

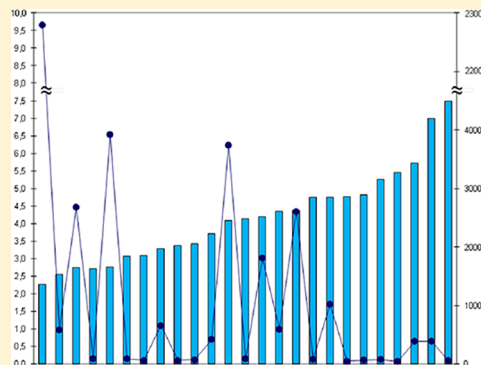
# Critical Test of Some Computational Chemistry Methods for Prediction of Gas-Phase Acidities and Basicities

Eve Toomsalu, Ilmar A. Koppel, and Peeter Burk\*

Institute of Chemistry, University of Tartu, Ravila 14a, Tartu, Estonia

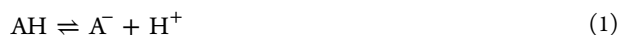
**S** Supporting Information

**ABSTRACT:** Gas-phase acidities and basicities were calculated for 64 neutral bases (covering the scale from 139.9 kcal/mol to 251.9 kcal/mol) and 53 neutral acids (covering the scale from 299.5 kcal/mol to 411.7 kcal/mol). The following methods were used: AM1, PM3, PM6, PDDG, G2, G2MP2, G3, G3MP2, G4, G4MP2, CBS-QB3, B1B95, B2PLYP, B2PLYPD, B3LYP, B3PW91, B97D, B98, BLYP, BMK, BP86, CAM-B3LYP, HSEh1PBE, M06, M062X, M06HF, M06L, mPW2PLYP, mPW2PLYPD, O3LYP, OLYP, PBE1PBE, PBE1PBE, tHCTHhyb, TPSSH, VSXC, X3LYP. The addition of the Grimmes empirical dispersion correction (D) to B2PLYP and mPW2PLYP was evaluated, and it was found that adding this correction gave more-accurate results when considering acidities. Calculations with B3LYP, B97D, BLYP, B2PLYPD, and PBE1PBE methods were carried out with five basis sets (6-311G\*\*, 6-311+G\*\*, TZVP, cc-pVTZ, and aug-cc-pVTZ) to evaluate the effect of basis sets on the accuracy of calculations. It was found that the best basis sets when considering accuracy of results and needed time were 6-311+G\*\* and TZVP. Among semiempirical methods AM1 had the best ability to reproduce experimental acidities and basicities (the mean absolute error (mae) was 7.3 kcal/mol). Among DFT methods the best method considering accuracy, robustness, and computation time was PBE1PBE/6-311+G\*\* (mae = 2.7 kcal/mol). Four Gaussian-type methods (G2, G2MP2, G4, and G4MP2) gave similar results to each other (mae = 2.3 kcal/mol). Gaussian-type methods are quite accurate, but their downside is the relatively long computational time.



## 1. INTRODUCTION

Proton transfer reactions are very important in chemical and biochemical processes where the acidity or basicity of the compounds is manifested. Gas-phase acid–base equilibria are important for understanding the properties of the substances, as those reactions reveal the impact of the substituent within the isolated molecule without any interference from the solvent. The gas-phase acidity of Brønsted acids is defined as the Gibbs energy  $\Delta G$  of reaction 1:



Similarly, the gas-phase basicity is defined as the  $\Delta G$  of reaction 2:



Sometimes deprotonation enthalpies and proton affinities are used instead of acidities or basicities. These correspond to the  $\Delta H$  of the above given reactions (1 and 2). Stronger acids have smaller numerical values of acidities and deprotonation enthalpies (the release of the proton is easier) while the stronger bases have numerically larger values of the gas-phase basicities and proton affinities (the binding of proton is stronger). During the last few decades, a considerable amount of experimental gas-phase acidities and basicities have been measured.<sup>1</sup>

The application of quantum chemical and density functional theory calculations in different fields of chemistry has been

explosively growing during the last 20 years. Gas-phase acidities and basicities can also be predicted with different computational methods and different computational descriptors can be used for rationalizing the trends in acidities and basicities.<sup>2–8</sup>

Many different computational methodologies have been developed over the last 20 years for more or less accurate calculations of different energetic properties: the Gaussian 1–4 family of methods,<sup>9–16</sup> different schemes for extrapolation to the complete basis set (CBS),<sup>17,18</sup> and many different density functional theory (DFT) methods.<sup>19</sup> The computational efficiency of DFT has motivated the development of a huge number of functionals and their implementation into available programs.

In the year 2001, J. Perdew proposed his classification of density functional methods, the so-called Jacob's ladder. Five rungs of the ladder represent the hierarchy of density approximations: on the first rung is the local density approximation (LDA), on the second rung is the generalized gradient approximation (GGA), on the third rung is the meta-generalized gradient approximation (M-GGA), on the fourth rung are the hybrid generalized gradient approximation (H-GGA) and the hybrid meta-generalized gradient approximation (MH-GGA), and finally on the fifth rung is the fully nonlocal

Received: May 14, 2013

description. Within this view, each rung should add something to the rung below. Although each functional has its own strengths and weaknesses, the quality of the results should increase, when moving from the lower rungs to the higher rungs. The accuracy of the results is also influenced by the basis sets used.<sup>20</sup>

As a result, the newcomer to the field of computational chemistry faces severe problems in choosing the right methodology for solving the problem at hand. The rational choice of the computational method inevitably requires a comparison of calculated results with experimental data to validate the methodology. In that respect, calculations of acidities and basicities are no exception.

There are many articles about using different computational methods for calculation of acidities and basicities, but usually a few compounds are investigated and only some comparison between different methods has been made.<sup>2,6,8,21–33</sup>

Dewar and Dieter have calculated proton affinities for 60 compounds and deprotonation enthalpies for 80 compounds by using the AM1 model. It was found that AM1 was an effective tool for studying processes involving deprotonation or protonation of neutral molecules. The main problems encountered with AM1 involved small anions, where the charge is largely concentrated on one atom, and anions formed by the deprotonation of oximes. It was also found that deprotonation energies of alcohols and proton affinities of amines are not accurately reproduced by AM1.<sup>21</sup> Burk et al. calculated gas-phase acidities for 175 compounds and proton affinities for 119 compounds with both the AM1 and PM3 method. They compared the results of the experiment, and calculations with AM1 and PM3 methods. It was found that, with some exceptions, the PM3 method seems to be a useful tool for calculating gas-phase acidities; however, for predicting proton affinities, AM1 performs better.<sup>22,23</sup>

Smith and Radom have tested Gaussian methods such as G2, G2(MP2), G2(MP2, SVP), and other methods (such as BP86, BLYP, B3P86, B3LYP, QCISD(T), MP2, MP4, and F4) to reproduce experimental proton affinities and deprotonation energies for a set of small molecules. They found that the accuracy of G2(MP2,SVP) is comparable to the G2 results for proton affinities, but the computational cost is significantly lower. Revision of the experimental proton affinities was suggested for some compounds. The G2, G2(MP2), and G2(MP2, SVP) methods were also successful for prediction of deprotonation energies. MP4 and F4 gave good results for conjugated systems, but needed a bigger basis set (6-311+G(3df,2p) was used). The quality of the B3LYP results was comparable with those of G2, but the computational cost was much lower. In the case of B3P86, the bigger basis set (6-311+G(3df,2p)) gave less accurate results, compared to the smaller one (6-311+G\*\*).<sup>24,25</sup>

Merrill and Kass have calculated deprotonation enthalpies for 35 acids, using seven different DFT methods and seven different basis sets. The best results were obtained with the B3PW91/6-311++G(df,dp) method.<sup>26</sup>

Burk et al. have calculated gas-phase acidities and basicities for 49 acids and 32 bases using the B3LYP functional with four different basis sets (6-31G\*, 6-31+G\*, 6-311+G\*\*, and 6-311+G(3df,3pd)). The best results were obtained with the 6-311+G(3df,3pd) basis set, where the average absolute errors were below 2.5 kcal/mol both for acidities and basicities.<sup>27</sup>

Koppel et al. have calculated acidities for classical strong mineral acids HClO<sub>4</sub>, CF<sub>3</sub>SO<sub>3</sub>H, FSO<sub>3</sub>H, H<sub>2</sub>SO<sub>4</sub>, HBF<sub>4</sub>, HPO<sub>3</sub>, and HNO<sub>3</sub> with G2 or G2(MP2) methods. The acidities for 39 neutral strong or superstrong Brønsted acids, Brønsted–Lewis

conjugate acids, and some compounds modeling the acidic clusters of zeolites were also calculated using the DFT B3LYP/6-311+G\*\* approach. It was found that the G2 and G2(MP2) methods are better than B3LYP/6-311+G\*\* for calculating acidities. The DFT method overestimates the acidity of strongly acidic compounds ( $\Delta G_{\text{acid}} \leq 340$  kcal/mol) and underestimates the acidity of the weakest acids ( $\Delta G_{\text{acid}} \geq 380$  kcal/mol) but gave quite good agreement with experimental values for moderately acidic compounds ( $340 \text{ kcal/mol} \leq \Delta G_{\text{acid}} \leq 380 \text{ kcal/mol}$ ).<sup>2</sup>

Vayner and Ball have calculated proton affinities (PAs) for aziridine, azetidine, pyrrolidine, and piperidine using different methods (HF, BLYP, B3LYP, SVWN, MP2) and basis sets (STO-3G, 4-31G, 6-31G\*, 6-311G\*\*). SVWN gave the most accurate results compared to the experiment, but BLYP and B3LYP were also as accurate as MP2, at less computational cost. The influence of basis set size was studied with the B3LYP method and it was shown that the bigger basis sets do not guarantee better accuracy. They concluded that the studied methods, in combination with moderately sized basis sets, were sufficient to predict accurate PAs of cyclic amines.<sup>28</sup>

Aue et al. have calculated gas-phase basicities (GBs) for 40 polynuclear aromatic hydrocarbons and related hydrocarbons using a semiempirical AM1 method, ab initio HF/6-31G, HF/6-31G\*, MP2/6-31G\*, MP2/6-31+G\*\* methods, and the DFT B3LYP/6-311+G\*\* method. Statistical analysis shows standard errors of 2.12, 1.67, 1.53, 1.36, and 1.55 kcal/mol for each of the five methods, respectively. The results permitted an evaluation of the reliability of experimental data and indicated possible experimental problems that need re-evaluation. Predictions were made for GBs for 12 new PAHs with GBs not yet measured.<sup>29</sup>

Pokon et al. have calculated deprotonation energies for 17 acids with the G3, CBS-QB3, and CBS-APNO methods. They determined mean absolute deviations of 0.84–1.26 kcal/mol and root-mean-square (RMS) errors of 1.06–1.49 kcal/mol. The CBS-APNO method had the best overall agreement with experimental values, followed by G3 and CBS-QB3 methods.<sup>30</sup>

Li et al. have calculated GBs for 20 amines and phosphines using different methods (HF, B3LYP, B3P86, BH&HLYP, PBEPBE, MPWPW91) and basis sets (6-31G\*, 6-31+G\*, 6-31G\*\*, 6-31++G\*\*, 6-311G(2df,2p), 6-31++G(2df,2p)). They found that HF could not reliably predict the gas-phase basicities (the mean error was 6.5 kcal/mol and the RMS error was 6.8 kcal/mol). Among all the investigated density functionals, it was found that the BH&HLYP method gave the worst predictions (mean error = 3.3 kcal/mol, RMS error = 2.7 kcal/mol). It was found that the PBEPBE/6-311++(2df,2p)/PBEPBE/6-31G\* method was sufficiently accurate (mean error = 0.0 kcal/mol, RMS error = 1.3 kcal/mol).<sup>31</sup>

Range et al. used five multilevel methods (CBS-QB3, G3B3, G3MP2B3, MCG3/3, and MC-QCISD/3) and six density functional methods (PBE0, B1B95, B3LYP, MPW1KCIS, PBE1KCIS, MPW1B95) to calculate the gas-phase basicities and proton affinities for a series of 17 molecules relevant to the study of phosphoryl transfer in biological systems. It was found that multicoefficient correlation methods (MCG3/3 and MC-QCISD) slightly outperformed the other methods tested. They found also that four density functional methods (PBE1KCIS, MPW1B95, PBE0, and B1B95) can perform nearly as well as the multilevel methods.<sup>32</sup>

He et al. have calculated the gas-phase basicities for 41 small molecules. Geometries were optimized at the MP2/aug-cc-pVTZ level. MP2 and CCSD(T) electron correlation methods

Table 1. Used Bases and Corresponding Experimental Gas-Phase Basicities

compound	gas-phase basicity, GB (kcal/mol)	molecular formula	compound	gas-phase basicity, GB (kcal/mol)	molecular formula
sulfuryl fluoride	139.9	SO <sub>2</sub> F <sub>2</sub>	butan-1-ol	181.4	C <sub>4</sub> H <sub>10</sub> O
bis-(2,2,2-trifluoromethyl)ether	144.2	C <sub>2</sub> F <sub>6</sub> O	isopropyl alcohol	182.3	C <sub>3</sub> H <sub>8</sub> O
sulfur dioxide	146.6	SO <sub>2</sub>	propionic acid	183.1	C <sub>3</sub> H <sub>6</sub> O <sub>2</sub>
ethyne	147.4	C <sub>2</sub> H <sub>2</sub>	nitrobenzene	183.9	C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub>
2,2,2-trifluoroethanal	149.6	C <sub>2</sub> HOF <sub>3</sub>	2-methylpropene	185.4	C <sub>4</sub> H <sub>8</sub>
difluoroamine	151.8	NHF <sub>2</sub>	benzenenitrile	186.6	C <sub>7</sub> H <sub>5</sub> N
water	158.3	H <sub>2</sub> O	acetone	186.9	C <sub>3</sub> H <sub>6</sub> O
pentafluorobenzene	158.4	C <sub>6</sub> HF <sub>5</sub>	butan-2-ol	187.5	C <sub>4</sub> H <sub>10</sub> O
1,1,3,3-tetrafluoro-2-propanone	159.9	C <sub>3</sub> H <sub>2</sub> F <sub>4</sub> O	phenol	187.9	C <sub>6</sub> H <sub>6</sub> O
2,2,2-trifluoroethanol	160.1	C <sub>2</sub> H <sub>3</sub> F <sub>3</sub> O	benzoic acid	188.8	C <sub>7</sub> H <sub>6</sub> O <sub>2</sub>
1,2,3,4-tetrafluorobenzene	160.8	C <sub>6</sub> H <sub>2</sub> F <sub>4</sub>	4-fluorobenzaldehyde	190.1	C <sub>7</sub> H <sub>5</sub> FO
trifluoroacetic acid	162.7	C <sub>2</sub> HF <sub>3</sub> O <sub>2</sub>	pentan-2-one	191.4	C <sub>5</sub> H <sub>10</sub> O
1,1,1-trifluoro-2-propanone	165.4	C <sub>3</sub> H <sub>3</sub> F <sub>3</sub> O	methoxybenzene	192.9	C <sub>7</sub> H <sub>8</sub> O
1,4-difluorobenzene	165.6	C <sub>6</sub> H <sub>4</sub> F <sub>2</sub>	heptan-4-one	194.9	C <sub>7</sub> H <sub>14</sub> O
1,2,3-trifluorobenzene	166.5	C <sub>6</sub> H <sub>3</sub> F <sub>3</sub>	methoxyethene	198.4	C <sub>3</sub> H <sub>6</sub> O
1,2-difluorobenzene	168.1	C <sub>6</sub> H <sub>4</sub> F <sub>2</sub>	ethanamide	199.0	C <sub>2</sub> H <sub>5</sub> NO
methanoic acid	169.8	CH <sub>2</sub> O <sub>2</sub>	4-nitrobenzenamine	199.4	C <sub>6</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>
butadiyne	170.4	C <sub>4</sub> H <sub>2</sub>	phenyl methyl sulfide	201.6	C <sub>7</sub> H <sub>8</sub> S
1,2,4,5-tetrafluorobenzene	171.8	C <sub>6</sub> H <sub>2</sub> F <sub>4</sub>	aniline	203.3	C <sub>6</sub> H <sub>7</sub> N
1,2,3,5-tetrafluorobenzene	172.0	C <sub>6</sub> H <sub>2</sub> F <sub>4</sub>	1-amino-3-methylbenzene	206.5	C <sub>7</sub> H <sub>11</sub> N
1,3-difluorobenzene	173.0	C <sub>6</sub> H <sub>4</sub> F <sub>2</sub>	3-aminophenol	207.2	C <sub>6</sub> H <sub>7</sub> NO
chlorobenzene	173.2	C <sub>6</sub> H <sub>5</sub> Cl	pyridazine	209.6	C <sub>4</sub> H <sub>4</sub> N <sub>2</sub>
benzene	173.4	C <sub>6</sub> H <sub>6</sub>	2-propanamine	212.5	C <sub>3</sub> H <sub>9</sub> N
bromobenzene	173.5	C <sub>6</sub> H <sub>5</sub> Br	<i>N,N</i> -dimethyl- <i>p</i> -tolylamine	219.4	C <sub>9</sub> H <sub>13</sub> N
fluorobenzene	173.7	C <sub>6</sub> H <sub>5</sub> F	trimethylamine	219.4	C <sub>3</sub> H <sub>9</sub> N
2-methoxy-1-ethanol	174.4	C <sub>3</sub> H <sub>8</sub> O <sub>2</sub>	1-methylpiperidine	224.7	C <sub>6</sub> H <sub>13</sub> N
chloroacetic acid	175.5	C <sub>2</sub> H <sub>3</sub> ClO <sub>2</sub>	1,4,5,6-tetrahydropyrimidine	231.3	C <sub>4</sub> H <sub>8</sub> N <sub>2</sub>
fluoroacetic acid	175.5	C <sub>2</sub> H <sub>3</sub> FO <sub>2</sub>	<i>N,N</i> -dimethyl- <i>N'</i> -(1-methylpropyl)-methanimidamide	235.6	C <sub>7</sub> H <sub>16</sub> N <sub>2</sub>
1-methyl-4-chlorobenzene	175.7	C <sub>7</sub> H <sub>7</sub> Cl	1,5-diazabicyclo[4.3.0]non-5-ene	240.4	C <sub>7</sub> H <sub>12</sub> N <sub>2</sub>
1-fluoro-4-methylbenzene	175.9	C <sub>7</sub> H <sub>7</sub> F	1,3,4,6,7,8-hexahydro-1-methyl-2H-pyrimido[1,2- $\alpha$ ]pyrimidine	246.2	C <sub>8</sub> H <sub>15</sub> N <sub>3</sub>
4-bromo-1-methylbenzene	178.3	C <sub>7</sub> H <sub>7</sub> Br	iminotris(dimethylamino)phosphorane	249.7	C <sub>6</sub> H <sub>19</sub> N <sub>4</sub> P
toluene	180.8	C <sub>7</sub> H <sub>8</sub>	tris(dimethylamino)(methylimino)-phosphorane	251.9	C <sub>7</sub> H <sub>21</sub> N <sub>4</sub> P

extrapolated to the complete basis set limit were used. The overall accuracy of calculated gas-phase basicities, in descending order, is CCSD(T)\_CBS > CCSD(T)/aug-cc-pVDZ > MP2/aug-cc-pVQZ  $\approx$  MP2\_CBS > HF/aug-cc-pVQZ. The best RMS error obtained was 1.0 kcal/mol at the CCSD(T)\_CBS//MP2/aug-cc-pVTZ level. Conformational effects were also found to be relevant in the case of several flexible molecules.<sup>33</sup>

Bras et al. have analyzed how well DFT functionals, which are often used to characterize complex and large molecules such as proteins, describe the proton affinity of the ionizable side chains of lysine, histidine, and arginine, as well as the anions of aspartate, cysteine, serine, and tyrosine anions. The reference values were determined at the CCSD(T)/CBS level. A new benchmarking database for PA and for proton transfer between different ionizable side chains was provided. Among the 64 density functionals tested, the MPW1B95-D3, XYG3, MPW1B95, B1B95-D3, BMK, BMK-D3, M06-2X, B1LYP, B1B95, PBE1PBE, CAM-B3LYP, B97-1, PBE1KCIS, B3P86, CAM-

B3LYP-D3, B3LYP, B98, M06-L, and M06 provided the most accurate PA values for all ionizable amino acids studied, with errors below 1.5 kcal/mol. It was concluded that M06-2X was the most-accurate density functional for proton transfers between different amino acids.<sup>8</sup>

Vesecchi and Galembeck have studied  $\gamma$ -butyrolactone and 2-pyrrolidinone using the composite chemical models G2, G2MP2, MP2, G3, CBS-Q, CBS-4, and CBS-QB3; the DFT methods B3LYP, B3P86, PW91PW91, mPW1PW, and B98; and the basis sets 6-31G(d), 6-31+ G(d), 6-31G(d,p), 6-31+G(d,p), 6-31++G(d,p), 6-311G(d), 6-311+G(d), 6-311G(d,p), 6-311+G(d,p), 6-311++G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ. Values obtained for the enthalpies of formation, proton affinity, and gas-phase basicity of the two target molecules were compared to the experimental data reported in the literature. The best results were achieved with the use of DFT models, and the B3LYP method led to the most-accurate data.<sup>6</sup>

Table 2. Used Acids and Corresponding Experimental Gas-Phase Acidities

compound	molecular formula	gas-phase acidity, GA (kcal/mol)	compound	molecular formula	gas-phase acidity, GA (kcal/mol)
trifluoromethanesulfonic acid	CF <sub>3</sub> SO <sub>3</sub> H	299.5	2-butanone	C <sub>4</sub> H <sub>8</sub> O	360.4
fluorosulfuric acid	HFO <sub>3</sub> S	299.8	1,3,5-trichlorobenzene	C <sub>6</sub> H <sub>3</sub> Cl <sub>3</sub>	360.8
sulfuric acid	H <sub>2</sub> SO <sub>4</sub>	302.3	1,2,4-trifluorobenzene	C <sub>6</sub> H <sub>3</sub> F <sub>3</sub>	362.6
2,4,6-trinitrophenol	C <sub>6</sub> H <sub>3</sub> N <sub>3</sub> O <sub>7</sub>	302.8	benzyl chloride	C <sub>7</sub> H <sub>7</sub> Cl	364.9
metaphosphoric acid	HPO <sub>3</sub>	303.3	pentafluoroethane	C <sub>2</sub> HF <sub>5</sub>	366.8
methanedisulfonyl difluoride	CH <sub>2</sub> F <sub>2</sub> O <sub>4</sub> S <sub>2</sub>	307.3	3,3,3-trifluoro-2-methyl-1-propene	C <sub>4</sub> H <sub>5</sub> F <sub>3</sub>	367.4
2,4-dinitrophenol	C <sub>6</sub> H <sub>4</sub> N <sub>2</sub> O <sub>5</sub>	308.1	1,2-bis(trifluoromethyl)-benzene	C <sub>8</sub> H <sub>4</sub> F <sub>6</sub>	368.5
1,1,1,5,5,5-hexafluoro-2,4-pentandione	C <sub>5</sub> H <sub>2</sub> F <sub>6</sub> O <sub>2</sub>	311.3	isopropylbenzene	C <sub>9</sub> H <sub>12</sub>	372.9
(trifluoromethyl)thiolacetic acid	C <sub>2</sub> HF <sub>3</sub> OS	312.5	isocyanomethane	C <sub>2</sub> H <sub>3</sub> N	373.1
3,5-bis(trifluoromethyl)benzoic acid	C <sub>6</sub> H <sub>4</sub> F <sub>6</sub> O <sub>2</sub>	317.4	toluene	C <sub>7</sub> H <sub>8</sub>	373.7
nitric acid	HNO <sub>3</sub>	317.8	nitrobenzene	C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub>	377.0
4-nitrobenzoic acid	C <sub>7</sub> H <sub>5</sub> NO <sub>4</sub>	321.1	N-methyl methanimine	C <sub>2</sub> H <sub>5</sub> N	378.0
trifluoromethanesulfonamide	CF <sub>3</sub> SO <sub>2</sub> NH <sub>2</sub>	321.3	2-methyl-1-propene	C <sub>4</sub> H <sub>8</sub>	379.6
dichloroacetic acid	C <sub>2</sub> H <sub>2</sub> Cl <sub>2</sub> O <sub>2</sub>	321.9	propene	C <sub>3</sub> H <sub>6</sub>	382.5
3,4,5-trichlorophenol	C <sub>6</sub> H <sub>3</sub> Cl <sub>3</sub> O	323.9	propanenitrile	C <sub>3</sub> H <sub>5</sub> N	384.0
2-(trifluoromethyl)-1,1,1,3,3,3-hexafluoropropane	C <sub>4</sub> HF <sub>9</sub>	326.6	water	H <sub>2</sub> O	384.1
benzoic acid	C <sub>7</sub> H <sub>6</sub> O <sub>2</sub>	333.0	chlorobenzene	C <sub>6</sub> H <sub>5</sub> Cl	386.7
2-propenoic acid	C <sub>3</sub> H <sub>4</sub> O <sub>2</sub>	337.2	chloromethane	CH <sub>3</sub> Cl	389.1
phenol	C <sub>6</sub> H <sub>6</sub> O	342.3	1,3-butadiene	C <sub>4</sub> H <sub>6</sub>	391.3
m-nitroaniline	C <sub>6</sub> H <sub>6</sub> N <sub>2</sub> O <sub>2</sub>	345.0	benzene	C <sub>6</sub> H <sub>6</sub>	392.9
4-aminophenol	C <sub>6</sub> H <sub>7</sub> NO	345.6	hydrogen	H <sub>2</sub>	394.2
4-(trifluoromethyl)benzenamine	C <sub>7</sub> H <sub>6</sub> F <sub>3</sub> N	346.0	methylamine	CH <sub>5</sub> N	394.5
pentafluorobenzene	C <sub>6</sub> HF <sub>5</sub>	349.0	trimethylamine	C <sub>3</sub> H <sub>9</sub> N	398.0
nitromethane	CH <sub>3</sub> NO <sub>2</sub>	349.7	1,1,1-trifluoropropane	C <sub>3</sub> H <sub>5</sub> F <sub>3</sub>	398.5
acetaldehyde	C <sub>2</sub> H <sub>4</sub> O	359.0	methoxyethene	C <sub>3</sub> H <sub>6</sub> O	403.3
aniline	C <sub>6</sub> H <sub>7</sub> N	359.1	ethane	C <sub>2</sub> H <sub>6</sub>	411.7
2-pentanone	C <sub>5</sub> H <sub>10</sub> O	359.6			

## 2. METHODS

In the current paper, we will present an overview of earlier publications on benchmarking different methodologies for the calculation of acidities and basicities and provide our own results for a diverse and sufficiently large dataset.

Calculations were performed using the GAUSSIAN 09 program package.<sup>34</sup> Full conformational search of all studied neutral molecules and corresponding ions was carried out with the B3LYP/6-311+G\*\* method to ensure that the global minima on the potential energy surfaces were found. Only conformers energetically within 3 kcal/mol from the most stable were considered with other methods. All geometries were fully optimized and the frequencies calculated at all investigated levels of theory. No scaling was applied to obtained frequencies for the calculation of thermodynamic parameters (but see the discussion in Section 3.5), which were corrected to 298.15 K. All stationary points were found to be true minima (Nimag = 0). Since the B2PLYPD/aug-cc-pVTZ calculations of frequencies were prohibitively expensive for us, we used the thermodynamic parameters from B2PLYPD/cc-pVTZ calculations to evaluate the frequencies.

The gas-phase acidities (GAs) and GBs were calculated for 64 neutral bases and 53 neutral acids. Compounds were chosen so they would evenly cover a wide range of scale. The basicity scale started with sulfuryl fluoride (139.9 kcal/mol) and ended with tris(dimethylamino)(methylimino)phosphorane (251.9 kcal/mol). The strongest acid was trifluoromethanesulfonic acid (299.5 kcal/mol), and the weakest was ethane (411.7 kcal/mol). Most of the acidities and basicities were taken from the National

Institute of Standards and Technology (NIST) database.<sup>1</sup> Experimental GB values for phosphazenes were taken from Kaljurand et al.<sup>35</sup> and GB values for weak bases (weaker than water) were taken from Leito et al.<sup>7</sup> The acidity of nitrobenzene was taken from Cheng and Grabowski<sup>36</sup> as the NIST value seems unreliable; it strongly deviates from our calculated results. The studied compounds and used acidities and basicities are given in Tables 1 and 2.

The following computational methods were used: semi-empirical methods AM1,<sup>37</sup> PM3,<sup>38,39</sup> PM6,<sup>40</sup> PDDG;<sup>41</sup> Gaussian type methods G2,<sup>11</sup> G2MP2,<sup>12</sup> G3,<sup>13</sup> G3MP2,<sup>14</sup> G4,<sup>15</sup> G4MP2,<sup>16</sup> complete basis set method CBS-QB3,<sup>17,18</sup> density functional theory methods from GGA methods B97D,<sup>42</sup> BLYP,<sup>43–45</sup> BP86,<sup>43,46</sup> OLYP,<sup>44,45,47,48</sup> PBEPBE<sup>49,50</sup> were used; from M-GGA methods M06L,<sup>51</sup> VSXC<sup>52</sup> were used; from H-GGA methods B2PLYP,<sup>53</sup> B3LYP,<sup>44,45,54,55</sup> B3PW91,<sup>54,56</sup> B98,<sup>57</sup> CAM-B3LYP,<sup>58</sup> HSEh1PBE,<sup>59–62</sup> mPW2PLYP,<sup>63</sup> O3LYP,<sup>44,47,48,54</sup> PBE1PBE,<sup>64</sup> X3LYP<sup>65</sup> were used; from HM-GGA methods B1B95,<sup>66</sup> BMK,<sup>67</sup> M06,<sup>68</sup> M06HF,<sup>69,70</sup> M062X,<sup>68</sup> tHCTHhyb<sup>71</sup> and TPSSH<sup>72</sup> were used. Also, the addition of the Grimmes empirical dispersion correction (D)<sup>73</sup> to B2PLYP and mPW2PLYP was evaluated.

Calculations with the B3LYP, B97D, BLYP, B2PLYPD and PBE1PBE methods were carried out with five basis sets (6-311G\*\*, <sup>74</sup> 6-311+G\*\*, <sup>74,75</sup> TZVP,<sup>76</sup> cc-pVTZ,<sup>77,78</sup> and aug-cc-pVTZ<sup>77,78</sup>) to evaluate the effect of the basis set on the accuracy of calculations. Calculated GA and GB values are given in the Supporting Information.



Table 3. Results of Correlation Analysis for Acidities and Basicities Together (in kcal/mol)<sup>a</sup>

	<i>b</i>	<i>a</i>	<i>R</i> <sup>2</sup>	<i>S</i> <sub>0</sub>	<i>N</i>	$\Delta$
AM1	0.99 ± 0.01	4.45 ± 2.77	0.9882	9.7	117	7.3
PDDG	1.00 ± 0.01	9.65 ± 3.63	0.9792	12.7	116	12.1
PM3	0.98 ± 0.01	18.37 ± 3.09	0.9844	11.0	116	11.3
PM6	1.07 ± 0.01	-2.71 ± 3.02	0.9875	9.9	116	8.8
CBS-QB3	0.98 ± 0.00	5.39 ± 1.08	0.9982	3.8	117	2.8
G2	0.99 ± 0.00	2.70 ± 1.12	0.9981	3.9	116	2.3
G2MP2	0.99 ± 0.00	2.51 ± 1.12	0.9981	3.9	116	2.3
G3	0.99 ± 0.00	3.31 ± 1.05	0.9983	3.7	116	2.4
G3MP2	0.99 ± 0.00	2.85 ± 1.06	0.9983	3.7	117	2.4
G4	0.99 ± 0.00	2.79 ± 1.08	0.9982	3.8	116	2.3
G4MP2	0.99 ± 0.00	1.94 ± 1.09	0.9982	3.8	117	2.3
B1B95	1.00 ± 0.00	-0.11 ± 1.27	0.9977	4.3	112	2.7
B2PLYP	1.00 ± 0.00	1.95 ± 1.20	0.9978	4.2	117	2.6
B2PLYPD	1.00 ± 0.00	1.39 ± 1.19	0.9978	4.2	117	2.5
B3LYP	1.01 ± 0.00	-1.34 ± 1.31	0.9975	4.5	117	3.1
B3PW91	1.01 ± 0.00	-2.19 ± 1.24	0.9977	4.3	117	3.0
B97D	1.03 ± 0.00	-9.55 ± 1.33	0.9975	4.4	117	4.1
B98	1.01 ± 0.00	-3.18 ± 1.27	0.9976	4.3	117	3.2
BLYP	1.03 ± 0.01	-5.45 ± 1.42	0.9971	4.8	117	4.7
BMK	0.99 ± 0.00	2.96 ± 1.33	0.9973	4.7	117	3.2
BP86	1.03 ± 0.00	-5.96 ± 1.29	0.9976	4.4	117	4.2
CAM-B3LYP	1.00 ± 0.00	3.04 ± 1.28	0.9975	4.5	117	3.3
HSEh1PBE	1.01 ± 0.00	-1.09 ± 1.23	0.9977	4.2	117	2.8
M06	1.01 ± 0.00	-0.73 ± 1.28	0.9975	4.4	117	3.3
M062X	0.99 ± 0.00	4.85 ± 1.29	0.9974	4.5	117	3.7
M06HF	0.99 ± 0.01	9.37 ± 1.87	0.9944	6.7	117	6.5
M06L	1.01 ± 0.00	-2.71 ± 1.22	0.9978	4.2	117	3.2
mPW2PLYP	1.00 ± 0.00	2.50 ± 1.21	0.9978	4.2	117	2.7
mPW2PLYPD	1.00 ± 0.00	2.20 ± 1.22	0.9977	4.3	117	2.6
O3LYP	1.01 ± 0.00	-4.75 ± 1.26	0.9977	4.3	117	3.5
OLYP	1.02 ± 0.00	-7.67 ± 1.33	0.9975	4.5	117	3.7
PBE1PBE	1.00 ± 0.00	-0.89 ± 1.22	0.9978	4.2	117	2.7
PBEPBE	1.03 ± 0.00	-5.29 ± 1.29	0.9976	4.4	117	4.2
tHCTHhyb	1.01 ± 0.00	-4.12 ± 1.26	0.9977	4.3	117	3.2
TPSSH	1.01 ± 0.00	-3.93 ± 1.28	0.9976	4.4	117	3.4
VSXC	1.02 ± 0.00	-6.42 ± 1.33	0.9975	4.5	117	3.6
X3LYP	1.01 ± 0.00	-0.73 ± 1.30	0.9975	4.5	117	3.1

<sup>a</sup>Legend: *a* is the intercept and *b* is the slope of the correlation line, *R*<sup>2</sup> is the square of correlation coefficient, *S*<sub>0</sub> is the standard deviation of the points from the correlation line, *N* is the number of points, and  $\Delta$  is the average absolute error. 6-311+G\*\* basis set is used with DFT methods.

The GAs were calculated as the Gibbs energy changes for reaction 1, and GBs were calculated as the Gibbs energy changes for reaction 2.

The statistical (linear regression) analysis of the relationships between the calculated and experimental GAs and GBs was performed using linear regression in the form:

$$\Delta G_{\text{exp}} = a + b\Delta G_{\text{calc}} \quad (3)$$

where  $\Delta G_{\text{exp}}$  is the experimental and  $\Delta G_{\text{calc}}$  the calculated Gibbs energy change of reaction 1 or 2. Also, the separate regression analyses were performed for acids and bases, since it has been shown<sup>27,79</sup> that simultaneous correlation of acidities and basicities can lead, especially for the lower levels of calculations, to the regression line with slope close to 1 and intercept 0, because of the systematic cancellation of errors, while both families have the slope and intercept of correlation line very different from the theoretical ones. The regression parameters, along with mean absolute error (mae), are given in Tables 3–6. The mae values were calculated according to the formula

$$\text{mae} = |\Delta G_{\text{exp}} - \Delta G_{\text{calc}}| \quad (4)$$

It should be kept in mind that the experimental values also contain errors, so that calculated mae values are dependent on the quality of calculated numbers, as well as the experimental errors.

The computational cost of different methods was evaluated based on the computer processing unit (CPU) time needed for the calculation (geometry optimization and calculation of frequencies) of a neutral toluene molecule. These times are also reported in the Supporting Information. Times reported in Figures 1 and 2 are obtained without applying the resolution of identity (RI) approximation. Computational times are reported in the Supporting Information, along with RI approximation, and they show accelerations by factors of 1.2–2.3.

### 3. RESULTS AND DISCUSSION

Analysis of the results reveals that, for several compounds, our calculated values with the most used methods strongly deviated from those reported by NIST. The case of nitrobenzene, for

Table 4. Results of Correlation Analysis for Acidities (in kcal/mol)<sup>a</sup>

	<i>b</i>	<i>a</i>	<i>R</i> <sup>2</sup>	<i>S</i> <sub>0</sub>	<i>N</i>	$\Delta$
AM1	0.88 ± 0.04	46.24 ± 15.54	0.8870	10.9	53	9.3
PDDG	0.93 ± 0.07	34.14 ± 23.34	0.7922	14.7	52	13.2
PM3	0.92 ± 0.05	35.78 ± 18.37	0.8590	12.1	52	10.8
PM6	0.90 ± 0.04	48.17 ± 14.22	0.9039	10.0	52	13.3
CBS-QB3	0.98 ± 0.02	8.60 ± 6.26	0.9836	4.2	53	2.8
G2	0.96 ± 0.02	13.44 ± 7.09	0.9790	4.7	52	3.0
G2MP2	0.97 ± 0.02	10.64 ± 7.21	0.9787	4.8	52	2.9
G3	0.97 ± 0.02	10.93 ± 6.11	0.9846	4.1	52	2.8
G3MP2	0.98 ± 0.02	5.50 ± 6.08	0.9848	4.0	53	2.6
G4	0.98 ± 0.02	6.61 ± 6.37	0.9836	4.2	52	2.7
G4MP2	0.99 ± 0.02	4.61 ± 6.67	0.9819	4.4	53	2.7
B1B95	0.94 ± 0.02	20.65 ± 6.91	0.9797	4.6	51	3.1
B2PLYP	0.95 ± 0.02	19.99 ± 6.64	0.9804	4.6	53	3.3
B2PLYPD	0.95 ± 0.02	19.77 ± 6.52	0.9811	4.5	53	3.1
B3LYP	0.95 ± 0.02	21.67 ± 6.68	0.9799	4.6	53	3.7
B3PW91	0.95 ± 0.02	19.11 ± 6.52	0.9812	4.5	53	3.0
B97D	0.97 ± 0.02	10.39 ± 6.99	0.9795	4.7	53	2.9
B98	0.95 ± 0.02	18.20 ± 6.71	0.9802	4.6	53	3.1
BLYP	0.97 ± 0.02	17.87 ± 7.40	0.9760	5.1	53	6.9
BMK	0.92 ± 0.02	29.75 ± 6.41	0.9806	4.6	53	4.0
BP86	0.98 ± 0.02	12.99 ± 7.14	0.9783	4.8	53	5.9
CAM-B3LYP	0.93 ± 0.02	26.18 ± 6.51	0.9804	4.6	53	4.0
HSEh1PBE	0.95 ± 0.02	19.13 ± 6.47	0.9814	4.5	53	3.0
M06	0.96 ± 0.02	17.95 ± 6.93	0.9789	4.8	53	4.1
M062X	0.94 ± 0.02	24.03 ± 6.85	0.9787	4.8	53	4.2
M06HF	0.91 ± 0.02	37.24 ± 7.82	0.9701	5.7	53	6.6
M06L	0.96 ± 0.02	13.57 ± 6.41	0.9823	4.4	53	3.0
mPW2PLYP	0.94 ± 0.02	21.70 ± 6.56	0.9806	4.6	53	3.3
mPW2PLYPD	0.94 ± 0.02	21.40 ± 6.53	0.9809	4.5	53	3.2
O3LYP	0.96 ± 0.02	15.03 ± 6.84	0.9798	4.7	53	3.1
OLYP	0.97 ± 0.02	12.45 ± 7.31	0.9773	4.9	53	3.2
PBE1PBE	0.95 ± 0.02	18.70 ± 6.49	0.9814	4.5	53	3.0
PBEPBE	0.98 ± 0.02	14.11 ± 7.18	0.9779	4.9	53	6.1
tHCTHhyb	0.95 ± 0.02	17.40 ± 6.61	0.9808	4.5	53	3.1
TPSSH	0.94 ± 0.02	20.18 ± 6.55	0.9809	4.5	53	3.1
VSXC	0.97 ± 0.02	11.72 ± 6.96	0.9795	4.7	53	3.0
X3LYP	0.95 ± 0.02	21.89 ± 6.67	0.9800	4.6	53	4.0

<sup>a</sup>Legend: *a* is the intercept and *b* is the slope of the correlation line, *R*<sup>2</sup> is the square of correlation coefficient, *S*<sub>0</sub> is the standard deviation of the points from the correlation line, *N* is the number of points, and  $\Delta$  is the average absolute error. 6-311+G\*\* basis set is used with DFT methods.

which an alternative experimental value is reported, was already mentioned previously. Eight (8) bases out of 64 and 11 acids out of 53 differed by more than 5 kcal/mol, compared to the experiment. In the following discussion, the numbers shown in parentheses indicate the number of methods with mae > 5 kcal/mol, out of the 57 methods tested. These compounds were the following bases: 1,1,3,3-tetrafluoro-2-propanone (26), 1,2,3,5-tetrafluorobenzene (30), sulfonyl fluoride (33), 2-methylpropene (36), bis-(2,2,2-trifluoromethyl)ether (47), methyl phenyl sulfide (47), 1,2,4,5-tetrafluorobenzene (53), 2-methoxy-1-ethanol (55), and the following acids: 2-(trifluoromethyl)-1,1,1,3,3,3-hexafluoropropane (31), methanedisulfonyl difluoride (29), trifluoromethanesulfonic acid (33), fluorosulfuric acid (35), trimethylamine (40), 2,4,6-trinitrophenol (36), 1,1,1,5,5,5-hexafluoro-pentane-2,4-dione (42), chlorobenzene (50), hydrogen (52), methoxyethene (57), and propanenitrile (57). Out of 10 deviating acids, 5 belong to very strong acids, for which the experimental values in the NIST database have been questioned.<sup>80–82</sup> The reliability of the experimental gas-phase acidities and basicities of these molecules deserves further

examination. However, in further analysis, those molecules were taken into account. The correlation analysis results would be much better if they were excluded from analysis.

**3.1. Semiempirical Methods.** Among the semiempirical methods, AM1, PM3, PM6, and PDDG were used to calculate gas-phase acidities and basicities. The best results for basicities were obtained with the PM6 method, where the mae was 5.1 kcal/mol. The results of the regression analysis indicate that the description of basicities is close to theoretical—the slope of the correlation line is  $1.01 \pm 0.04$  and the intercept is  $-1.95 \pm 7.52$ . With other methods, the results were less accurate: AM1 gave a mae value of 5.7 kcal/mol; PDDG, mae = 11.1 kcal/mol; and PM3, 11.8 kcal/mol. In all cases, the slopes and intercepts of the correlation lines were close to their ideal values (1 and 0). The acidity of the hydrogen molecule was not considered in the analysis, because semiempirical methods cannot describe the relatively high negative charge density of the H<sup>+</sup> ion accurately and the mae values for acidities were in the range of 49.3–109.0 kcal/mol. The AM1 method was the best for estimating acidities, with mae = 9.3 kcal/mol. The mae values for other methods were

Table 5. Results of Correlation Analysis for Basicities (in kcal/mol)<sup>a</sup>

	<i>b</i>	<i>a</i>	<i>R</i> <sup>2</sup>	<i>S</i> <sub>0</sub>	<i>N</i>	$\Delta$
AM1	1.01 ± 0.04	0.36 ± 7.23	0.9151	7.8	64	5.7
PDDG	0.98 ± 0.06	12.55 ± 10.05	0.8295	11.0	64	11.1
PM3	1.02 ± 0.05	5.54 ± 9.43	0.8568	10.1	64	11.8
PM6	1.01 ± 0.04	−1.95 ± 7.52	0.9107	8.0	64	5.1
CBS-QB3	0.94 ± 0.02	12.60 ± 2.79	0.9844	3.3	64	2.8
G2	0.96 ± 0.01	7.77 ± 2.54	0.9878	2.9	64	1.8
G2MP2	0.96 ± 0.01	7.23 ± 2.53	0.9879	2.9	64	1.8
G3	0.95 ± 0.01	9.84 ± 2.73	0.9856	3.2	64	2.1
G3MP2	0.95 ± 0.01	10.83 ± 2.71	0.9856	3.2	64	2.2
G4	0.96 ± 0.02	8.36 ± 2.85	0.9845	3.3	64	2.0
G4MP2	0.97 ± 0.02	5.63 ± 2.84	0.9851	3.3	64	1.9
B1B95	0.95 ± 0.02	8.80 ± 2.86	0.9851	3.3	61	2.3
B2PLYP	0.95 ± 0.01	11.15 ± 2.65	0.9862	3.1	64	2.1
B2PLYPD	0.94 ± 0.01	11.19 ± 2.63	0.9863	3.1	64	2.1
B3LYP	0.95 ± 0.02	9.50 ± 2.98	0.9829	3.5	64	2.5
B3PW91	0.96 ± 0.02	6.45 ± 2.91	0.9842	3.3	64	3.0
B97D	0.97 ± 0.02	0.99 ± 3.18	0.9823	3.5	64	5.1
B98	0.96 ± 0.02	6.49 ± 2.91	0.9842	3.3	64	3.3
BLYP	0.97 ± 0.02	4.94 ± 3.34	0.9796	3.8	64	3.0
BMK	0.92 ± 0.02	15.33 ± 2.87	0.9830	3.5	64	2.6
BP86	0.99 ± 0.02	1.32 ± 3.13	0.9828	3.5	64	2.8
CAM-B3LYP	0.93 ± 0.02	13.99 ± 2.89	0.9831	3.5	64	2.7
HSEh1PBE	0.95 ± 0.02	8.05 ± 2.91	0.9840	3.4	64	2.5
M06	0.98 ± 0.02	4.62 ± 3.30	0.9801	3.8	64	2.6
M062X	0.93 ± 0.02	15.03 ± 3.00	0.9815	3.6	64	3.3
M06HF	0.84 ± 0.02	33.77 ± 3.96	0.9606	5.3	64	6.4
M06L	0.99 ± 0.02	0.11 ± 3.42	0.9798	3.8	64	3.5
mPW2PLYP	0.94 ± 0.01	12.23 ± 2.68	0.9857	3.2	64	2.2
mPW2PLYPD	0.94 ± 0.01	12.45 ± 2.74	0.9850	3.3	64	2.2
O3LYP	0.97 ± 0.02	3.25 ± 2.98	0.9840	3.4	64	3.9
OLYP	0.98 ± 0.02	−0.41 ± 3.22	0.9821	3.6	64	4.2
PBE1PBE	0.96 ± 0.02	7.23 ± 2.91	0.9841	3.4	64	2.5
PBEPBE	0.99 ± 0.02	1.50 ± 3.10	0.9830	3.5	64	2.7
tHCTHhyb	0.97 ± 0.02	4.47 ± 2.99	0.9837	3.4	64	3.3
TPSSH	0.96 ± 0.02	4.56 ± 3.01	0.9834	3.4	64	3.7
VSXC	0.98 ± 0.02	1.24 ± 3.51	0.9784	3.9	64	4.2
X3LYP	0.95 ± 0.02	9.96 ± 2.96	0.9830	3.5	64	2.4

<sup>a</sup>Legend: *a* is the intercept and *b* is the slope of the correlation line, *R*<sup>2</sup> is the square of correlation coefficient, *S*<sub>0</sub> is the standard deviation of the points from the correlation line, *N* is the number of points,  $\Delta$  is the average absolute error. 6-311+G\*\* basis set is used with DFT methods

10.8 kcal/mol (PM3), 13.2 kcal/mol (PM6), and 13.3 kcal/mol (PDDG). The correlation line parameters are not close to theoretical for the acids: the slope was  $0.88 \pm 0.04$  and the intercept is  $46.42 \pm 15.54$  for the AM1 method.

The best results, when considering calculated acidities and basicities together, were obtained with the AM1 method (mae = 7.3 kcal/mol). The regression line parameters of the overall correlations are ideal.

### 3.2. Gaussian-Type and Complete Basis Set Methods.

Among Gaussian-type methods G2, G2MP2, G3, G3MP2, G4 and G4MP2 were used to calculate gas-phase acidities and basicities. All these methods gave almost similar results. For the basicities, G2 and G2MP2 methods gave the best results with mae = 1.8 kcal/mol. With other methods, the mae was in the range of 1.9–2.2 kcal/mol. It was found that the G3MP2 method gave the best results for acidities (mae = 2.6 kcal/mol), whereas, for other methods, errors were in the range of 2.7 and 3.0 kcal/mol. When considering acidities and basicities together, similar results were obtained with four methods: G2, G2MP2, G4, G4MP2, where the mae was 2.3 kcal/mol. Gaussian-type

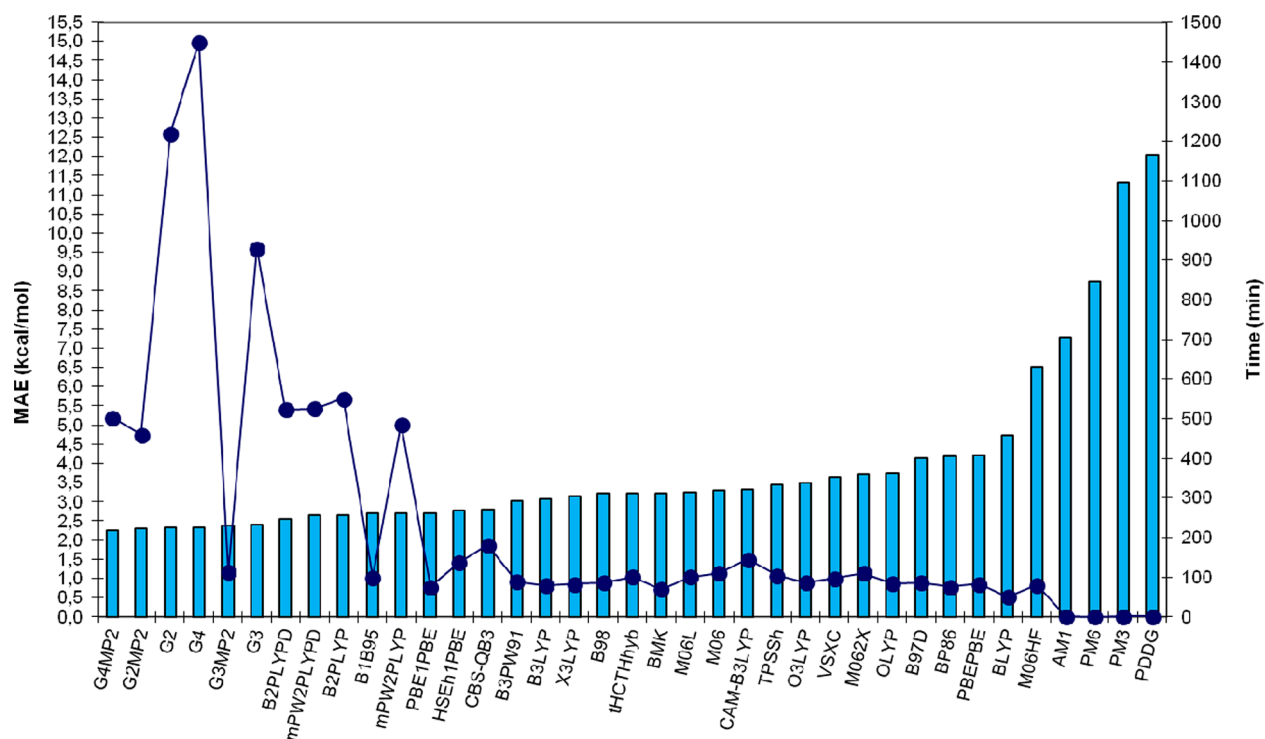
methods have, for the acidities, the slopes of the correlation line that are almost ideal in all cases (the intercepts are still remarkable  $-4.13$  kcal/mol), while, for basicities, the correlation line parameters are somewhat worse: slopes are smaller than 1 and intercepts are also significantly greater than 0. Gaussian-type methods are quite accurate, but their downside is the relatively long computational time (see Figure 1). It can be seen from Figure 1 that it is more reasonable to use GXMP2 methods than pure GX methods, because calculation times are shorter while the accuracies are similar. The errors are bigger in the current work compared to previous results,<sup>13,14</sup> probably due to the bigger dataset and the inclusion of strongly deviating values.

From the complete basis set methods, only one method was tested: CBS-QB3. This method evaluated acidities and basicities similarly, giving a mae value of 2.8 kcal/mol. The correlation line parameters for acidities are close to their theoretical values (slope =  $0.98 \pm 0.02$  and intercept =  $8.60 \pm 6.26$ ); for basicities, the parameters are not so good (slope =  $0.94 \pm 0.02$  and intercept =  $12.60 \pm 2.79$ ). Compared to Gaussian-type methods, the CBS-QB3 method is less accurate when calculating basicities, but

Table 6. Results of Correlation Analysis for Acidities and Basicities Together with Different Methods and Basis Sets<sup>a</sup>

method	basis set		<i>b</i>	<i>a</i>	<i>R</i> <sup>2</sup>	<i>S</i> <sub>0</sub>	<i>N</i>	$\Delta$
B2PLYPD	aug-cc-pVTZ	all	1.01 ± 0.00	−1.46 ± 1.02	0.9985	3.5	117	2.3
B2PLYPD	6-311+G**	all	1.00 ± 0.00	1.39 ± 1.19	0.9978	4.2	117	2.5
B3LYP	aug-cc-pVTZ	all	1.02 ± 0.00	−4.45 ± 1.08	0.9983	3.7	117	2.7
PBE1PBE	6-311+G**	all	1.00 ± 0.00	−0.89 ± 1.22	0.9978	4.2	117	2.7
PBE1PBE	aug-cc-pVTZ	all	1.01 ± 0.00	−4.13 ± 1.06	0.9984	3.6	117	2.8
B3LYP	6-311+G**	all	1.01 ± 0.00	−1.34 ± 1.31	0.9975	4.5	117	3.1
B3LYP	TZVP	all	0.99 ± 0.00	0.88 ± 1.26	0.9976	4.4	117	3.1
B2PLYPD	TZVP	all	0.98 ± 0.00	2.89 ± 1.18	0.9979	4.1	117	3.3
BLYP	TZVP	all	1.01 ± 0.00	−2.76 ± 1.36	0.9973	4.7	117	3.4
PBE1PBE	TZVP	all	0.99 ± 0.00	0.97 ± 1.21	0.9978	4.2	117	3.4
BLYP	cc-pVTZ	all	1.01 ± 0.00	−3.91 ± 1.32	0.9974	4.5	117	3.7
B2PLYPD	cc-pVTZ	all	0.98 ± 0.00	2.61 ± 1.13	0.9980	4.0	117	4.1
B97D	6-311+G**	all	1.03 ± 0.00	−9.55 ± 1.33	0.9975	4.4	117	4.1
BLYP	aug-cc-pVTZ	all	1.04 ± 0.00	−8.78 ± 1.12	0.9982	3.8	117	4.2
B3LYP	cc-pVTZ	all	0.99 ± 0.00	−0.03 ± 1.22	0.9978	4.2	117	4.4
B97D	aug-cc-pVTZ	all	1.04 ± 0.00	−12.79 ± 1.08	0.9984	3.6	117	4.4
BLYP	6-311+G**	all	1.03 ± 0.01	−5.45 ± 1.42	0.9971	4.8	117	4.7
PBE1PBE	cc-pVTZ	all	0.99 ± 0.00	−0.54 ± 1.21	0.9978	4.2	117	4.7
BLYP	6-311G**	all	0.99 ± 0.01	−1.87 ± 1.49	0.9967	5.1	117	4.8
B97D	TZVP	all	1.01 ± 0.00	−6.82 ± 1.30	0.9976	4.4	117	4.8
B3LYP	6-311G**	all	0.98 ± 0.01	1.48 ± 1.45	0.9968	5.0	117	5.3
PBE1PBE	6-311G**	all	0.97 ± 0.00	2.35 ± 1.35	0.9972	4.7	117	5.5
B2PLYPD	6-311G**	all	0.96 ± 0.00	4.92 ± 1.34	0.9972	4.7	117	5.7
B97D	cc-pVTZ	all	1.01 ± 0.00	−8.33 ± 1.27	0.9977	4.3	117	7.0
B97D	6-311G**	all	0.99 ± 0.01	−5.55 ± 1.46	0.9969	5.0	117	7.5

<sup>a</sup>Values determined in units of kcal/mol. Legend: *a* is the intercept and *b* is the slope of the correlation line, *R*<sup>2</sup> is the square of correlation coefficient, *S*<sub>0</sub> is the standard deviation of the points from the correlation line, *N* is the number of points, and  $\Delta$  is the average absolute error. The 6-311+G\*\* basis set is used with DFT methods.

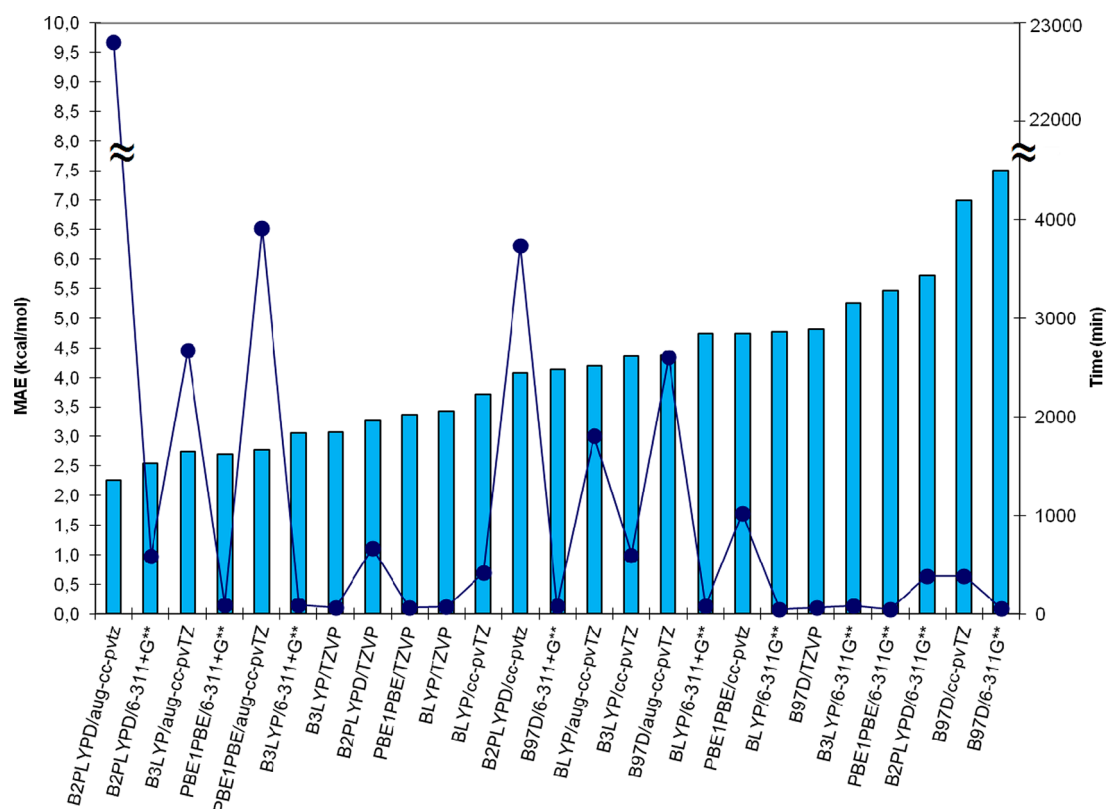


**Figure 1.** Mean absolute errors (maes) for both basicities and acidities (columns), and calculation time of neutral toluene molecule (line) with different methods.

results for acidities are as good as with Gaussian-type methods. In addition, the CBS-QB3 method is quite fast, compared to GX methods (except G3MP2, see Figure 1).

**3.3. Density Functional Theory Methods.** It was found that there was no single GGA method equally good at calculating both acidities and basicities. The methods that were good for





**Figure 2.** Mean absolute errors (maes) for basicities and acidities together (columns), and calculation time of neutral toluene molecule (line) with different basis sets.

calculating acidities were not so good for basicities and vice versa. PBEPBE was best for calculating basicities (mae = 2.7 kcal/mol). The next best methods then were BP86 (mae = 2.8 kcal/mol), BLYP (mae = 3.0 kcal/mol), OLYP (mae = 4.2 kcal/mol), and B97D (mae = 5.1 kcal/mol). For the bases, the correlation between experimental and calculated basicities was ideal in most of the cases (except BLYP and B97D). B97D was best for calculating acidities (mae = 2.9 kcal/mol), followed by OLYP (mae = 3.2 kcal/mol), BP86 (mae = 5.9 kcal/mol), PBEPBE (mae = 6.1 kcal/mol), and BLYP (mae = 6.9 kcal/mol). It could be seen with the latter methods that there is a systematic error (all acidities values are underestimated, compared to the experiment). The comparisons of regression line parameters indicate that the acidities do not have correlation lines as close to the theoretical one as bases: slopes are smaller than 1 and intercepts are significantly greater than 0. When considering both acidities and basicities, the OLYP method gave the best results (mae = 3.7 kcal/mol), followed by B97D (mae = 4.1 kcal/mol), PBEPBE (mae = 4.2 kcal/mol), BP86 (mae = 4.2 kcal/mol), and BLYP (mae = 4.7 kcal/mol). The correlation line parameters were again not close to their theoretical values for overall correlation. Because there is no universal method that is good for evaluating both acidities and basicities, different methods should be suggested for calculating basicities (PBEPBE) and acidities (B97D). It should be noted that, among DFT methods, GGA methods had correlation lines parameters that were closest to the ideal. Computational times are similar for all GGA methods. The fastest method was BP86, and the slowest method was B97D.

The results with two M-GGA methods were quite similar, M06L was more accurate. The mae value with the M06L method was 3.5 kcal/mol for basicities, and VSXC gave a mae value of 4.2 kcal/mol. The methods gave the same results when calculating

acidities (mae = 3.0 kcal/mol). When considering both acidities and basicities then the mae values were 3.2 and 3.6 kcal/mol, respectively, with M06L and VSXC. The given methods can predict acidities better than basicities. The correlation line parameters were also closer to the ideal values for acids than bases. Computational time was shorter for the VSXC method.

The mae values of the calculated basicities were in the range of 2.1–3.9 kcal/mol with the hybrid-GGA methods. The best results were obtained with the B2PLYP method (mae = 2.1 kcal/mol). The mPW2PLYP (mae = 2.2 kcal/mol) and X3LYP (mae = 2.4 kcal/mol) also gave good results. Among H-GGA methods, the least accurate was O3LYP (mae = 3.9 kcal/mol).

Three methods gave good results for acidities: PBE1PBE, B3PW91, and HSEh1PBE (mae = 3.0 kcal/mol), whereas the values with other methods were between 3.1 and 4.0 kcal/mol. When considering both acidities and basicities, the mae values were in the range of 2.6–3.5 kcal/mol. The accuracy decreases in the order B2PLYP (mae = 2.6 kcal/mol), mPW2PLYP (mae = 2.7 kcal/mol), PBE1PBE (mae = 2.7 kcal/mol), HSEh1PBE (mae = 2.8 kcal/mol), B3PW91 (mae = 3.0 kcal/mol), B3LYP (mae = 3.1 kcal/mol), X3LYP (mae = 3.1 kcal/mol), B98 (mae = 3.2 kcal/mol), CAM-B3LYP (mae = 3.3 kcal/mol), and O3LYP (mae = 3.5 kcal/mol). Some methods (O3LYP and B98) predict acidities better than basicities. The correlation line parameters for basicities ( $a = 0.93$ – $0.97$ ,  $b = 3$ – $14$ ) and acidities ( $a = 0.93$ – $0.96$ ,  $b = 15$ – $26$ ) are different from the ideal, so overall correlation lines parameters that are closer to their theoretical values can again be attributed to the favorable cancellation of errors.

B2PLYP and mPW2PLYP were compared with the B2PLYPD and mPW2PLYPD methods, respectively, to evaluate the influence of adding the Grimmes empirical dispersion correction

(D). It was found that adding this correction gave more-accurate results when considering acidities (B2PLYPD mae = 3.1 kcal/mol, B2PLYP mae = 3.3 kcal/mol; mPW2PLYPD mae = 3.2 kcal/mol, mPW2PLYP mae = 3.3 kcal/mol); for basicities, the results were the same. The computational time for the methods including empirical dispersion correction was only somewhat longer.

The computational time for H-GGA methods was variable as some of them include MP2 calculation. The fastest method was PBE1PBE and the slowest was B2PLYP.

Among HM-GGA methods, the most accurate for basicities was B1B95 (mae = 2.3 kcal/mol), and the least accurate was M06HF (mae = 6.4 kcal/mol). Mean absolute errors with other methods were in the range of 2.6–3.7 kcal/mol. The B1B95, tHCTHhyb, and TPSSh methods gave the best results for acidities (mae = 3.1 kcal/mol). B1B95 almost has ideal regression line parameters when considering both acidities and basicities. However, B1B95 had convergency failures with some compounds. The mae value with the latter method was 2.7 kcal/mol. Computational times for HM-GGA methods were quite short. Comparison of correlation line parameters indicates that the worst cases among DFT functionals found belong to this group (M06HF, BMK, M062X).

When looking at all of the results in the context of Jacob's ladder, it is evident that those methods that are higher on the ladder do not always give better results, compared to those methods that are lower on the ladder. Some functionals were found to be considerably better for describing acidities (e.g., B97D, M06L, O3LYP) or basicities (e.g., B1B95, B2PLYP, B3LYP). This probably stems from the ability of functional to describe relatively better electron density further away from atomic nuclei (the acidities are closer to experiment) or closer to nuclei (basicities are described better).

Among DFT methods, B2PLYPD was the most accurate for both basicities and acidities. However, it had long computational times, as did some other methods, which are the next in accuracy. The next best method was B1B95, but it had convergency failures with some compounds. The best method considering accuracy, robustness, and computation time was PBE1PBE.

**3.4. Basis Set Analysis.** It is well-known that the accuracy of calculated properties depends not only on the chosen DFT functional, but also on the basis set. The computational time is also basis-set dependent. So we have compared the performance of five triple- $\zeta$  quality basis sets on the accuracy of calculated acidities and basicities with 5 DFT functionals B3LYP, B97D, BLYP, B2PLYPD, and PBE1PBE. The chosen basis sets were 6-311G\*\*, 6-311+G\*\*, TZVP, cc-pVTZ, and aug-cc-pVTZ. Comparison of the results obtained with five DFT methods and five basis sets revealed that with different methods the accuracy order of basis sets was not the same. The mae values for the same basis set with different methods were calculated to identify basis set influence.

BLYP behaved quite differently, compared to other four DFT methods, with respect of the basis sets influence on the final accuracy. Four other (B3LYP, B97D, PBE1PBE, B2PLYPD) methods displayed similar dependence of the accuracy, with respect to the chosen basis set: the aug-cc-pVTZ and 6-311+G\*\* basis sets had the lowest errors, within 0.1–0.3 kcal/mol of each other, followed by TZVP, usually with bigger errors (by ~0.4–0.7 kcal/mol) (the only exception being the B3LYP functional, where TZVP gave results as good as those observed with the 6-311+G\*\* basis set), and cc-pVTZ and 6-311G\*\* with bigger errors. The BLYP has quite a different dependence on the basis

set as the best results were unexpectedly obtained with the TZVP method, closely followed by cc-pVTZ, and the rest of the studied basis sets gave already worse results. It is also worth pointing out that the influence of the basis set was weakest with the BLYP method.

When considering both acidities and basicities, it was found that the best basis set is aug-cc-pVTZ, where the mean absolute error was 3.3 kcal/mol. It was followed by 6-311+G\*\* (3.4 kcal/mol), TZVP (3.6 kcal/mol), cc-pVTZ (4.8 kcal/mol), and 6-311G\*\* (5.7 kcal/mol).

For basicities, the mae value was lowest with 6-311+G\*\* (3.0 kcal/mol). It was followed by TZVP (3.2 kcal/mol), aug-cc-pVTZ (3.4 kcal/mol), cc-pVTZ (4.2 kcal/mol), and 6-311G\*\* (4.4 kcal/mol).

It is usually believed that the diffuse functions present in the aug-cc-pVTZ and 6-311+G\*\* basis sets are especially important for calculating acidities (for correct description of anions). Indeed, for four methods, except BLYP, those basis sets are among the best, whereas for BLYP they are the worst.

For acidities, the mean absolute error for the same basis set with different methods was lowest with aug-cc-pVTZ (mae = 3.1 kcal/mol), followed by 6-311+G\*\* (mae = 3.9 kcal/mol), TZVP (mae = 4.0 kcal/mol), cc-pVTZ (mae = 5.5 kcal/mol), and 6-311G\*\* (mae = 7.3 kcal/mol) basis sets.

One important aspect when choosing a basis set is computational time. From the Supporting Information, it can be seen that the longest computational times correspond to the aug-cc-pVTZ basis set independent of the used method. The computational times for TZVP, 6-311G\*\*, and 6-311+G\*\* basis sets were considerably shorter.

It can be seen that TZVP, 6-311+G\*\*, and aug-cc-pVTZ gave similar results with all five DFT methods. Comparing those three basis sets, it was found that aug-cc-pVTZ is most accurate but has the longest computational time, while the other two give similar results with a much shorter computational time.

Considering computational times and accuracy, 6-311+G\*\* or TZVP should be the optimal basis set for calculating acidities and basicities, the first being somewhat more accurate with the above-suggested PBE1PBE functional, and the second somewhat faster.

**3.5. Application of the Scaling Factors to the Frequencies.** The calculated quantum chemical harmonic vibrational frequencies are typically larger than those observed experimentally, mainly because of neglect of anharmonicity effects in the theoretical treatment, but also because of incomplete incorporation of electron correlation and the use of finite basis sets. The relatively uniform nature of the overestimation of frequencies for a particular theoretical procedure allows the application of frequency scale factors to correct the discrepancy between experiment and calculations. The downside of such an approach is that the scaling factor for each method (e.g., each combination of a basis set and DFT functional, and perhaps also the grid used) needs to be determined. Moreover, different scaling factors seem to be necessary for correct calculations of different frequency-related quantities (fundamentals, low-frequency vibrational frequencies, zero-point vibrational energies, and thermal contributions to enthalpy and entropy).<sup>83</sup>

It seems reasonable to believe that the acidities and basicities can also be estimated more correctly using those correction factors as they contain contributions from zero-point vibrational energies, and vibrational contributions to the thermal enthalpies and entropies. We have tested this assumption for the B3LYP/6-311+G\*\* calculations using scaling factors from Radom et al.<sup>83</sup> The comparison of acidities and basicities calculated with scaled

and unscaled vibrations (see Tables S14–S16 in the Supporting Information) reveals that the differences are minor (typically 0.1 kcal/mol) and the correlation between experimental and calculated values does not change. Thus it can be concluded that for the calculation of acidities and basicities the application of scaled frequencies is not necessary.

#### 4. CONCLUSIONS

Our study shows that, among semiempirical methods, the AM1 method had the best ability to reproduce experimental acidities and basicities. The PM6 method was best for predicting basicities, and the AM1 method was best for acidities. Semiempirical methods predicted basicities better than acidities. These methods are very fast and suitable when one needs a quick estimate of gas-phase acidities and basicities (the AM1 method gave a mean absolute error of mae = 7.3 kcal/mol).

Four Gaussian-type methods (G2, G2MP2, G4, G4MP2) gave similar results to each other (the mae = 2.3 kcal/mol). Gaussian-type methods are quite accurate, but their downside is their relatively long computational time that is needed. It is more reasonable to use the GXMP2 method than the pure GX method, because calculation times are much shorter but the accuracies are similar. G2 and G2MP2 were the best methods for predicting basicities and G3MP2 for acidities. The CBS-type method CBS-QB3 predicted acidities and basicities with similar accuracy, but compared to Gaussian-type methods, the CBS-QB3 was less accurate for basicities. The CBS-QB3 method was quite fast, compared to most Gaussian-type methods.

Density functional theory (DFT) methods were categorized according to the Jacob's ladder into four groups: GGA, M-GGA, H-GGA, HM-GGA. Among DFT methods, B2PLYPD/6-311+G\*\* was the most accurate for basicities and acidities. However, it had a long computational time, as did some other methods, which are the following in accuracy. B1B95/6-311+G\*\* gave also a low mae value, but it had convergence failures with some compounds. The best method considering accuracy, robustness, and computation time was PBE1PBE/6-311+G\*\*.

Application of Grimmes empirical dispersion correction D gave somewhat better results than methods without it for acidities.

Calculations were also carried out with five basis sets (6-311G\*\*, 6-311+G\*\*, TZVP, cc-pVTZ, aug-cc-pVTZ) and with the B3LYP, B97D, BLYP, B2PLYPD, and PBE1PBE methods to evaluate the effect of the basis on the accuracy of calculations. It was found that the best basis sets when considering the accuracy of results and time needed were 6-311+G\*\* and TZVP.

The use of scaled DFT frequencies for calculating zero-point vibrational energies, and thermal contributions to enthalpy and entropy seems unnecessary for getting correct acidities and basicities.

In some cases, we found that calculated basicity or acidity values deviated strongly from experimentally determined ones from the NIST database.

#### ■ ASSOCIATED CONTENT

##### Supporting Information

Results of correlation analysis and calculation times are given in the Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### ■ AUTHOR INFORMATION

##### Corresponding Author

\*E-mail: [peeter.burk@ut.ee](mailto:peeter.burk@ut.ee).

##### Notes

The authors declare no competing financial interest.

#### ■ ACKNOWLEDGMENTS

This work was supported by the Estonian Science Foundation (Grant No. 8809) and the Estonian Ministry of Education and Research Targeted Financing Project No. SF0180120s08. This study was in part carried out using High Performance Computing Center of the University of Tartu. I.A.K. acknowledges the financial support from UT Centre of Excellence HIGH-TECHMAT, project SLOKT117T.

#### ■ REFERENCES

- (1) NIST Chemistry Webbook, NIST Standard Reference Database No. 69, <http://webbook.nist.gov> (accessed May 14, 2013).
- (2) Koppel, I. A.; Burk, P.; Koppel, I.; Leito, I.; Sonoda, T.; Mishima, M. *J. Am. Chem. Soc.* **2000**, *122*, 5114–5124.
- (3) Burk, P.; Sillar, K.; Koppel, I. A. *J. Mol. Struct.: THEOCHEM* **2001**, *543*, 223–231.
- (4) Koppel, I. A.; Schwesinger, R.; Breuer, T.; Burk, P.; Herodes, K.; Koppel, I.; Leito, I.; Mishima, M. *J. Phys. Chem. A* **2001**, *105*, 9575–9586.
- (5) Koppel, I. A.; Burk, P.; Koppel, I.; Leito, I. *J. Am. Chem. Soc.* **2002**, *124*, 5594–5600.
- (6) Vessecchi, R.; Galembeck, S. E. *J. Phys. Chem. A* **2008**, *112*, 4060–4066.
- (7) Leito, I.; Koppel, I. A.; Burk, P.; Tamp, S.; Kutsar, M.; Mishima, M.; Abboud, J.-L. M.; Davalos, J. Z.; Herrero, R.; Notario, R. *J. Phys. Chem. A* **2010**, *114*, 10694–10699.
- (8) Bras, N. F.; Perez, M. A. S.; Fernandes, P. A.; Silva, P. J.; Ramos, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 3898–3908.
- (9) Pople, J. A.; Headgordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622–5629.
- (10) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1990**, *93*, 2537–2545.
- (11) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221–7230.
- (12) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 1293–1298.
- (13) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.
- (14) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1999**, *110*, 4703–4709.
- (15) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 084108.
- (16) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *127*, 124105.
- (17) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822–2827.
- (18) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532–6542.
- (19) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757–10816.
- (20) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings*, Antwerp, Belgium, June 8–10, 2000; Doren, V. E., Alsenoy, K., Geerlings, P., Eds.; 2001; 577, 1–20.
- (21) Dewar, M. J. S.; Dieter, K. M. *J. Am. Chem. Soc.* **1986**, *108*, 8075–8086.
- (22) Burk, P.; Koppel, I. A. *Theor. Chim. Acta* **1993**, *86*, 417–427.
- (23) Burk, P.; Herodes, K.; Koppel, I. *Int. J. Quantum Chem.* **1993**, *48*, 633–641.
- (24) Smith, B. J.; Radom, L. *Chem. Phys. Lett.* **1995**, *245*, 123–128.
- (25) Smith, B. J.; Radom, L. *J. Phys. Chem.* **1995**, *99*, 6468–6471.



- (26) Merrill, G. N.; Kass, S. R. *J. Phys. Chem.* **1996**, *100*, 17465–17471.
- (27) Burk, P.; Koppel, I. A.; Koppel, I.; Leito, I.; Travníkova, O. *Chem. Phys. Lett.* **2000**, *323*, 482–489.
- (28) Vayner, E.; Ball, D. W. *J. Mol. Struct.: THEOCHEM.* **2000**, *496*, 175–183.
- (29) Aue, D. H.; Guidoni, M.; Betowski, L. D. *Int. J. Mass Spectrom.* **2000**, *201*, 283–295.
- (30) Pokon, E. K.; Liptak, M. D.; Feldgus, S.; Shields, G. C. *J. Phys. Chem. A* **2001**, *105*, 10483–10487.
- (31) Li, Z.; Matus, M. H.; Velazquez, H. A.; Dixon, D. A.; Cassady, C. J. *Int. J. Mass Spectrom.* **2007**, *265*, 213–223.
- (32) Range, K.; Lopez, C. S.; Moser, A.; York, D. M. *J. Phys. Chem. A* **2006**, *110*, 791–797.
- (33) He, X.; Fusti-Molnar, L.; Merz, K. M., Jr. *J. Phys. Chem. A* **2009**, *113*, 10096–10103.
- (34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.02*; Gaussian, Inc.: Wallingford, CT, 2009.
- (35) Kaljurand, I.; Koppel, I. A.; Kutt, A.; Room, E.-I.; Rodima, T.; Koppel, I.; Mishima, M.; Leito, I. *J. Phys. Chem. A* **2007**, *111*, 1245–1250.
- (36) Cheng, X.; Grabowski, J. J. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 34–36.
- (37) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (38) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (39) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221–264.
- (40) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (41) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601–1622.
- (42) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (44) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (45) Miehlisch, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200–206.
- (46) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (47) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403–412.
- (48) Hoe, W. M.; Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2001**, *341*, 319–328.
- (49) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (50) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396.
- (51) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
- (52) Van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998**, *109*, 400–410.
- (53) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
- (54) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (55) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (56) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.
- (57) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
- (58) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (59) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207–8215.
- (60) Heyd, J.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *120*, 7274–7280.
- (61) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 104103.
- (62) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 224106.
- (63) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
- (64) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (65) Xu, X.; Goddard, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2673–2677.
- (66) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040–1046.
- (67) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–3416.
- (68) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (69) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 5121–5129.
- (70) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.
- (71) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559–9569.
- (72) Tao, M. J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (73) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (74) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
- (75) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. W. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (76) Schafer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
- (77) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (78) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (79) Koppel, I. A.; Molder, U. H.; Palm, V. A. *Org. React. (Tartu)* **1985**, *22*, 3–45.
- (80) Gutowski, K. E.; Dixon, D. A. *J. Phys. Chem. A* **2006**, *110*, 12044–12054.
- (81) Koppel, I. A.; Taft, R. W.; Anvia, F.; Zhu, S. Z.; Hu, L. Q.; Sung, K. S.; Desmarteau, D. D.; Yagupolskii, L. M.; Yagupolskii, Y. L.; Ignatev, N. V.; Kondratenko, N. V.; Volkonskii, A. Y.; Vlasov, V. M.; Notario, R.; Maria, P. C. *J. Am. Chem. Soc.* **1994**, *116*, 3047–3057.
- (82) Leito, I.; Raamat, E.; Kutt, A.; Saame, J.; Kipper, K.; Koppel, I. A.; Koppel, I.; Zhang, M.; Mishima, M.; Yagupolskii, L. M.; Garlyauskayte, R. Y.; Filatov, A. A. *J. Phys. Chem. A* **2009**, *113*, 8421–8424.
- (83) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683–11700.