

Critical test of PM3 calculated gas-phase acidities

Peeter Burk and Ilmar A. Koppel

Department of Analytical Chemistry, Tartu University EE2400, Jakobi 2, Tartu, Estonia

Received September 21, 1992/Accepted February 11, 1993

Summary. Gas-phase acidities have been calculated for 175 compounds using the PM3 semiempirical molecular orbital model. With some exceptions, PM3 seems to be a useful tool for the investigation of gas-phase acidities. The main problems encountered involve two rather different classes of acids: one which generates small anions (e.g., halide ions, hydride ion, etc.), in which the charge is localized on one atom, and, a second, represented by anions that contain bulky electron acceptor substituents characterized by an extensive negative charge delocalization. In some cases (anilines, amides, alcohols, and phenols) the average error in predicted gas-phase acidity can be significantly reduced by employing an empirically derived correction.

Comparison with AM1 results shows that both methods are of roughly equal quality with the exception of hypervalent molecules where PM3 is better (averaged unsigned errors are 11.8 and 17.0 kcal/mol for PM3 and AM1, respectively).

Key words: Gas-phase acidities – PM3 – AM1

1 Introduction

In recent years quantum chemical calculations have become a widely used tool for the investigation of a wide range of problems [1]. Both *ab initio* and semiempirical self-consistent field molecular orbital calculations are often used. The widespread application of the semiempirical MINDO/3 [2], MNDO [3], AM1 [4], and PM3 [5] methods is due to the fact that they often give reliable answers, are easy to use, require relatively little computer time, and can handle large molecules.

Proton transfer reactions play a basic role in chemistry and in biochemistry. As a result, besides the gas phase basicities or proton affinities of neutral bases, also numerous calculations of gas phase acidities of neutral Bronsted acids have appeared [1, 6, 7 and references therein]. These computations require the heats of formation of the acid and the conjugated base, the latter being an anion for uncharged acids. It is known [1, 8], that *ab initio* methods require the addition of diffuse orbitals in order to obtain reliable energies for anions. This requirement

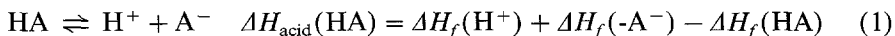
further limits the size of a molecule that can be calculated in a reasonable time. Semiempirical methods, at the same time, are reported to reproduce the energies of anions almost as well as for neutral molecules [4, 6, 9–13]. Known exceptions are small anions with charge predominantly localized on a single atom, such as CH_3^- and OH^- where the calculated energies are consistently too positive. This error is probably [14] due to failure to allow for orbital expansion accompanying large localized negative charge.

However, there have been only a few systematic investigations [1, 6, 7, 11, 15–17] testing the reliability of different methods for prediction of gas-phase acidities and none dealing with the PM3 method. Therefore the current investigation was undertaken to evaluate the reliability of the PM3 method for predicting gas-phase acidities.

A useful way to consider agreement between experiment and theory is to calculate the linear regression between the theoretical and experimental quantities and to calculate the mean difference (unsigned average error) between the experimental and theoretical quantities. The degree of agreement between the two is then reflected by the slope and intercept of the correlation line, the unsigned average error, the standard deviation from the correlation line and the correlation coefficient. The last two depend on the scatter of the points about the correlation line. A *slope different from unity* and a *non-zero intercept* imply systematic deviation between experiment and theory and means that the theoretical values of relative acidities will be systematically either too high or too low. A non-zero mean difference between theory and experiment reflects an overall bias in the absolute values of the calculated acidity.

2 Method

The acidity (ΔH_{acid}) of compound HA was found as the heat of reaction for the proton abstraction equilibrium to form the conjugate base A^- :



where $\Delta H_f(\text{H}^+)$, $\Delta H_f(\text{HA})$, and $\Delta H_f(\text{A}^-)$ are the heats of formation for proton, acid, and its conjugated anion. In case of proton the experimental heat of formation (367.2 kcal/mol [19]) was used instead of the calculated value (353.6), because it improves the calculated acidities considerably. We acknowledge that this is not quite a correct procedure, despite the fact that such a way was also used in earlier works on semiempirical calculations of acidities and basicities [6, 11].

The calculations were carried out on a MicroVAX II computer using the standard PM3 procedure, as implemented in the MOPAC 6.0 program package [20]. All geometries were fully optimized by minimizing the energy with respect to all geometrical variables without using the aid of symmetry. In order to avoid premature completion of calculations in local minimum, the calculations were carried out starting off from several different initial geometries.

3 Results and discussion

Acidities for 175 compounds, calculated using the PM3 method, are given in Table 1, along with corresponding experimentally measured values when avail-

Table 1. Comparison with experiment of PM3 heats of formation and acidities (kcal/mol)

	calc ΔH_f		ΔH_{acid}			exp ΔH_f (HA)	error in calc ΔH_f (HA)
	HA	A ⁻	calc	exp ^a	error		
C-H acids							
Methane	-13.0	51.5	430.2	416.6	13.6	-17.8	4.8
Ethane	-18.1	31.7	415.5	421.0	-5.5	-20.1	2.0
Propane	-23.6	14.8	404.1	419.0	-14.9	-25.0	1.4
Cyclopropane	16.3	60.2	409.7	412.0	-2.3	12.7	3.6
<i>t</i> -Butane	-24.4	0.0	390.1	414.0	-23.9	-32.1	7.7
Ethene	16.6	61.7	410.8	406.0	4.8	12.5	4.1
Propene	6.4	25.2	384.5	387.5	-3.0	4.8	1.6
2-Methylpropene	-2.8	16.1	384.6	390.3	-5.7	-4.0	1.2
2-Methyl-1,3-butadiene	22.8	39.0	381.9	385.7	-3.8	17.9	4.9
Cyclopentadiene	31.8	15.9	349.8	353.9	-4.1	31.0	0.8
Ethyne	50.7	76.4	391.4	376.7	14.7	54.5	-3.8
Propyne	40.2	64.2	389.7	381.1	8.6	44.6	-4.4
1-Pentyne	40.2	54.5	380.0	379.8	0.2	34.4	5.8
<i>t</i> -Bu-acetylene	31.8	47.4	381.3	378.0	3.3	25.0	6.8
Phenylacetylene	74.7	86.8	377.8	370.7	7.1	73.1	1.5
Benzene	23.5	51.7	393.9	400.7	-6.8	19.8	3.7
Toluene	14.1	18.5	370.1	380.8	-10.7	12.0	2.1
1,4-Dimethylbenzene	4.8	8.7	369.6	381.9	-12.3	4.3	0.5
Ethylbenzene	9.5	8.5	364.7	379.8	-15.1	6.9	2.5
<i>i</i> -Propylbenzene	4.9	-1.2	359.6	379.0	-19.4	1.0	3.9
Cycloheptatriene	46.4	43.1	362.3	375.2	-12.9	43.7	2.7
Fluorene	49.0	27.7	344.4	353.3	-8.9	44.9	4.1
9-Ph-fluorene	79.6	43.6	329.7	343.3	-13.6	76.0 ^b	3.6
Fluoradene	10.0	72.9	328.7	332.5	-3.8		
Diphenylmethane	43.0	26.8	349.5	363.6	-14.1	33.5	9.6
1,4-DiPh-cyclo-pentadiene	78.4	42.5	329.8	337.7	-7.9	77.0 ^b	1.4
CH ₃ Br	-5.0	27.2	397.9	392.7	5.2	-9.1	4.1
CH ₃ Cl	-14.7	20.7	401.1	396.0	5.1	-19.6	4.9
CH ₂ Cl ₂	-17.1	-7.9	374.9	374.6	0.3	-22.9	5.8
CHCl ₃	-20.9	-35.2	351.4	357.1	-5.7	-25.0	4.1
ClCH ₂ -acetylene	38.2	50.1	377.6	368.1	9.5	39.0 ^b	-0.8
Cl-benzene	16.7	35.8	384.8	387.2	-2.4	13.0	3.7
Nitromethane	-15.9	-43.3	338.3	356.4	-18.1	-17.9	2.0
Nitroethane	-19.7	-53.3	332.1	356.0	-23.9	-24.4	4.7
2-Nitropropane	-27.1	-61.8	331.0	356.1	-25.1	-33.2	6.1
Nitromethyl- <i>t</i> -butane	-51.1	-68.2	348.6	355.2	-6.6	-45.2 ^b	-6.0
<i>m</i> NO ₂ -toluene	5.1	-7.7	352.9	362.9	-10.0	7.0	-1.9
HCN	33.0	27.7	360.5	351.2	9.3	32.3	0.7
Cyanomethane	23.3	28.5	370.9	372.9	-2.0	18.0	5.3
Cyanoethane	18.5	17.1	364.2	375.0	-10.8	12.3	6.3
2-Cyanopropane	13.4	6.1	358.4	375.2	-16.8	5.8	7.6
Cyanocyclopropane	52.2	53.7	367.2	375.4	-8.2	44.0	8.2
2-Cyanopropene	41.0	42.6	367.4	370.7	-3.3	31.0	10.0
<i>p</i> CN-toluene	49.0	34.7	351.4	360.8	-9.4		
Cyanotoluene	43.2	28.6	351.1	351.9	-0.8		
PhCH(CN) ₂	94.7	47.5	318.5	322.1	-3.6		
FCH ₂ CN	-13.1	-20.5	358.3	369.2	-10.9		
Metoxycyanomethane	-10.2	-13.7	362.1	371.8	-9.7	-8.0 ^b	-2.2
<i>m</i> CF ₃ -C ₆ H ₄ CH ₂ CN	-105.3	-138.2	332.8	341.9	-9.1	-115.9 ^b	10.6

Table 1. (continued)

	calc ΔH_f		ΔH_{acid}			exp ΔH_f (HA)	error in calc ΔH_f (HA)
	HA	A ⁻	calc	exp ^a	error		
<i>p</i> COCN-toluene	19.7	-2.6	343.4	345.7	-2.3	3.6 ^b	16.2
Methyl ether	-46.9	-5.0	407.7	407.0	0.7	-44.0	-2.9
Ethanal	-44.2	-39.9	370.0	365.8	4.2	-39.6	-4.6
Propanal	-47.1	-51.4	361.5	365.3	-3.8	-44.8	-2.3
Propanone	-53.0	-48.6	370.0	369.1	0.9	-51.9	-1.1
Phenylmethylketone	-17.7	-20.2	363.2	361.3	1.9	-20.8	3.1
Benzylmethylketone	-23.0	-43.7	345.0	351.8	-6.8	-23.4	0.4
Ethylphenylketone	-22.8	-20.6	358.9	360.5	-1.6	-26.1	3.3
MeCO-acetylene	9.4	15.9	372.2	360.1	12.1	15.6	-6.2
MeOCO-acetylene	-29.6	-26.9	368.4	358.8	9.6	-24.0 ^b	-5.6
4-Methoxyphenylmethylketone	-54.8	-58.5	362.0	362.8	-0.8	-58.3	3.5
Methyl acetate	-92.8	-92.8	365.7	371.9	-6.2	-98.0	5.2
Dimethyl ethane-amide	-52.2	-49.0	368.9	374.9	-6.0	-55.9	3.7
Me ₂ S	-10.4	2.3	378.4	393.2	-14.8	-9.0	-1.4
PhSCH ₃	23.9	25.0	366.8	381.6	-14.8	23.4	0.5
MeSOM _e	-10.4	2.3	378.4	373.5	4.9	-36.2	25.8
MeSO ₂ Me	-76.2	-90.6	351.3	365.8	-14.5	-89.0	12.8
<i>p</i> SO ₂ Me-toluene	-51.0	-71.7	345.0	358.8	-13.8	-65.2	14.2
CF ₃ SO ₂ Me	-207.7	-250.5	322.9	347.0	-24.1		
CF ₃ SO ₂ CH = CHMe	-192.8	-239.0	319.5	343.3	-23.8	-222.0 ^b	29.2
<i>p</i> SO ₂ CF ₃ -toluene	-182.8	-217.7	330.8	347.4	-16.6		
Pyridine	30.4	48.2	383.5	391.9	-8.4	33.0	-2.6
Pyrimidine	38.0	43.6	371.4	383.3	-12.4	47.0	-9.0
Me ₃ P	-29.8	-4.3	391.2	391.3	-0.1	-24.1	-5.7
CH ₂ = Se(Me)CH ₃	3.5	29.8	392.0	385.6	6.4	5.0	-1.5
SiMe ₄	-53.5	-18.7	400.5	397.0	3.5	-55.7	2.2
			N-H acids				
Ammonia	-3.1	38.3	407.1	403.6	3.5	-11.0	7.9
Methylamine	-5.2	38.3	407.1	403.6	3.5	-11.0	7.9
Methylamine	-5.2	21.7	392.6	403.3	-10.7	-5.5	0.3
Ethylamine	-11.1	17.0	393.8	399.3	-5.5	-11.3	0.2
Dimethylamine	-8.0	7.8	381.5	396.2	-14.7	-4.4	-3.6
Me ₃ SiNH ₂	-57.9	-41.6	382.0	378.8	3.2		
Aniline	21.3	14.3	358.7	366.4	-7.7	20.8	0.5
<i>p</i> CH ₃ -aniline	12.1	4.7	358.3	367.3	-9.0	14.1 ^b	-2.1
<i>m</i> Cl-aniline	14.4	1.5	352.8	358.9	-6.1	13.0 ^b	1.4
<i>p</i> Cl-aniline	14.4	1.2	352.5	360.4	-7.9	13.0 ^b	1.4
<i>m</i> OCH ₃ -aniline	15.2	5.2	355.7	366.8	-11.1	12.9 ^b	2.3
<i>p</i> OCH ₃ -aniline	-15.4	-25.8	355.3	367.1	-11.8	-14.1 ^b	-1.3
<i>p</i> CHO-aniline	-13.6	-34.8	344.5	349.6	-5.1	-8.1 ^b	-5.4
<i>m</i> NO ₂ -aniline	12.5	-10.4	342.7	352.3	-9.6	15.0	-2.4
<i>p</i> NO ₂ -aniline	10.7	-24.1	330.9	343.5	-12.6	13.0	-2.3
<i>m</i> SO ₂ CF ₃ -aniline	-174.9	-202.7	337.9	346.9	-9.0		
<i>p</i> SO ₂ CF ₃ -aniline	-178.4	-209.5	336.1	338.6	-2.5		
Pyrrrole	27.1	11.5	350.0	358.7	-8.7	25.9	1.2
HCONH ₂	-41.8	-50.5	357.0	359.9	-2.9	-44.0	2.1
PhCONH ₂	-16.1	-28.7	353.2	354.1	-0.9	-24.0	7.9
H ₂ NCONH ₂	-45.8	-53.8	357.7	362.6	-4.9	-58.8	13.0
CF ₃ CONH ₂	-196.2	-223.8	338.1	343.8	-5.7	-200.0 ^b	3.8
Cl ₃ C CONH ₂	-55.7	-81.1	340.3	343.2	-2.9	-56.2 ^b	0.4

Table 1. (continued)

	calc ΔH_f		ΔH_{acid}			exp ΔH_f (HA)	error in calc ΔH_f (HA)
	HA	A ⁻	calc	exp ^a	error		
HCONHMe	-43.7	-53.0	356.3	360.4	-4.1	-45.0 ^b	1.3
Acetaminobenzene	-22.0	-47.2	340.4	347.5	-7.2	-30.8	8.9
CF ₃ CONHPh	-165.2	-209.4	321.4	333.4	-12.0	-168.7 ^b	3.6
O-H acids							
Water	-43.4	-17.5	401.6	390.7	10.9	-57.8	4.4
Methanol	-51.1	-37.9	378.8	380.6	-1.8	-48.2	-2.9
Ethanol	-58.7	-44.8	379.7	378.0	1.7	-56.1	-2.6
Propanol	-64.1	-51.4	378.3	376.0	2.3	-60.9	-3.2
Isopropanol	-65.8	-54.9	376.6	375.4	1.2	-65.1	-0.7
Isobutanol	-67.7	-56.9	376.5	374.7	1.8	-67.8	0.1
<i>t</i> -Butanol	-65.7	-54.9	376.6	374.6	2.0	-74.7	9.0
<i>t</i> -Bu-methanol	-75.5	-58.8	382.3	372.6	9.7	-76.0 ^b	0.5
1- <i>t</i> -Bu-ethanol	-79.2	-64.2	380.7	371.1	9.6	-83.9 ^b	4.7
1- <i>t</i> -Bu-propanol	-81.2	-74.0	372.8	370.0	2.8	-88.7 ^b	7.5
1- <i>t</i> -Bu-isobutanol	-89.6	-76.2	379.0	368.5	10.5	-93.7 ^b	4.1
1- <i>t</i> -Bu-methanol	-88.2	-78.9	375.0	366.4	8.6	-99.2 ^b	11.0
Phenylmethanol	-22.4	-18.6	369.5	370.0	-0.5	-23.9	1.5
2-Methoxyethanol	-92.7	-83.6	374.7	373.8	0.9	-87.0 ^b	-5.7
F ₂ CHCH ₂ OH	-151.8	-155.6	361.9	366.4	-4.5	-148.2 ^b	-3.6
CF ₃ CH(Me)OH	-214.4	-231.5	348.5	360.3	-11.8	-216.3 ^b	1.9
(CF ₃) ₂ CHOH	-360.6	-398.1	328.2	344.9	-16.7	-367.0 ^b	6.4
CF ₃ CH ₂ CH ₂ OH	-305.6	-325.7	345.5	355.4	-9.9	-313.9	7.4
(CF ₃) ₃ COH	-516.6	-572.6	309.7	331.6	-21.9	-549.0 ^b	32.4
Et ₃ SiOH	-133.2	-125.8	373.1	360.5	12.6	-133.6 ^b	0.4
Phenol	-21.7	-44.1	343.3	349.2	-5.9	-23.0	1.3
<i>p</i> Me-phenol	-30.9	-53.9	342.8	351.6	-8.8	-29.9	-1.0
<i>m</i> Et-phenol	-35.6	-56.5	344.8	349.3	-4.5	-34.9	-0.7
<i>m</i> F-phenol	-65.4	-94.5	336.5	344.5	-8.0	-71.0 ^b	5.6
<i>p</i> CF ₃ -phenol	-180.6	-219.1	327.2	337.2	-10.0	-182.8 ^b	2.2
<i>p</i> CN-phenol	13.1	-25.8	326.9	332.2	-5.3	10.3 ^b	2.8
<i>m</i> NO ₂ -phenol	-29.9	-68.7	326.8	334.6	-7.8	-27.0 ^b	-2.9
<i>p</i> NO ₂ -phenol	-31.7	-81.5	315.8	327.8	-12.0	-28.0 ^b	-3.7
<i>m</i> COMe-phenol	-63.0	-91.5	337.2	342.5	-5.3	-78.5	15.7
<i>m</i> SO ₂ Me-phenol	-85.6	-121.5	329.9	336.1	-6.2	-105.9 ^b	20.3
<i>p</i> SO ₂ CF ₃ -phenol	-219.6	-275.7	309.6	322.6	-13.0		
CH ₃ CH=NOH	-2.9	-14.9	353.7	365.6	-11.9	-4.8	1.9
C ₆ H ₅ CH=NOH	30.6	3.4	338.5	352.9	-14.4	25.8 ^b	4.8
(CH ₃) ₃ CCH=NOH	-17.5	-31.1	352.1	364.6	-12.5	-32.3 ^b	14.8
(CH ₃) ₂ C=NOH	-11.6	-24.9	352.5	366.1	-13.6	-15.0 ^b	3.4
Et ₂ NOH	-27.7	-14.8	378.6	370.6	8.0	-8.6 ^b	-19.1
Formic acid	-90.1	-110.9	344.9	345.2	-0.3	-90.5	0.4
Acetic acid	-99.3	-119.6	345.3	348.5	-3.2	-103.3	4.0
Propanoic acid	-104.3	-124.9	345.1	347.3	-2.2	-107.0	2.7
<i>t</i> -BuCOOH	-113.8	-129.6	349.9	344.6	5.3	-122.0 ^b	8.2
MeOCH ₂ COOH	-133.3	-154.3	344.7	341.6	3.1	-132.9 ^b	-0.4
FCH ₂ COOH	-139.7	-165.8	339.6	338.8	0.8	-140.1 ^b	0.4
CHF ₂ COOH	-186.6	-222.8	329.5	330.0	-0.5	-196.9 ^b	10.3
CF ₃ COOH	-244.0	-290.4	319.3	322.7	-3.4	-246.3	2.3
CF ₃ CH ₂ COOH	-254.9	-287.2	333.4	334.8	-1.4	-259.3 ^b	4.4
ClCH ₂ COOH	-101.3	-127.4	339.6	336.3	3.3	-104.0	2.7

Table I. (continued)

	calc ΔH_f		ΔH_{acid}			exp ΔH_f (HA)	error in calc ΔH_f (HA)
	HA	A ⁻	calc	exp ^a	error		
CHCl ₂ COOH	-102.4	-135.1	333.0	327.3	5.7	-102.1 ^b	-0.4
Cl(CH ₂) ₃ COOH	-109.0	-135.0	339.7	345.4	-5.7	-119.7 ^b	10.8
BrCH ₂ COOH	-89.6	-116.4	338.9	334.0	4.9	-94.4 ^b	4.8
MeCHBrCOOH	-99.0	-125.2	339.5	336.2	3.3	-103.3 ^b	4.2
EtCHBrCOOH	-103.9	-130.4	339.2	336.4	2.8	-114.0 ^b	10.1
Benzoic acid	-66.2	-89.7	342.2	338.3	3.9	-70.3	4.1
pCH ₃ -benzoic acid	-223.1	-254.9	333.9	332.4	1.5	-233.3 ^b	10.2
3,5-DiCF ₃ -benzoic acid	-378.7	-416.8	327.7	324.4	3.3	-391.3 ^b	12.5
pCN-benzoic acid	-30.0	-61.4	334.3	328.5	5.8	-37.8 ^b	7.8
mNO ₂ -benzoic acid	-66.4	-97.1	335.0	329.5	5.5	-94.3	27.9
pNH ₂ -benzoic acid	-69.2	-91.5	343.4	341.1	2.3	-70.0	0.8
mOH-benzoic acid	-110.9	-136.3	340.3	337.9	2.4	-112.3 ^b	1.4
S-H acids							
H ₂ S	-0.9	-15.9	350.8	351.2	-0.4	-4.9	4.0
MeSH	-5.5	-22.1	349.1	356.9	-7.8	-5.5	0.0
EtSH	-8.4	-25.2	348.9	355.2	-6.3	-11.1	2.7
<i>n</i> -PrSH	-13.7	-31.1	348.3	354.2	-5.9	-16.2	2.5
<i>i</i> -PrSH	-14.4	-28.1	352.0	353.4	-1.4	-18.2	3.8
<i>n</i> -BuSH	-24.5	-42.3	347.9	353.7	-5.8	-21.1	-3.4
<i>i</i> -BuSH	-16.0	-33.5	348.2	353.1	-4.9	-23.3	7.3
<i>t</i> -BuSH	-17.8	-35.7	347.8	353.0	-5.2	-26.2	8.4
<i>t</i> -BuCH ₂ SH	-24.7	-43.1	347.3	351.7	-4.4	-30.8	6.2
Other acids							
H ₂	-13.2	91.8	470.9	400.4	70.5	0.0 ^c	-13.4
Me ₃ SiH	-37.2	-50.7	352.2	381.2	-29.0	-39.0	1.8
SiH ₄	13.1	-2.8	349.8	372.3	-22.5	8.0	5.1
PH ₃	0.2	-14.5	351.0	370.9	-19.9	1.3	-1.1
HF	-59.7	-17.1	409.8	371.4	38.4	-65.1	5.4
HCl	-20.5	-51.2	336.5	333.7	2.8	-22.1	1.6
HBr	5.3	-56.2	305.7	322.4	-16.7	-8.7	14.0
HI	28.8	-64.6	273.8	314.3	-40.5	6.3	22.5

^a - from Ref. [19]^b - estimated in Ref. [19]^c - by definition

able. These particular compounds were chosen because they cover a wide range of acidities for different classes of compounds (C-H, O-H, N-H, S-H acids; alcohols, amines, anilines, phenols, hydrocarbons, etc., fluoro-, nitro-, cyano-, etc. substituted acids, etc.). In Table 2 the analogous results, calculated with AM1 method, for 11 hypervalent compounds are presented.

The results of statistical (regression) analysis according to the formula:

$$\Delta H_{\text{acid}}(\text{calc}) = a + b \cdot \Delta H_{\text{acid}}(\text{exp}) \quad (2)$$

are presented in Table 3, where N is the number of points, $\Delta H_{\text{acid}}^{\text{min}}$ is minimal and $\Delta H_{\text{acid}}^{\text{max}}$ maximal experimental gas-phase acidity in considered group; $\delta \Delta H_{\text{acid}}$ is the average unsigned error in acidities, R is the correlation coefficient, σ is the

Table 2. Comparison with experiment of AM1 heats of formation and acidities (kcal/mol) for hypervalent compounds

	calc ΔH_f		ΔH_{acid}			exp ΔH_f (HA)	error in calc ΔH_f (HA)
	HA	A ⁻	calc	exp ^a	error		
<i>m</i> SO ₂ Me-phenol	-82.9	-115.9	334.2	336.1	-1.9	-105.9	23.0
<i>p</i> SO ₂ Me-toluene	-46.6	-72.8	341.0	358.8	-17.8	-65.2	18.7
MeSO ₂ Me	-70.0	-94.0	343.2	365.8	-22.6	-89.0	19.0
MeSOMe	-8.7	8.6	384.5	373.5	11.0	-36.2	27.5
<i>p</i> SO ₂ CF ₃ -phenol	-222.4	-282.7	307.0	322.6	-15.6		
CF ₃ SO ₂ CH=CHMe	-193.8	-251.8	309.2	343.3	-34.1	-222.0	28.2
<i>m</i> SO ₂ CF ₃ -aniline	-177.1	-202.1	342.2	346.9	-4.7		
CF ₃ SO ₂ Me	-204.3	-262.2	309.3	347.0	-37.7		
<i>p</i> SO ₂ CF ₃ -toluene	-184.7	-277.5	324.3	347.4	-23.1		
<i>p</i> SO ₂ CF ₃ -aniline	-186.5	-213.6	340.1	338.6	1.5		

^a - from Ref. [19]**Table 3.** Results of statistical analysis of gas-phase acidities, calculated with PM3 method

	<i>N</i>	$\Delta H_{\text{acid}}^{\text{min}}$	$\Delta H_{\text{acid}}^{\text{max}}$	$\delta \Delta H_{\text{acid}}$	<i>a</i>	<i>b</i>	<i>R</i>	σ	γ
All	175	314.3	421.0	8.2	-30.4	1.1	0.903	11.2	7.3
All ^a	165	314.3	421.0	7.5	-4.8	1.0	0.937	7.9	6.3
All ^b	80	338.3	421.0	7.6	-8.6	1.0	0.900	9.2	7.2
C-H acids	75	322.1	421.0	8.8	-33.8	1.0	0.923	9.4	7.0
C-H acids ^b	46	351.2	421.0	8.2	-8.4	1.0	0.891	10.2	7.6
N-H acids	25	333.4	403.6	7.2	-14.3	1.0	0.975	4.7	3.6
N-H acids ^b	10	347.6	403.6	7.8	-28.9	1.0	0.977	5.0	3.1
Amines	5	378.8	403.6	7.5	110.6	0.7	0.691	8.7	7.7
Amides	8	333.4	362.6	5.1	-80.6	1.2	0.979	2.8	1.7
Anilines	11	338.6	367.3	8.4	30.0	0.8	0.959	2.9	2.5
O-H acids	58	322.6	390.7	6.2	-36.2	1.0	0.934	7.5	5.6
Alcohols	29	322.6	380.6	6.8	-113.9	1.3	0.970	5.7	3.1
Carboxylic acids	22	322.7	348.5	3.2	53.3	0.8	0.897	3.2	2.9
Nitro-compounds	9	327.8	362.9	13.2	123.5	0.6	0.717	8.4	11.1
Diatomic hydrides	5	314.3	400.4	34.1	-410.3	2.2	0.995	8.5	2.4
Hypervalent comp.	10	322.6	373.5	11.8	-65.8	1.1	0.895	9.1	6.3
Trifluoromethyl-sulfonyl compounds	6	322.6	347.4	14.8	71.0	0.7	0.662	9.1	8.5

^a - without Si-H and P-H acids, H₂, C₂H₂, CH₄, H₂O, HF, and HCN^b - compounds, calculated by Dewar [6] with AM1 and used by us to compare AM1 and PM3

standard deviation, and γ is the average unsigned error, when ΔH_{acid} is corrected using formula:

$$\Delta H_{\text{calc.}}^{\text{cor.}} = \frac{\Delta H_{\text{calc.}} - a}{b} \quad (3)$$

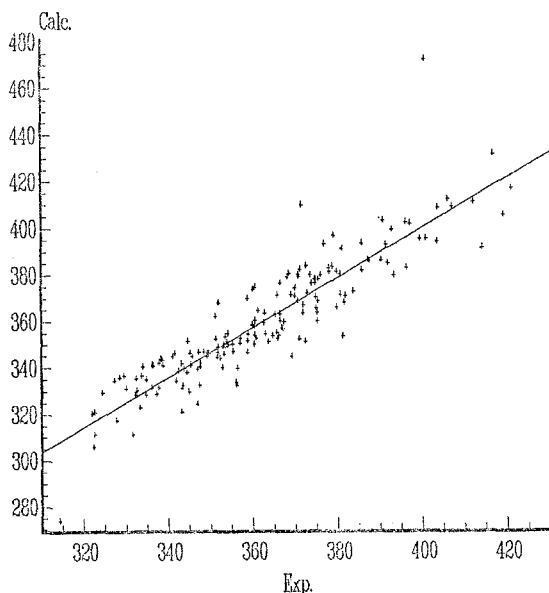
Table 4 gives the analogous results for AM1 calculated acidities from Ref. [6], [11], and from this work.

Table 4. Results of statistical analysis of gas-phase acidities, calculated with AM1 method

	<i>N</i>	$\Delta H_{\text{acid}}^{\text{min}}$	$\Delta H_{\text{acid}}^{\text{max}}$	$\delta \Delta H_{\text{acid}}$	<i>a</i>	<i>b</i>	<i>R</i>	σ	γ
All	97	338.3	421.0	7.9	-28.7	1.1	0.894	9.6	7.2
All ¹	80	338.3	421.0	8.2	-31.6	1.1	0.892	10.1	7.4
C-H acids	63	350.3	421.0	7.7	-25.1	1.1	0.894	9.5	6.8
C-H acids ^a	46	351.2	421.0	8.1	-19.0	1.0	0.890	10.4	7.4
N-H acids	10	352.7	403.2	6.7	-136.5	1.4	0.975	6.4	2.9
O-H acids	24	338.3	390.8	9.0	-87.3	1.2	0.860	9.6	5.5
Hypervalent compounds	10	322.6	373.5	17.0	-68.9	1.1	0.740	16.5	11.5

^a - compounds, calculated by Dewar [6] with AM1 and used by us to compare AM1 and PM3

The average unsigned error in the heats of formation for all 175 neutrals is 5.1 kcal/mol. This error is fairly typical for the PM3 method and is in keeping with what has previously been found [18]. Knowing that the heats of formation of the anions are reproduced more poorly, it is not surprising, that the calculated absolute acidities are not very reliable. At the same time, the average unsigned error in acidities of the same collection of acids is 8.1 kcal/mol, compared with 7.9 kcal/mol for AM1 method obtained by us (Table 4) from analysis of results of calculations for 97 compounds presented in Ref. [6] and [11]. But when we compare the results only for compounds calculated by both methods, the errors are 8.2 kcal/mol and 7.6 kcal/mol for AM1 and PM3 respectively. *It should be mentioned that in case of PM3 nearly one third of calculated acidities have errors greater than 10 kcal/mol and more than half deviate more than 5 kcal/mol (see also Fig. 1).*

**Fig. 1.** Calculated and experimental gas-phase acidities (kcal/mol) for all acids surveyed

From Tables 1 and 2 it can be seen, that the behavior of the five calculated diatomic hydrides (H_2 , HF, HCl, HBr, HI) is quite different from that for the rest of the compounds. For these molecules the average error is as big as 34.1 kcal/mol, but when we use empirical correction according Eq. (3), the error reduces to 2.4 kcal/mol. It should also be mentioned, that in case of these compounds the slope of the correlation line is 2.21 and intercept is -410.4 which both display the largest deviations of these quantities from their ideal values, i.e. the unity slope and zero intercept.

For a few calculated Si-H and P-H acids and for acetylene, methane, water and HCN, the calculated values are also far from their experimental values. When these compounds are excluded from the statistical analysis, the average unsigned error reduces to 7.3 kcal/mol. The poor results for these acids can be attributed to the failure of methods from the MNDO family to allow for orbital expansion on atoms bearing large negative charge. The calculated heats of formation are expected to be, and are, too positive whereas the formal charge in an anion is largely concentrated on a single atom.

On the other hand, it should be mentioned that the PM3 tends to overestimate the charge delocalization in a relatively bulky anions, such as $(CH_3)_3C^-$, $CF_3SO_2CH_2^-$, $(CF_3)_3CO^-$, etc., which results in too negative heats of formation. In turn this leads to too high acidities for these acids. This can be also the reason for which the PM3 method and also AM1, as mentioned by Dewar [6], mostly overestimate the stabilizing effect of methyl and phenyl substitution at the anionic center.

On the basis of the results of the overall correlation analysis, one can conclude that the errors in calculated acidities are not systematic, while the slope b of the correlation line is practically unity and the intercept a is also reasonably small. But such an inference turns out to be false, while examining the results of correlation analyses obtained for different classes of compounds (compare with Ref. [7], where the theoretical values for the slope b and intercept a were obtained for the simultaneous comparison of the calculated and experimental acidities of neutral and cationic Bronsted acids). It becomes evident that in many cases (amides, alcohols, and phenols) there are excellent correlations between calculated and experimental acidities with the intercept and slope considerably differing from theoretical (see Table 3). So, it can be concluded that overall good agreement with theoretical regression with zero intercept and unity slope is due to using in the same correlation several classes of compounds. However, in different classes, correlated one by one, the differences from the theoretical correlation are large enough to decide that there are still remaining systematic errors.

The average unsigned error in calculated acidities for the *carbon acids* is 8.8 kcal/mol, which is comparable with that found by us for the AM1 method (7.7 kcal/mol). But when we compare the results only for compounds calculated by both methods, the errors are 8.1 kcal/mol and 8.2 kcal/mol for AM1 and PM3, respectively.

The average unsigned error for acidities of 26 *nitrogen acids* is 7.2 kcal/mol. That is slightly better than the overall average unsigned error and average unsigned error for C-H acids and worse than that found by us for the AM1 method (6.7 kcal/mol). When one compares the results only for compounds calculated by both methods, the errors are 6.7 kcal/mol and 7.8 kcal/mol for AM1 and PM3, respectively.

The reasonable closeness of the slope of correlation line and the intercept to the theoretical values is due to fact that in the same regression are amines,

anilines, and amides. If we consider these classes of compounds separately, a different view appears. Amines have an average unsigned error of 7.5 kcal/mol and a rather scattered correlation. At the same time, amides have an excellent fit with a slope and intercept different from expected ideal values. In case of anilines, *p*-CF₃SO₂-aniline deviates considerably. After the exclusion of this point a very good correlation with slope and intercept close to their theoretical values appears (see Table 3).

For the *oxygen acids* listed in Table 1 the average unsigned error in acidities is 6.2 kcal/mol, compared with 9.0 kcal/mol found by us using AM1. When we compare the results only for compounds calculated by both methods, the errors didn't change. The reasons for big errors in calculated acidities of water and (CF₃)₃COH have already been discussed. The errors for oximes are also significantly larger than those for the other compounds. Excluding water, (CF₃)₃COH and oximes from the statistical analysis, the average unsigned error for the remaining 51 O-H acid falls to 5.2 kcal/mol. This value compares favorably with the corresponding error (4.4 kcal/mol) in the heats of formation of the corresponding neutrals and is only slightly larger than that of N-H acids and smaller than C-H acids. This is a bit surprising, as Dewar has reported that for the AM1 method the acidities of the oxygen acids are reproduced worse than C-H and N-H acids.

As in the case of N-H acids, the different classes of O-H acids show different correlations between calculated and experimentally measured acidities. The correlation for the carboxylic acids is rather poor, while in case of phenols and alcohols the fit is significantly better (see Table 2).

The average unsigned error of nine calculated S-H acids is 4.7 kcal/mol. However, no conclusions about effectiveness of PM3 to estimate the acidities of S-H acids can be made since a very narrow (351.2–356.9 kcal/mol) range of experimentally measured acidities was available.

For the nitrosubstituted acids given in Table 1, the average unsigned error is 13.2 kcal/mol, which is too big compared with the overall average unsigned error. For the AM1 model Dewar has attributed this mainly to the failure of method to predict the heats of formation rather neutrals than anions, which is supported by his calculation results. For the PM3 method, however, for nitrocompounds, Stewart had reported a great improvement of calculated heats of formation, which was also confirmed by us in this work (average unsigned error is 6.4 kcal/mol). However, the errors in the calculated acidities of nitrosubstituted acids remain too big. It must be pointed out that all the calculated acidities for nitrocompounds are too high and the bigger errors are associated mostly with the small compounds.

For the calculated acidities of six trifluoromethylsulfonylsubstituted acids the average unsigned error is 14.8 kcal/mol, which is too large when compared with the overall error. At the same time we should not forget that we deal here with hypervalent compounds for which the reported [18] error of calculated heats of formation of neutrals is 13.6 kcal/mol.

4 Conclusions

With a few exceptions, PM3 seems to be a useful tool for the investigation of gas-phase acidities. The errors in calculated acidities (8.1 kcal/mol) are comparable with those reported [18] in calculated heats of formation of the neutrals

(7.8 kcal/mol). However, it should be mentioned that in case of PM3 nearly one third of calculated acidities have errors greater than 10 kcal/mol and more than half deviate more than 5 kcal/mol.

The main problems encountered involve relatively small, often diatomic, neutral hydrides whose anions are usually characterized by the charge localization on one atom, as well as acids whose conjugated anions contain bulky electron acceptor substituents, capable to the extensive delocalization of the negative charge from the protonization center to the other regions of the anions. Probably for the same reason the large errors accompany also the introduction of methyl and phenyl substituents at anionic centers.

In some cases (anilines, amides, alcohols, and phenols) the average error in acidity can be significantly reduced by employing an empirically derived correction.

Comparison with AM1 results show that both methods are of roughly equal quality with the exception of hypervalent molecules, where PM3 is clearly better (average unsigned errors are 17.0 and 11.8 kcal/mol for AM1 and PM3, respectively). Shortcomings of both MNDO family-based methods for the correct description of hypervalent compounds are rather evident and another approach with proper inclusion of *d*-orbitals [21] is needed.

References

1. Hehre WJ, Radom L, Schlyer, PRv, Pople JA (1986) *Ab Initio* molecular orbital theory. Wiley, NY
2. Bingham RC, Dewar MJS, Loo DH (1975) *J Am Chem Soc* 97:1285
3. Dewar MJS, Thiel W (1977) *J Am Chem Soc* 99:4899
4. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Am Chem Soc* 107:3902
5. Stewart JJP (1989) *J Comput Chem* 10:209
6. Dewar MJS, Dieter KM (1986) *J Am Chem Soc* 108:8075 and references therein
7. Koppel IA, Mölder UH, Palm VA (1985) *Org React (Tartu)* 21:3
8. Kollmar H (1978) *J Am Chem Soc* 100:2665
9. Bartmess JE, Burnham RD (1984) *J Org Chem* 49:1382
10. Bartmess JE, Caldwell G, Rozeboom D (1983) *J Am Chem Soc* 105:340
11. Kass SR (1990) *J Comp Chem* 11:94
12. Dewar MJS, Fox MA, Nelson DJ (1980) *J Organomet Chem* 185:157
13. Gordon MS, Davis LP, Burggaf LF, Damrauer M (1986) *J Am Chem Soc* 108:7889
14. Olivella S, Urpi F, Vilarrasa J (1984) *J Comput Chem* 5:230
15. Siggel MR, Thomas TD, Saethre LJ (1988) *J Am Chem Soc* 110:91
16. Koppel IA, Comisarow MB (1980) *Org React (Tartu)* 17:498
17. Koppel IA, Mölder UH (1981) *Org React (Tartu)* 18:42
18. Stewart JJP (1989) *J Comput Chem* 10:221
19. Lias SG, Bartmess JE, Liebman JF, Holmes JL, Levin RD, Mallard WG (1988) *J Phys Chem Ref Data* 17, Suppl No. 1
20. Stewart JJP (1983) MOPAC Program Package – QPCE 455
21. Thiel W, Voityuk AA (1992) *Theor Chim Acta* 81:391