Proton Affinity of Five-Membered Heterocyclic Amines: Assessment of Computational Procedures†

J. SRINIVASA RAO, G. NARAHARI SASTRY
Molecular Modeling Group, Organic Chemical Sciences, Indian Institute of Chemical Technology, Tarnaka, Hyderabad 500 007, India

Received 29 August 2005; accepted 2 September 2005
Published online 4 November 2005 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.20862

ABSTRACT: Ab initio quantum chemical calculations, G3B3, second-order Møller–Plesset (MP2), and the hybrid density functional method B3LYP were employed to compute the proton affinities of 24 heterocyclic amines. A range of basis sets are employed, starting from double-ζ polarization quality to triple-ζ quality basis set with augmented diffuse and polarization function. Experimental values were used to calibrate the performance of various theoretical models. The regioselectivity for the protonation has been unambiguously established by performing B3LYP/6-31G* calculations on the possible putative sites of attack. For the given series of compounds the performance of B3LYP/6-31+ +G** and G3B3 levels of theory have been in excellent agreement with the experimental results with the deviations are of the order comparable with the experimental error. © 2005 Wiley Periodicals, Inc. Int J Quantum Chem 106: 1217–1224, 2006

Key words: assessment; DFT; G3B3; regioselectivity; proton affinity; heterocyclic amine

Introduction

Heterocyclic systems are of widespread occurrence in nature, particularly in such natural products as nucleic acids, plant alkaloids, vitamins, proteins, anthocyanins, and flavones, and the heme pigments and chlorophyll. Heterocyclic azoles are also known as the building blocks of many antibiotics, anticancer agents, fungicides, and several other drugs [1–3]. Acidity and basicity of molecules are fundamental concepts in chemistry and are central to understanding the chemical reactivity [4–6]. The kinetics and thermodynamics of gas-phase proton transfer reactions are crucial in analyzing the preferred site for protonation in the polyfunctional molecules, which aids in understanding the reaction mechanisms occurring in several biological systems [7–10]. Yáñez et al. [11] studied the effect of
basis set on the protonation of three-membered heterocyclic compounds. These heterocyclic bases are of particular interest for the following reasons. Prediction of correct site for protonation is challenging problem from theoretical perspective [12]. Computational quantum chemistry results designate the importance of employing the good quality basis sets especially with polarization functions on hydrogen [13–16]. In recent years, computational chemistry is playing an important role in supporting and supplementing the experimental research, and its complementarity to the gas phase results is often required to rationalize and explain the observed results.

The present study considers 24 heterocyclic amines whose proton affinity values are available experimentally (Scheme 1). The considered molecules include pyrrole and a series of substituted pyrazoles, imidazoles, triazoles, and thiazoles. The experimental values are taken from Hunter and Lias [17]. Some of the compounds studied were subjected to theoretical investigations earlier, albeit not uniformly at the quantitative levels of accuracy undertaken in the present work. In their elegant

SCHEME 1. Optimized geometries of neutral heterocyclic amines at B3LYP/6-31+G** level of theory. All values are given in Angstroms (Å). The numbers are given in the descending order of proton affinity values.
experimental and theoretical calculations, Huang and Rodgers [18] reported the proton affinities of pyrrole, pyrazole, and imidazole at the Møller–Plesset (MP2) level of theory. Schmiedekamp et al. [19] reported the density functional theory (DFT) calculations on a series of 1,2,4-triazole. Proton affinity determination has been one of the most actively pursued topics by computational chemists, and a range of computational methods has been employed over the years to model this quantity. The importance of including electron correlation in reliable prediction of the proton affinities has been unambiguously established in the past [20–22]. As accurate estimation of the proton affinity values of the chosen set of cyclic amines is of high importance, we undertook a systematic analysis to estimate the relative performance of current line of computational methods. In recent years, DFT, especially the B3LYP method, appears to be the method of choice, as compared with the conventional electron correlation method (MP2), because of its computational economy. Thus, wherever possible, it has become our practice to employ this more economical variant compared with the MP2 and CCSD(T) methodologies. Therefore, the main objective of the present study is to assess the performance of the hybrid density functional method B3LYP, as compared with the ab initio methods. Although the site of attack is not controversial in most cases, we undertook a systematic study to assess the relative proton affinities of various positions in each of the isomers. Therefore, the regioselectivity is clearly established; this study also provides an idea about the next possible site of protonation in each of the isomers considered in the study.

**COMPUTATIONAL DETAILS**

All calculations were carried out with the Gaussian 03 suite of programs [23]. All structures were optimized without imposing any symmetry constraints at the B3LYP level, using 6-31G* and 6-31+ +G** basis sets. Frequency calculations were done to ascertain the nature of stationary points on the potential energy surface (PES). Hartree–Fock (HF), MP2, and B3LYP single-point calculations were done on B3LYP/6-31+ +G** optimized geometries by employing the 6-311 + +G** basis set. Zero-point vibrational energies (ZPVE) calculated at the B3LYP/6-31+ +G** level of theory were used for all single-point calculations. In addition to these methods, we have used the G3B3 [24] method to calculate the proton affinities. G3B3 (G3//B3LYP) is a variant of G3 theory in which geometries are obtained at the B3LYP/6-31G* level instead of MP2(FU)/6-31G*; ZPVE are calculated at the B3LYP/6-31G* level and are scaled by 0.96. The G3B3 calculation is done in the following sequence: (a) geometry optimization and frequency calculation at the B3LYP/6-31G(d) level; (b) QCISD(T,FC)/6-31G(d)//B3LYP/6-31G(d) single-point energy evaluation; (c) MP4(FC)/6-31+ +G(d)//B3LYP/6-31G(d) single-point energy; (d) MP4(FC)/6-31G(2df,p)//B3LYP/6-31G(d) single point; and (e) MP2(Full)/6-311 + +G(2df)//B3LYP/6-31G(d) single point. In addition, this calculation includes some empirical spin-orbit and higher-level corrections to accurately model the atomization energies. Therefore, this procedure is an effective alternative to a QCISD(T) calculation with a large basis set and the performance of G3 series of calculations seem to be among the best computational estimates for energetics of various kind.

Proton affinities are computed by employing the following equation [25]:

\[
\text{Proton affinity} = \Delta E_{\text{ele}} + \text{ZPVE} + 5RT/2,
\]

where \( \Delta E_{\text{ele}} = [E(B)-E(BH^+)] \), ZPVE refers to the zero-point vibrational energy, and the constant 5RT/2 is the classical estimation of the effect of gaining three transitional degrees of freedom (3RT/2) for the proton plus RT, the PV term for the proton. The negative of standard enthalpy change (\( -\Delta H_{\text{298}} \)) during the following reaction is considered as the proton affinity of that particular heterocyclic amine:

\[
\text{BH}^+ \rightarrow \text{B} + \text{H}^+.
\]

**Results and Discussion**

First, we report the results of proton affinity values of all possible sites of attack of the 24 amines considered in the study. Table I presents the proton affinity values computed at the B3LYP/6-31G* level of theory. The preferential order of proton affinities for each site is given in Scheme 1. The proton binds preferably to the in-plane lone pair of the nitrogen in 23 cases; in the absence of such a possibility, as in pyrrole (1), it binds to a ring C. While the N containing the in-plane lone pair has been the site of highest proton affinity, the differences between the
ring C and the N with the \( \pi \)-type lone pair are smaller. In most cases, ring C has higher affinity compared with the ring N center. Radom and colleagues [26] rationalized the difference in the proton binding affinity of the two ring carbons in pyrrole based on the number of resonance structures of the protonated species.

Thus, the rationalization of the most preferred protonation site is well established, and the variation in the proton affinities ranges from 214 for pyrrole (1) all the way up to 239 kcal/mol for 19. Some of the observations made during the calculations of the various protonated regioisomers are mentioned below. For the compound 1,2,3-triazole (4) there are three putative sites, N1, N2, and N3; the B3LYP computations indicate that protonation at the 2 (N2) position leads to cleavage of the heterocyclic ring. The compound 1,2,4-triazole (5) also has two nitrogen atoms with localized lone pair of electrons, while protonation on N placed between the two carbon atoms is preferred by 10.7 kcal/mol. In 2-aminothiazole (24), in addition to the ring nitrogen there is one more site, in the form of a side chain -NH\(_2\) substituent, outside the ring. However, protonation at ring nitrogen is favored over side-chain nitrogen by \( \sim \)25 kcal/mol; this may be traced to the electron-rich nature of the ring N atom. Except in 1 and 24, the proton affinity of the second site in most isomers is not very competitive, as their values are lower by \( \geq \)10 kcal/mol. In a large number of cases, the differences between the first and second sites are well over 20 kcal/mol.

After unambiguously identifying the most stable regioisomer for all 24 molecules considered, higher-level calculations were carried out on the most stable protonated isomer to estimate the proton affinities. Increasing the basis set from 6-31G* to 6-31G** led to substantial reduction of the proton affinity values in most cases; agreement with the experimental values is much better (Table II). However, further increasing the quality to the 6-311G** level has only a minor effect. Taking the B3LYP/6-31++G** as the reference geometry, single-point calculations were done at the HF, B3LYP, and MP2 levels of theory using the 6-311++G** level of theory. Finally, G3B3 calculations were done on all the molecules considered. The mean absolute deviation values for each level

### Table I

Calculated proton affinities (kcal/mol) for all sites in the heterocyclic ring at B3LYP/6-31G* level of theory.

<table>
<thead>
<tr>
<th>Compound</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>214.4</td>
<td>210.3</td>
<td>192.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>217.4</td>
<td>196.9</td>
<td>188.9</td>
<td>187.8</td>
<td>183.6</td>
</tr>
<tr>
<td>3</td>
<td>229.6</td>
<td>201.7</td>
<td>198.0</td>
<td>197.7</td>
<td>178.3</td>
</tr>
<tr>
<td>4</td>
<td>216.8</td>
<td>204.5</td>
<td>187.0</td>
<td>185.2</td>
<td>174.5</td>
</tr>
<tr>
<td>5</td>
<td>214.9</td>
<td>204.2</td>
<td>171.7</td>
<td>169.9</td>
<td>168.0</td>
</tr>
<tr>
<td>6</td>
<td>222.7</td>
<td>203.9</td>
<td>194.1</td>
<td>191.1</td>
<td>188.1</td>
</tr>
<tr>
<td>7</td>
<td>221.4</td>
<td>197.9</td>
<td>197.1</td>
<td>195.3</td>
<td>190.1</td>
</tr>
<tr>
<td>8</td>
<td>226.0</td>
<td>204.2</td>
<td>202.0</td>
<td>198.9</td>
<td>195.3</td>
</tr>
<tr>
<td>9</td>
<td>227.5</td>
<td>213.2</td>
<td>195.1</td>
<td>194.5</td>
<td>193.1</td>
</tr>
<tr>
<td>10</td>
<td>228.0</td>
<td>212.1</td>
<td>201.0</td>
<td>200.9</td>
<td>198.6</td>
</tr>
<tr>
<td>11</td>
<td>235.1</td>
<td>211.8</td>
<td>204.6</td>
<td>199.4</td>
<td>185.8</td>
</tr>
<tr>
<td>12</td>
<td>233.9</td>
<td>208.1</td>
<td>203.7</td>
<td>203.7</td>
<td>202.4</td>
</tr>
<tr>
<td>13</td>
<td>221.9</td>
<td>204.6</td>
<td>196.4</td>
<td>196.2</td>
<td>188.3</td>
</tr>
<tr>
<td>14</td>
<td>225.6</td>
<td>205.0</td>
<td>204.6</td>
<td>202.2</td>
<td>194.9</td>
</tr>
<tr>
<td>15</td>
<td>226.8</td>
<td>213.0</td>
<td>201.7</td>
<td>197.2</td>
<td>194.5</td>
</tr>
<tr>
<td>16</td>
<td>227.2</td>
<td>212.1</td>
<td>201.3</td>
<td>196.9</td>
<td>195.9</td>
</tr>
<tr>
<td>17</td>
<td>231.2</td>
<td>219.3</td>
<td>201.6</td>
<td>201.2</td>
<td>201.1</td>
</tr>
<tr>
<td>18</td>
<td>234.0</td>
<td>207.9</td>
<td>205.1</td>
<td>205.0</td>
<td>183.1</td>
</tr>
<tr>
<td>19</td>
<td>238.9</td>
<td>216.3</td>
<td>211.4</td>
<td>206.4</td>
<td>190.0</td>
</tr>
<tr>
<td>20</td>
<td>237.8</td>
<td>216.8</td>
<td>210.7</td>
<td>205.0</td>
<td>182.9</td>
</tr>
<tr>
<td>21</td>
<td>237.7</td>
<td>213.4</td>
<td>212.1</td>
<td>208.2</td>
<td>187.9</td>
</tr>
<tr>
<td>22</td>
<td>217.3</td>
<td>189.2</td>
<td>185.3</td>
<td>182.3</td>
<td>164.6</td>
</tr>
<tr>
<td>23</td>
<td>222.9</td>
<td>199.9</td>
<td>189.5</td>
<td>186.2</td>
<td>176.2</td>
</tr>
<tr>
<td>24</td>
<td>226.5</td>
<td>220.2</td>
<td>193.9</td>
<td>193.1</td>
<td>179.7</td>
</tr>
</tbody>
</table>
The proton affinities of most of the azoles considered are much higher compared with pyrrole. Two factors may be put forward to explain such a behavior. First, azoles are electron rich compared with pyrrole; second, and more importantly, the protonation disrupts aromaticity in pyrrole. Optimized geometries of the most stable protonated amines are depicted in Figure 1. In contrast, protonation of the azole lone pair leaves the \( \pi \) framework intact; thus, the aromaticity in azoles is retained even after protonation. Among the series of compounds under study, imidazoles have substantially higher proton affinity compared with the rest. The higher stability of the protonated imidazoles are to be traced to the preferential topological charge stabilization and more strongly delocalized \( \pi \) framework. Similarly, the higher substitution by the alkyl group makes the ring electron rich, thereby increasing the proton affinity of the substituted compounds.

Take, for example, pyrazole (2), which has a proton affinity value of 214 kcal/mol. While a single methyl substitution (6, 7, 13, and 14) increases the proton affinity value by almost 5 kcal/mol, substitution by two methyl groups (8, 9, 15, and 16) increases the proton affinity value by another 5 kcal/mol. Substitution by three methyl groups (10 and 17) leads to a further increase in the proton affinity value. Very similar observations were made in the case of imidazoles as well, as the unsubstituted imidazole has a proton affinity value of \( \sim 225 \) kcal/mol.

### Proton Affinity of Five-Membered Heterocyclic Amines

<table>
<thead>
<tr>
<th>Compound</th>
<th>B3LYP 6-31+ + G**</th>
<th>MP2/6-311 + + G**a</th>
<th>HF/6-311 + + G**a</th>
<th>G3B3</th>
<th>Exptb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>210.9</td>
<td>209.6</td>
<td>203.5</td>
<td>219.1</td>
<td>209.1</td>
</tr>
<tr>
<td>2</td>
<td>214.8</td>
<td>215.9</td>
<td>210.5</td>
<td>221.3</td>
<td>213.9</td>
</tr>
<tr>
<td>3</td>
<td>226.6</td>
<td>227.7</td>
<td>224.2</td>
<td>232.5</td>
<td>226.0</td>
</tr>
<tr>
<td>4</td>
<td>214.2</td>
<td>215.3</td>
<td>213.5</td>
<td>219.0</td>
<td>213.6</td>
</tr>
<tr>
<td>5</td>
<td>212.3</td>
<td>213.5</td>
<td>209.9</td>
<td>218.8</td>
<td>212.3</td>
</tr>
<tr>
<td>6</td>
<td>218.5</td>
<td>221.3</td>
<td>215.0</td>
<td>226.8</td>
<td>218.8</td>
</tr>
<tr>
<td>7</td>
<td>219.0</td>
<td>220.3</td>
<td>214.8</td>
<td>224.9</td>
<td>217.6</td>
</tr>
<tr>
<td>8</td>
<td>223.5</td>
<td>224.7</td>
<td>222.8</td>
<td>229.9</td>
<td>222.6</td>
</tr>
<tr>
<td>9</td>
<td>225.0</td>
<td>226.3</td>
<td>219.4</td>
<td>232.1</td>
<td>224.0</td>
</tr>
<tr>
<td>10</td>
<td>227.3</td>
<td>227.3</td>
<td>220.6</td>
<td>232.6</td>
<td>226.5</td>
</tr>
<tr>
<td>11</td>
<td>232.0</td>
<td>233.1</td>
<td>229.0</td>
<td>238.3</td>
<td>231.2</td>
</tr>
<tr>
<td>12</td>
<td>229.6</td>
<td>232.2</td>
<td>228.4</td>
<td>236.7</td>
<td>230.4</td>
</tr>
<tr>
<td>13</td>
<td>220.0</td>
<td>220.5</td>
<td>215.0</td>
<td>225.8</td>
<td>218.9</td>
</tr>
<tr>
<td>14</td>
<td>222.9</td>
<td>224.6</td>
<td>219.0</td>
<td>229.4</td>
<td>222.4</td>
</tr>
<tr>
<td>15</td>
<td>224.6</td>
<td>225.9</td>
<td>219.7</td>
<td>231.3</td>
<td>223.5</td>
</tr>
<tr>
<td>16</td>
<td>224.6</td>
<td>225.8</td>
<td>219.6</td>
<td>231.4</td>
<td>223.3</td>
</tr>
<tr>
<td>17</td>
<td>229.0</td>
<td>230.4</td>
<td>223.6</td>
<td>236.1</td>
<td>227.6</td>
</tr>
<tr>
<td>18</td>
<td>231.4</td>
<td>232.6</td>
<td>229.1</td>
<td>237.6</td>
<td>230.9</td>
</tr>
<tr>
<td>19</td>
<td>236.3</td>
<td>237.5</td>
<td>233.3</td>
<td>242.6</td>
<td>235.6</td>
</tr>
<tr>
<td>20</td>
<td>235.1</td>
<td>236.2</td>
<td>232.5</td>
<td>241.0</td>
<td>234.3</td>
</tr>
<tr>
<td>21</td>
<td>235.4</td>
<td>236.6</td>
<td>232.8</td>
<td>241.3</td>
<td>234.8</td>
</tr>
<tr>
<td>22</td>
<td>216.4</td>
<td>217.1</td>
<td>214.9</td>
<td>220.4</td>
<td>215.8</td>
</tr>
<tr>
<td>23</td>
<td>221.8</td>
<td>222.4</td>
<td>219.5</td>
<td>226.2</td>
<td>220.9</td>
</tr>
<tr>
<td>24</td>
<td>223.8</td>
<td>225.0</td>
<td>220.0</td>
<td>229.8</td>
<td>223.0</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>1.43</td>
<td>2.57</td>
<td>2.48</td>
<td>7.71</td>
<td>0.91</td>
</tr>
</tbody>
</table>

---

* a Single point calculations on B3LYP/6-31 + + G** optimized geometries.
  b Data taken from Hunter and Lias, also available at webook.nist.gov.
  c Mean absolute deviation (\( \Delta \)) taken with respect to the experimental value.
kcal/mol, and the dimethyl-substituted has enhanced its proton affinity ability by almost 10 kcal/mol.

A quick review of Table II indicates that B3LYP/6-31+G** proton affinity values are in excellent agreement with experimental values in most cases. However, while most of the computed results are within 1 kcal/mol, for the compounds 4, 6, 12, 18, and 23, the calculated proton affinities have shown some noticeable deviations within the range of 1.5–3.4 kcal/mol.

Figure 2 depicts the correlation between experimental and computed proton affinities at different levels. Excellent linear correlation was obtained between the experimental and theoretical values at the B3LYP/6-31+G** and G3B3 levels of theory. Thus, the computational study is consistent with the experimental results and also establishes the protonation site in each of the isomers. Not surprisingly, our systematic computations reveal that HF calculations deviate substantially from the experimental values. However, it is rather surprising to see that the MP2 level fails to improve substantially over the HF results. The correlation coefficients ($r$-value) were found to be 0.9487 and 0.9655 for HF and MP2 methods, respectively. However, B3LYP
calculations have led to substantial improvement in agreement with the experimental results. The G3B3 calculations as well as B3LYP calculations augmented with polarization and diffuse functions on hydrogen have been found to be in excellent agreement with the experimental results.

**Conclusions**

Known proton affinity values of 24 five-membered heterocyclic amines were considered to assess the performance of the computational methods. The theoretical study also unambiguously establishes the regioselectivity of protonation in all 24 heterocyclic amines considered. G3B3 calculations have shown excellent agreement with experimental results. For the series of compounds studied in this article, the performance of B3LYP/6-31+G** level of theory is excellent in modeling proton affinities and is comparable to the performance of G3B3 level. While the traditional correlated levels require large basis sets to reproduce the results correctly, a basis set of double-ζ with polar-
ization and diffuse functions on the heavy atoms and hydrogens appears to be sufficient to model the systems.

ACKNOWLEDGMENTS

J. S. R. thanks CSIR for the fellowship.

References