. However, the calculated absolute values from 6-31G** are about 6.6 kcal/mol too large as judged

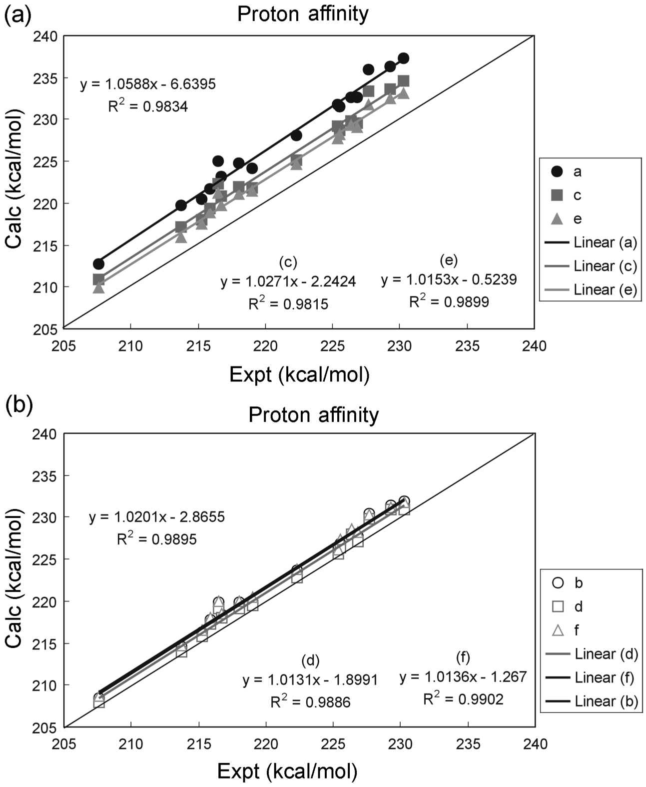


Figure 1. Gas phase proton affinities (PA) of heterocyclic molecules calculated with a basis set of a) (a) 6-31G**, (c) pp-PVDZ, (e) cc-PVTZ(-f) and b) (b) 6-31++G**, (d) aug-cc-PVDZ, and (f) aug-cc-PVTZ(-f).

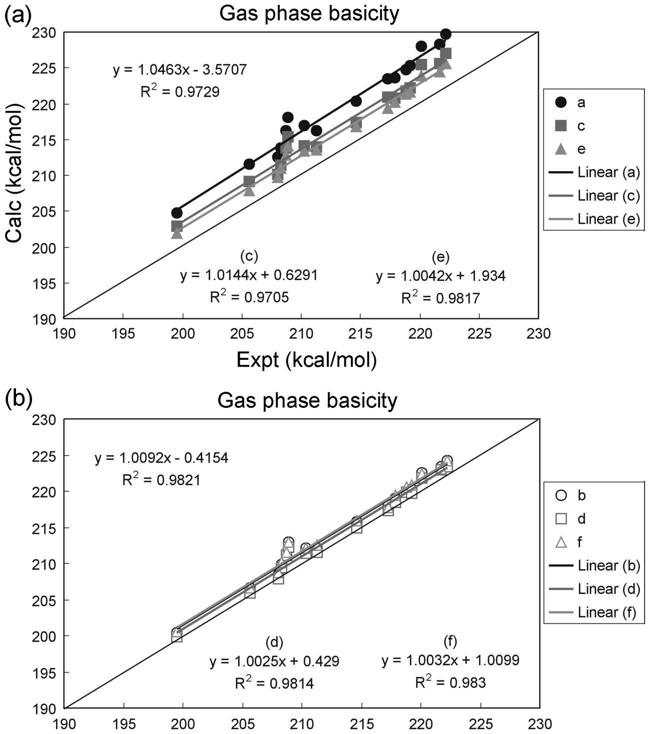


Figure 2. Gas phase basicity (GB) of heterocyclic molecules calculated with a basis set of a) (a) 6-31G**, (c) pp-PVDZ, (e) cc-PVTZ(-f) and b) (b) 6-31++G**, (d) aug-cc-PVDZ, and (f) aug-cc-PVTZ(-f).

from the *y*-intercept. This discrepancy can be reduced either by using different basis set (c, or e) or by including diffuse function (b, d or f). The inclusion of the diffuse function (See (b)) has greatly affected the precision of the theoretical calculations as reported by others.[17] This is due to the poorer description of a neutral base in comparison with its conjugated cationic acid.[17] However the calculated relative PA values are sufficient as shown their R^2 values. Judging from the slope and the *y*-intercept together, inclusion of the diffuse function is better treatment for the direct reproduction of PA.

Gas phase basicity (GB). In Figure 2, we depicted gas phase basicity, the Gibbs free energy change corresponding to the protonation reaction for the 16 molecules shown in scheme 2. Similar argument can be applied to GB as in the case of PA.

pK_a. As discussed earlier, the accuracy of pK_a calculations depends on both the gas phase basicity and the solvation energy. The former factor was checked against the gas phase experiment in the previous 2 sections. Here we check the latter part.

Shown in Figure 3 are pK_a values for the 16 molecules shown in Scheme 2. The inclusion of the diffuse function (See b)) has greatly affected the precision of the theoretical calculations as reported by others. Of three basis set, we choose aug-cc-PVDZ (= cc-PVDZ++) for the production

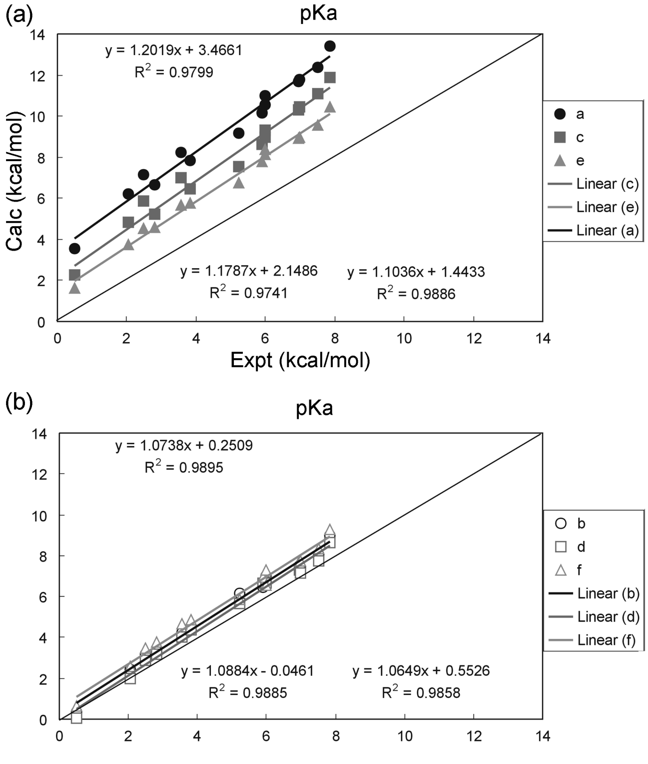


Figure 3. pK_a values of heterocyclic molecules calculated with a basis set of a) (a) 6-31G**, (c) pp-PVDZ, (e) cc-PVTZ(-f) and b) (b) 6-31++G**, (d) aug-cc-PVDZ, and (f) aug-cc-PVTZ(-f).

calculation due to its relative small computation cost and smallest deviation. Moreover, the correlation consistent basis sets are one of the most widely used basis sets due to their convergent behavior upon increasing basis set size.[18]

For diprotic acids having two protons available, we use the following dissociative equilibria.[19,20]

$$K_{a1} = K_{a,A} + K_{a,B} \quad (8)$$

$$pK_{a1} = -\log(K_{a,A} + K_{a,B}) = -\log(10^{-pK_{a,A}} + 10^{-pK_{a,B}}) \quad (9)$$

where $K_{a,A}$ and $K_{a,B}$ is the site-specific, microscopic pK_a value for proton site A and B, respectively, and K_{a1} is the apparent, overall pK_a value.

Summary

We performed B3LYP flavor of DFT calculations with six different basis sets on the nitrogen-containing heterocyclic aromatic compounds for the pK_a calculations. The inclusion of the diffuse function would greatly improve the quality of the numerical results. The cc-pVDZ++ basis set would be of choice for the accuracy and the size of the resources required. The calculated values are in good agreement with the experimental data within 0.4 pK_a units of mean unsigned error.

Future directions in the research to obtain further improvements in the theoretical calculation of the pK_a value is that for the nonaqueous, organic solvent and mixed solvent.[21] We are also currently working on this line of research.[22-24]

Acknowledgement. This work was supported by Korea Research Foundation (KRF-2002-003-C00073). Parts of the calculations in this work were performed by using the supercomputing resources of Korea Institute of Science and Technology Information.

References

1. Mondal, B.; Puranik, V. G.; Lahiri, G. K. *Inorg. Chem.* **2002**, *41*, 5831.
2. Chanda, N.; Mondal, B.; Puranik, V. G.; Lahiri, G. K. *Polyhedron* **2002**, *21*, 2033.
3. Barron, J. A.; Glazier, S.; Bernhard, S.; Takada, K.; Houston, P. L.; Abruna, H. D. *Inorg. Chem.* **2003**, *42*, 1448.
4. Lo, K. K. W.; Chung, C. K.; Zhu, N. Y. *Chemistry-a European Journal* **2003**, *9*, 475.
5. McMillin, D. R.; Moore, J. J. *Coord. Chem. Rev.* **2002**, *229*, 113.
6. Nazeeruddin, M. K.; Zakeeruddin, S. M.; Humphry-Baker, R.; Kaden, T. A.; Graetzel, M. *Inorg. Chem.* **2000**, *39*, 4542.
7. Roitzsch, M. R.; Lippert, B. *J. Am. Chem. Soc.* **2004**, *126*, 2421.
8. Rabbe, C.; Mikhalko, V.; Dognon, J. P. *Theor. Chem. Acc.* **2000**, *104*, 280.
9. Matyus, P.; Fuji, K.; Tanaka, K. *Tetrahedron* **1994**, *50*, 2405.
10. Jang, Y. H.; Goddard III, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *Chem. Res. Toxicol.* **2002**, *15*, 1023.
11. Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Phys. Chem. A* **1997**, *101*, 10075.
12. Lehtonen, O.; Ikkala, O.; Pietila, L.-O. *J. Mol. Struct. (THEOCHEM)* **2003**, *663*, 91.
13. Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M. N.; Goddard III, W. A.; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875.
14. Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775.
15. Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
16. Jang, Y. H.; Goddard III, W. A.; Noyes, K. T.; Sowers, L. C.; Hwang, S.; Chung, D. S. *J. Phys. Chem. B* **2003**, *107*, 344.
17. Kallies, B.; Mitzner, R. *J. Phys. Chem. B* **1997**, *101*, 2959.
18. Sinha, P.; Boesch, S. E.; Gu, C.; Wheeler, R. A.; Wilson, A. K. *J. Phys. Chem. A* **2004**, *108*, 9213.
19. Adam, K. R. *J. Phys. Chem. A* **2002**, *106*, 11963.
20. Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pK_a Prediction for Organic Acids and Bases*; Chapman and Hall: London, 1981.
21. Almerindo, G. I.; Tondo, D. W.; Pietila, L.-O. *J. Phys. Chem. A* **2004**, *108*, 166.
22. Hwang, S.; Jang, Y. H.; Chung, D. S. *Chem. Lett.* **2001**, *30*, 1182.
23. Jang, Y. H.; Hwang, S.; Chung, D. S. *Chem. Lett.* **2001**, *30*, 1316.
24. Park, I.; Jang, Y. H.; Hwang, S.; Chung, D. S. *Chem. Lett.* **2003**, *21*, 376.