

# A theoretical calculation of relative pKas of phosphonic acids in an aqueous solution using idodesmic reaction and the Polarisable Continuum Model

Rachid Bouaddi<sup>1</sup>, Lahcen El Maimouni<sup>1</sup> And Abdellatif El Hammadi<sup>2</sup>

<sup>1</sup>Equipe de Matériaux et Physico Chimie de l'Atmosphère et du Climat (MPAC),

<sup>2</sup>Equipe Physicochimie des milieux naturels, Matériaux et Environnement (PCMNE), Laboratoire de Biotechnologie, Matériaux et Environnement (LBME), faculty of science, University Ibn Zohr, B.P 8106 Cité Dakhla Agadir, Morocco.

Corresponding Author; Rachid Bouaddi

**ABSTRACT:** The geometries of the neutral and deprotonated structures of 24 phosphonic acids  $R-PO_3H_2$  are optimized at the HF and B3LYP level using 6-31<sup>++</sup>G<sup>\*\*</sup> and LANL2DZ basis set and SMD solvation model (R=H, CH<sub>3</sub>, C<sub>2</sub>H<sub>5</sub>, n-C<sub>3</sub>H<sub>7</sub>, n-C<sub>4</sub>H<sub>9</sub>, n-C<sub>6</sub>H<sub>13</sub>, iso-C<sub>3</sub>H<sub>7</sub>, iso-C<sub>4</sub>H<sub>9</sub>, s-C<sub>4</sub>H<sub>9</sub>, t-C<sub>4</sub>H<sub>9</sub>, CH<sub>2</sub>CH(CH<sub>3</sub>)<sub>3</sub>, C(CH<sub>3</sub>)<sub>2</sub>C<sub>2</sub>H<sub>5</sub>, CH<sub>2</sub>C<sub>6</sub>H<sub>5</sub>, CH=CH<sub>2</sub>, CH=CHC<sub>6</sub>H<sub>5</sub>, C<sub>6</sub>H<sub>5</sub>, CH<sub>2</sub>OH, CF<sub>3</sub>, CH<sub>2</sub>Cl, CHCl<sub>2</sub>, CCl<sub>3</sub>, CH<sub>2</sub>Br, CH<sub>2</sub>I, and CH<sub>2</sub>Si(CH<sub>3</sub>)<sub>3</sub>). Relative  $pka_1$  and  $pka_2$  of these systems are then calculated at the same level by isodesmic reaction in aqueous phase. The calculated values of the pKas are compared to the corresponding experimental values. The average error of pKas varies between  $-1.62 \pm 1.92$  and  $-0.01 \pm 1.28$  pKa unit. The correlations are linear and quite high between calculated pKas ( $0.92 \leq R^2 \leq 0.99$ ) and calculated pKas versus experimental ones ( $0.73 \leq R^2 \leq 0.91$ ). In addition, the experimental  $pka_1$  values of phosphonic acids are scaled by comparison with the pKa of corresponding carboxylic acids. The scaled  $pka_1$  values made it possible to improve the correlation between the experiment and the theory ( $0.91 \leq R^2 \leq 0.97$ ). This study shows that the calculations of the relative dissociation constant pKa of phosphonic acid derivatives via isodesmic reaction at the HF and B3LYP level using the bases 6-31<sup>++</sup>G<sup>\*\*</sup> and LANL2DZ with the solvation model SMD conduct to results with a precision of better quality.

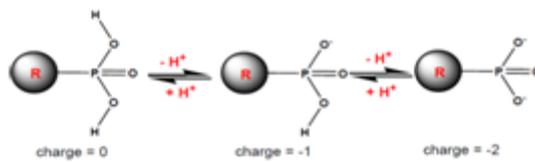
**KEYWORDS:** Deprotonation, isodesmic reaction, pKa dissociation constant, SMD solvation model, Phosphonic Acid

Date of Submission: 23-02-2019

Date of acceptance: 14-03-2019

## I. INTRODUCTION

Phosphonic acids constitute a special class of molecules within phosphorus-containing compounds. The tetrahedral P atom is connected via a P=O double bond to O, while it forms two P-O single bonds with OH groups. The P atom also has a single bond with carbon (P-C), the latter originating from an aliphatic or aromatic fragment R. The phosphonic acid group exhibits variable acidity.



**Figure 1:** chemical formula and deprotonation processes in the phosphonic acids.

There are two stepwise deprotonation processes, as shown in Figure 1. It should be noted that these deprotonation reactions are feebly studied. Only a few determinations of dissociation constants of some phosphonic acids are recorded in literature (P. Nylen, 1930)(P. Rumpf1947) (J. R. White, 1950)(Crofts, P. C., 1953). The purpose of the latest investigation of Crofts et al. was the study of the effects of variations within the alkyl group of the acids on the apparent dissociation constants of these acids.



$$pk_{ai} = -\log Ka_i = \frac{\Delta G_{i(aq)}^*}{2.303RT} \quad [3]$$

$\Delta G_{i(aq)}^*$  is the Gibbs free energy of the dissociation reaction, R is the ideal gases constant ( $R=1.987 \cdot 10^{-3} \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ ) and T is the absolute temperature in Kelvin ( $T=298.15\text{K}$ ).

The computation of acid dissociation constants is a demanding and arduous process because an error of 1.36 kcal/mol in the change of free energy of reaction 1a or 1b results in an error of 1 pKa unit (Liptak, 2001), (Liptak, 2002).

Therefore, to optimize errors, the  $pK_{ai}$  relative acidity constants are evaluated with respect to that of a reference system  $BH_2$  according to the isodesmic reaction pathway (4). The reference system  $BH_2$  chosen and used here is the phosphonic acid  $H-PO_3H_2$ . The choice is based on its chemical structure, which is similar to the studied compounds and the fact offers the possibility of quantifying the substituent effects on the acidity constants

## 2.2 Computational details

### 2.2.1 The calculation levels

All calculations have been carried out on 24 phosphonic acid compounds using the Gaussian 16 software (M. J. Frisch, 2016), at Hartree-Fock (HF), density functional theory (DFT) using Becke's three parameter (B3) exchange and the Lee–Yang–Parr (LYP) correlation functional (Becke AD, 1993). (Lee C, 1988). (Miehlich B, 1989) and AM1 level (M. J. S. Dewar, 1977) (M. J. S. Dewar, 1985).

The 6-31++G(d,p) Pople-type basis sets (P. C. Hariharan, 1974) (P. C. Hariharan, 1973), and LanL2DZ pseudopotential basis sets (Hay, 1985) (T. H. Dunning Jr, 1976) (W. R. Wadt, 1985) (P. J. Hay, 1985) integrated in Gaussian 16 were used. The use of polarization and diffuse OAs is justified by the presence of negative charges and intramolecular hydrogen bonding in deprotonated systems. The solvation energies were obtained at the same theory levels using the SMD model (V. Barone, 1998) (M. Cossi, 2003) (Marenich, 2009).

### 2.2.2 Geometry optimization

The HF and DFT calculations were performed using the gas phase starting geometries obtained with the AM1 semi empirical method (Dewar, 1977) (Dewar, 1985). Thereafter, the scanning of the potential energy surface was carried out for the  $(OH)_2P=O$  and the CX group together by the variation of each dihedral angle D1, D2 and D3 of  $0^\circ$  to  $360^\circ$  with a step of  $10^\circ$  Figure 3.

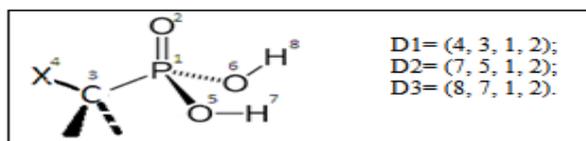
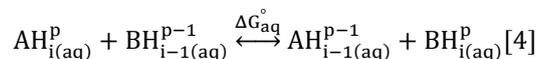


Figure 3: Dihedral angles involved in the scanning of a phosphonic acid potential energy surface.

### 2.2.3 Gibbs free energy calculation

The proton exchange reaction is a process in which the proton liberated by the acid  $AH_{i(aq)}^p$  is fixed by the reference base  $BH_{i-1(aq)}^{p-1}$  according to equations [4]:



Where p and p-1 represent the global charge of the acid species and its conjugate base respectively for  $i=2$  first dissociation and  $p=0$  and  $i=1$ ,  $p=-1$  second dissociation.

The  $\Delta G_{i(aq)}^*$  free energy reaction is determined according to Equation (5):

$$\Delta G_{i(aq)}^* = G_{aq}^*(AH_{i-1(aq)}^{p-1}) + G_{aq}^*(BH_{i(aq)}^p) - G_{aq}^*(AH_{i(aq)}^p) - G_{aq}^*(BH_{i-1(aq)}^{p-1}) \quad [5]$$

The geometrical optimization and the calculation of the frequencies of the reactants and products in solution allow determining the standard free energy  $G_{aq}^*(X_{(aq)})$  at 298.15K.

### III. RESULTATS AND DISCUSSION

The calculated acid dissociation constant pKa values for 24 phosphonic acids obtained as described above are compared with the experimental values identified in literature table 1 (appendix). Also, the average value of pKa calculated at different levels is carried with the standard deviation for comparison. The obtained pKa values are usually in the order of magnitude of the corresponding experimental values except 18, 19, 20, and 21 compounds. The latter have significant pKa<sub>1</sub> negative values at the HF and B3LYP levels. Table 2 (appendix) shows the errors of calculated pKa values relative to experimental one and table 3 illustrates the statistical parameters concerning some errors. The SMD-HF and SMD-B3LYP methods overestimate or underestimate the pKa values of the phosphonic acids studied. The maximum positive error is 2.50 pKa, while the negative error is quite high at -9.45 pKa. These two values are obtained with the method of B3LYP / LANL2DZ. The mean signed error MSE values are between -0.01±1.28 and -1.62±1.92 pKa unit. The values of these means become equal to -0.19 ±1.39 and -0.90 ±1.69 pKa units respectively without the 18, 20 and 21 compounds which have a large negative error. The mean unsigned error MUE values are 0.51±0.76 and 1.69±1.69 pKa units respectively. The MSE values of ΔpKa<sub>1</sub>ΔpKa<sub>2</sub> (mean value) are respectively -0.84±2.41 and -0.67±1.57 and strongly

**Table 3.** Statistical parameters: Mean Signed Errors (MSE), Mean Unsigned Errors (MUE) and Standard Deviation (SD) in pKa's of phosphonic acids calculated at SMD-HF and SMD-B3LYP level (N=21 for 6-31++G\*\* and N=23 for LANL2DZ basis set).

		MSE	MUE	SD
HF/B1	pKa <sub>1</sub>	-0.47	1.56	2.31
	pKa <sub>2</sub>	-0.17	0.83	1.22
HF/B2	pKa <sub>1</sub>	-1.11	1.6	2.26
	pKa <sub>2</sub>	-0.01	0.89	1.28
B3LYP/B1	pKa <sub>1</sub>	-0.06	1.64	2.26
	pKa <sub>2</sub>	-1.15	1.96	2.18
B3LYP/B2	pKa <sub>1</sub>	-1.6	1.76	2.7
	pKa <sub>2</sub>	-1.62	2.03	1.92
Mean Value	pKa <sub>1</sub>	-0.84	1.48	2.41
	pKa <sub>2</sub>	0.67	1.12	1.57
Whitout compound 18, 20 and 21.				
HF/B1	pKa <sub>1</sub>	0.4	0.86	1.21
	pKa <sub>2</sub>	0.26	0.51	0.76
HF/B2	pKa <sub>1</sub>	-0.19	0.75	1.39
	pKa <sub>2</sub>	0.35	0.66	0.99
B3LYP/B1	pKa <sub>1</sub>	0.79	1.05	1.21
	pKa <sub>2</sub>	-0.45	1.29	1.5
B3LYP/B2	pKa <sub>1</sub>	0.64	0.83	1.33
	pKa <sub>2</sub>	-0.9	1.6	1.69
Mean Value	pKa <sub>1</sub>	0.02	0.63	1.24
	pKa <sub>2</sub>	-0.15	0.66	0.95

reduced to 0.02±1.24 and 0.15±0.95) without the compounds 18, 20 and 21. The MUE values are also reduced to 0.63±1.24 and 0.66±0.95 for ΔpKa<sub>1</sub>,ΔpKa<sub>2</sub> (mean value). These results, obtained without the explicit use of water molecules, are in the order of magnitudes of the values obtained by (Bishnu. T, 2016) in the case of pKa calculations of thiols with three water molecules via the use of SMD-SMD B3LYP / 6-311++G '(d, p). In design, the large errors recorded for compounds 18, 19, 20 and 21 and others are probably due to the difficulties of experimental measurements or to the calculation methods used or to both. Asidefrom taking experimental measurements of pKa for phosphonic acids, there is a way to correct these errors and give more reliable pKa values for these compounds. The method is modeling via linear regression analysis where Y is pKa response variable, X is pKa prediction variable, A and B are fitting parameters:

In the following, are presented and discussed, X is pKa prediction variable, A and B are fitting parameters:

$$Y = (A \pm \Delta A) + (B \pm \Delta B) * X \quad [6]$$

In the following, it is presented and discussed, the results of the relations between calculated vs. calculated pKa values and those of calculated vs. experimental.

### 3.2 Correlation of theory-theory

Table 4 lists the statistical parameters of linear regressions established. The  $R^2$  values are 0.92, 0.99 and SD 0.29, 0.89 pKa units indicating a best correlation theory – theory.

Figure 4 and 5, shows a linear correlation between the calculated pKa at the HF and B3LYP level using the bases 6-31++G\*\* and LANL2DZ.

**Table 4.** Statistical Parameters of Linear Correlations calculated pKa vs calculated pKa of phosphonic acids:  $Y=(A\pm\Delta A) +(B\pm\Delta B)*X$ ; Correlation coefficient  $R^2$ ; standard deviation SD

Y vs X	A± ΔA	B± ΔB	R <sup>2</sup>	SD
HF/B2 vs HF/B1.	-0,93±0.19	1,12±0.06	0,95	0,74
B3LYP/B1 vs HF/B1.	0,45±0.08	0,98±0.02	0,99	0,29
B3LYP/B2 vs HF/B1	-1,41±0.16	1,15±0.05	0,97	0,63
B3LYP/B1 vs HF/B2.	1,31±0.18	0,82±0.05	0,93	0,76
B3LYP/B2 vs HF/B2.	-0,48±0.11	1,00±0.03	0,98	0,49
B3LYP/B2 vs B3LYP/B1.	-1,91±0.21	1,16±0.06	0,95	0,74
HF/B2 vs HF/B1.	0,41±0.41	0,97±0.05	0,94	0,57
B3LYP/B1 vs HF/B1.	-3,60±0.52	1,36±0.07	0,95	0,73
B3LYP/B2 vs HF/B1	-3,42±0.54	1,24±0.07	0,94	0,76
B3LYP/B1 vs HF/B2.	-3,62±0.50	1,25±0.07	0,95	0,69
B3LYP/B2 vs HF/B2.	-3,91±0.52	1,36±0.07	0,92	0,89
B3LYP/B2 vs B3LYP/B1.	0,10±0.30	0,90±0.04	0,96	0,64

The slope values of pKa1 and pKa2 are close to unity varying from 0.82±0.05 to 1.16±0.06 pKa1 and from 0.90±0.04 to 1.36±0.07 pKa2. As for the intercept values, they vary in a range of -1.91±0.2 to 1.31±0.18 and -3.91±0.52, 0.41±0.41 for pKa1 and pKa2 respectively.

The slight variation of the slope (increase or decrease with respect to the unit) is probably due to underestimation of solvation energy of the conjugate bases of phosphonic acids by the SMD salvation model.

**Table 5.** Statistical Parameters of Linear Correlations calculated mean vs experiment of pKa1 and pKa2 dissociation constants of phosphonic

Y	A± ΔA	B ± ΔB	R <sup>2</sup>	SD	Without
HF/B1	-2.79±0.22	2.40± 0.22	0.89	0.44	14, 18, 20, 21.
HF/B2	-1.71±0.71	1.83± 0.30	0.73	0.60	14, 19, 20, 21, 23, 24.
B3LYP/B1	-2.59±0.48	2.49± 0.20	0.91	0.40	14, 18, 20, 21
B3LYP/B2	-1.97±0.56	1.70± 0.24	0.78	0.48	14, 18, 20, 21, 23, 24.
HF/B1	-5.10±1.06	1.65± 0.14	0.88	0.84	16.
HF/B2	-4.91±1.06	1.64± 0.14	0.87	0.86	16.
B3LYP/B1	-10.45±1.75	2.23± 0.23	0.84	1.39	16.
B3LYP/B1	-8.80±1.72	1.95± 0.20	0.82	1.30	23.

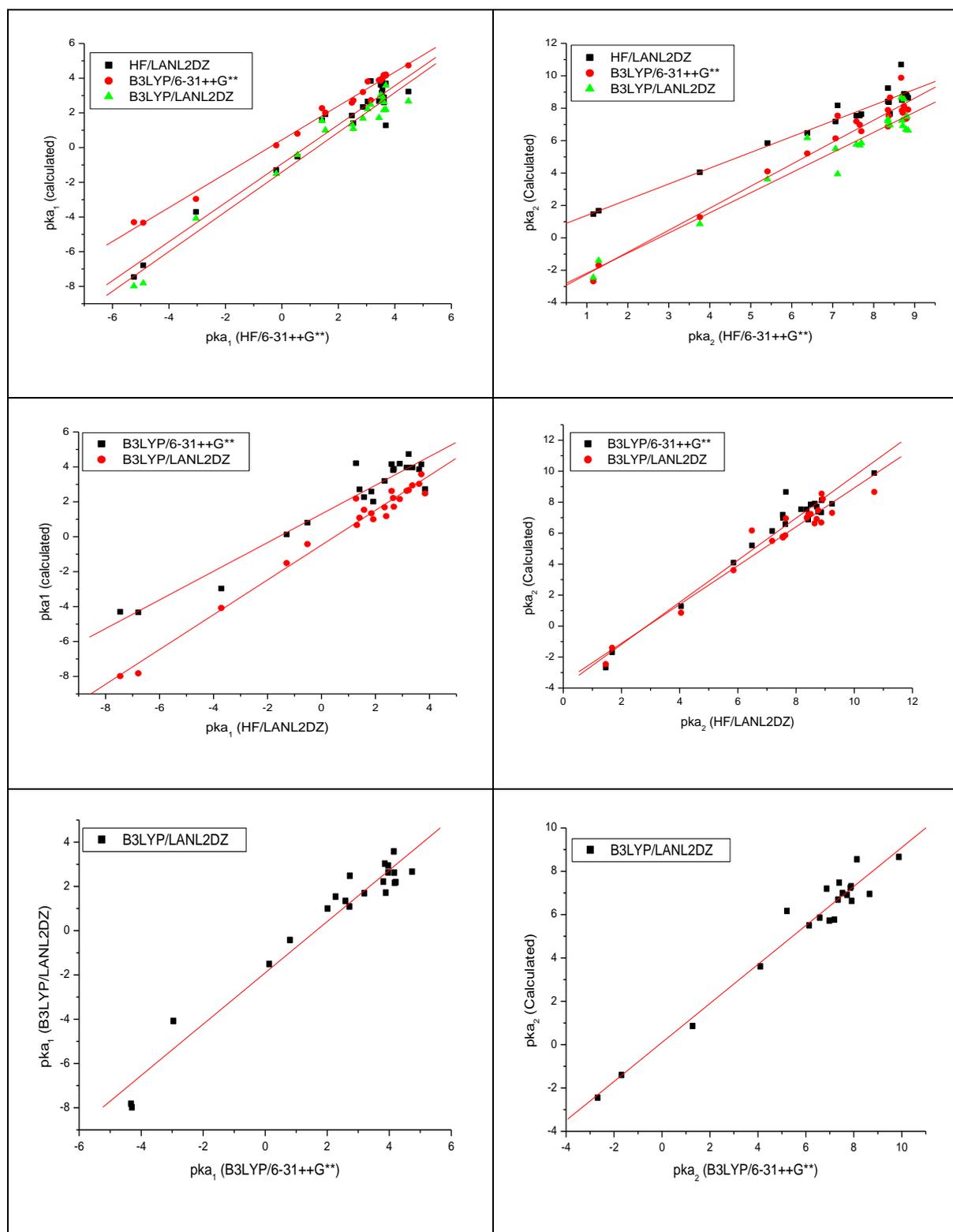
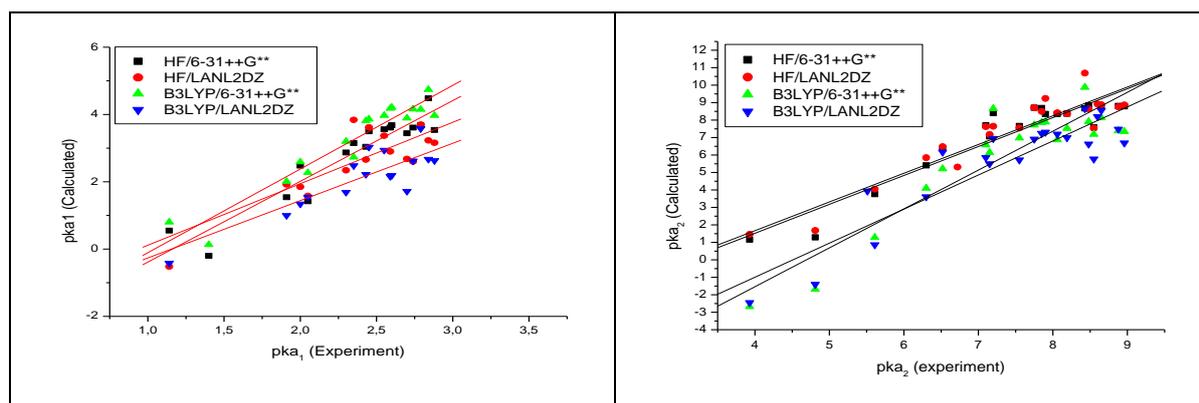


Figure 5: Linear correlation of calculated pKa<sub>1</sub> and pKa<sub>2</sub> value's at HF and B3LYP using 6-31++G\*\* and LANL2DZ basis set with PCM-SMD method of phosphonic acid derivatives.

### 3.3 Correlation of theory-experimental

In order to establish a prediction model for pKa of phosphonic acids, the pKa values calculated at the HF and B3LYP level and their mean values are correlated with the experimental ones available in the literature. The statistical analysis data are summarized in Table 5 and correlation lines are plotted on Figure 6.

The coefficient  $R^2$  of 0.8 to 0.9 with a SD of 0.4 to 0.6 pKa unit indicate a good correlation theory vs experience. However, it should be noted that  $R^2=0.73$  is relatively low in the case of calculation HF/ LANL2DZ. Moreover, the results show the effect of the AO's basis on the parameters A and B. Compared to the all-electron base 6-31++ G\*\*, the pseudo-potential base LANL2DZ decreases the slope of the correlation line of 0.60 at the HF level and 0.80 at the B3LYP level in the case of pKa<sub>1</sub>, whereas these variations are quite small in pKa<sub>2</sub> 0.01 and 0.26. The relatively small variations between the HF and DFT values + 0.09 for 6-31++G\*\* and -0.13 for LANL2DZ indicates an insignificant correlation electronic effect on the pKa values.



**Figure 6.** Linear correlation of experiment pKa<sub>1</sub> and pKa<sub>2</sub>'s dissociation constants versus calculated values at HF and B3LYP using 6-31++G\*\* and LANL2DZ basis set with PCM-SMD method of phosphonic acid derivatives

The intercept is 1.04 for HF and 0.62 for B3LYP in pKa<sub>1</sub> and 0.19 for HF and 1.65 for B3LYP in pKa<sub>2</sub>.

Moreover, compounds 14, 16, 18, 19, 20, 21, 23 and 24 have wide residues (X-X') or (Y-Y') (X' and Y' are the coordinates of the compound concerned on the correlation line). The difference is due to a strong overestimation or underestimation of the calculated or experimental pKa value or both.

Table 6 shows the statistical parameters of a linear correlation between the calculated means (HF, DFT) and the experimental values of pKa<sub>1</sub> and pKa<sub>2</sub> of the phosphonic acids. It is found that the coefficients of correlation  $R^2 = 0.91$  and  $R^2 = 0.86$  for pKa<sub>1</sub> and pKa<sub>2</sub> respectively, the most very weak deviations standard equal to 0.8 and 0.78 for pKa<sub>1</sub> and pKa<sub>2</sub> respectively.

**Table 6.** Statistical Parameters of Linear Correlations calculated mean vs experiment of pKa<sub>1</sub> and pKa<sub>2</sub> dissociation constants of phosphonic acids.

	pKa <sub>1</sub>	pKa <sub>2</sub>
A ± ΔA	-1.91±0.39	-0.80±0.94
B ± ΔB	2.01±0.16	1.08±0.12
R <sup>2</sup>	0.91	0.86
SD	0.80	0.78
N	17	16
Without	14, 18, 19, 20, 21 and 24.	13, 15, 17, 18, 19, 20 and 22.

Indeed, from Table 1, it should be noted that the experimental values of pKa used in the correlations for compounds 14 and 16 were not measured by (Freedman, L.D, 1957). Their pKa values are from reference [D.E.C. Corbridge, Phosphorus: Chemistry, Biochemistry and Technology, Sixth Edition, p 379]. As for the pKa's halogenated compounds 18, 19, 20, 21 and 23, their pKa<sub>1</sub> values are certainly quite low one (pH unit) in view of the relatively stronger acidity because of halogen substituent. In addition, some pKa aliphatic compounds were deduced by extrapolation considering the number of carbon atoms.

To scale, the pKa<sub>1</sub> constant of all R-PO<sub>3</sub>H<sub>2</sub> phosphonic acids including R=H, the pKa<sub>1</sub> experimental values are compared to those of the corresponding R-COOH carboxylic acids. Table 7(appendix) lists the dissociation constant values with the relative difference ΔpKa. It is noted that the compounds 14, 18, 19, 20, 21, 22 and 23 have relatively, small deviations compared to those of aliphatic ones. The mean relative difference values (MRD) with and without these compounds are 1.73±0.95 and 2.21±0.54 pKa units respectively.

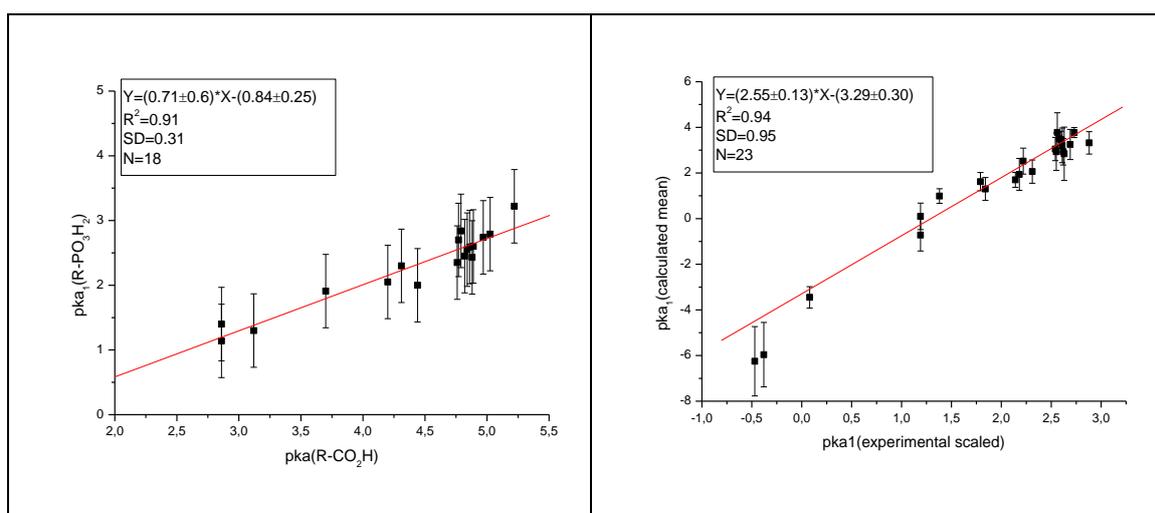
The scaled pKa's of phosphonic acids show good linear correlation with the corresponding pKa<sub>1</sub> mean values calculated at SMD-HF and B3LYP levels with R<sup>2</sup> value of 0.94 figure 7. However, the pKa<sub>1</sub> values of 18, 20 and 21 compounds are relatively underestimated. All the same, the MSE is reduced to 0.57±1.85 with N=24 compounds versus -0.06 ± 1.03 pKa units without 18, 20 and 21 compounds.

#### IV. CONCLUSIONS

The relative pKa<sub>1</sub> and pKa<sub>2</sub> of 24 phosphonic acids in aqueous solution are carried out at SMD-HF/6-31++G\*\*//SMD-HF/6-31++G\*\*, SMD-HF/LANL2DZ//SMD-HF/LANL2DZ, SMD-B3LYP/6-31++G\*\*//SMD-B3LYP/6-31++G\*\*, SMD-B3LYP/LANL2DZ//SMD-HF/LANL2DZ.

The correlations between calculated pKa and experimental ones are of satisfactory quality.

The results obtained give confidence in the calculation methods and offer a means of predicting experimental pKa values for phosphonic acids, which present difficulties with experimental measurements.



**Figure7:** (a) Linear correlation between experimental dissociation constant values of phosphonic acids (pKa<sub>1</sub>) and carboxylic acids (pKa). (b) Linear correlation between experimental scaled and calculated mean dissociation constant values of phosphonic acids.

#### Acknowledgement

The Polydisciplinary Faculty of Taroudant- Ibn Zohr University was supported this work.

#### Supporting Information

The calculations of the 24 compounds studied are carried out within the research team "Physic chemistry of Natural Environments, Materials and Environment (PCMME)", Polydisciplinary Faculty of Taroudant - University Ibn Zohr.

#### REFERENCES

- [1]. Moore, R., Lopes, J., 1999. Paper templates. In TEMPLATE'06, 1st International Conference on Template Production. SCITEPRESS.
- [2]. Smith, J., 1998. The book. The publishing company. London, 2<sup>nd</sup> edition. general calculation: [Bishnu Thapa and H. Bernhard Schlegel]. *J. Phys. Chem. A* 2016, 120, 5726-5735
- [3]. P. Nylen, Dissertation, Uppsala, 1930, pp. 148-149.
- [4]. P. Rumpf and V. Chavanne, *Compt. rend.*, 114, 919 (1947).
- [5]. J. R. White, *THIS JOURNAL* 7, 2, 1859 (1950).
- [6]. Crofts, P. C. and Kosolapoff, G.M.: *J. Am. Chem. Soc.* 75, 3379 (1953).
- [7]. Crofts, P. C. and Kosolapoff, G.M.: *J. Am. Chem. Soc.* 75, 5738 (1953).
- [8]. Liptak, M.D., Shields, G.C. *J. Am. Chem. Soc.* 2001, 123(30), 7314-9.
- [9]. Liptak, M.D., Gross, K.C., Seybold, P.G., Feldgus, S., Shields, G.C. *J. Am. Chem. Soc.* 2002, 124(22), 6421.
- [10]. Gaussian 16 revision A.03 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R.
- [11]. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J.
- [12]. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L.
- [13]. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson,
- [14]. D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R.
- [15]. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A.

- [16]. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016
- [20]. Becke AD, Density-functional thermochemistry. III. The role of exact exchange, *J Chem. Phys.* 98:5648–5652, 1993.
- [21]. Lee C, Yang W, Parr RG, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys Rev B*, 37:785–789, 1988.
- [22]. Michlich B, Savin A, Stoll H, Preuss H, Results obtained with the correlation-energy density functionals of Becke and Lee, Yang and Parr, *ChemPhys Lett* 157:200–206, 1989.
- [23]. M. J. S. Dewar, W. Thiel, *J. Am. Chem. Soc.* 99 (1977) 4899.
- [24]. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* 107 (1985) 3902.
- [25]. P. C. Hariharan, J. A. Pople, *Mol. Phys.* 27 (1974) 209.
- [26]. P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc.* 28 (1973) 213.
- [27]. P. J. Hay and W. R. Wadt, “Ab initio effective core potentials for molecular calculations - potentials for the transition-metal atoms Sc to Hg,” *J. Chem. Phys.*, 82 (1985) 270-83
- [28]. T. H. Dunning Jr. and P. J. Hay, in *Modern Theoretical Chemistry*, Ed. H. F. Schaefer III, Vol. 3 (Plenum, New York, 1976) 1-28.
- [29]. W. R. Wadt and P. J. Hay, “Ab initio effective core potentials for molecular calculations - potentials for main group elements Na to Bi,” *J. Chem. Phys.*, 82 (1985) 284-98.
- [30]. P. J. Hay and W. R. Wadt, “Ab initio effective core potentials for molecular calculations - potentials for K to Au including the outermost core orbitals,” *J. Chem. Phys.*, 82 (1985) 299-310
- [31]. Barone, M. Cossi, *J. Phys. Chem. A* 102 (1998) 1995.
- [32]. M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.* 24 (2003) 669
- [33]. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* 2009, 113, 6378–6396].
- [34]. M. J. S. Dewar, W. Thiel, *J. Am. Chem. Soc.* 99 (1977) 4899.
- [35]. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* 107 (1985) 3902.
- [36]. BishnuThapa and H. Bernhard Schlegel *J. Phys. Chem. A* 2016, 120, 5726-5735
- [37]. Freedman, L.D.; Doak, G.O. *Chem. Rev.* 1957, 57, 479.
- [38]. Riddick, J.A., W.B. Bunger, Sakano T.K. *Techniques of Chemistry* 4th ed., Volume II. Organic Solvents. New York, NY: John Wiley and Sons. 1985. p. 360
- [39]. Serjeant, E.P., Dempsey B.; *Ionisation Constants of Organic Acids in Aqueous Solution*. International Union of Pure and Applied Chemistry (IUPAC).
- [40]. IUPAC Chemical Data Series No. 23, 1979. New York, New York: Pergamon Press, Inc., p. 989
- [41]. Riddick, J.A., W.B. Bunger, Sakano T.K. *Techniques of Chemistry* 4th ed., Volume II. Organic Solvents. New York, NY: John Wiley and Sons. 1985. p. 1325
- [42]. Dean, J.A. *Handbook of Organic Chemistry*. New York, NY: McGraw-Hill Book Co., 1987. p. 8-45
- [43]. Dean, J.A. *Handbook of Organic Chemistry*. New York, NY: McGraw-Hill Book Co., 1987. p. 8-31
- [44]. Kortum G et al; *Dissociation Constants of Organic Acids in Aqueous Solution*. International Union of Pure and Applied Chemistry. London: Butterworth (1961).
- [45]. Haynes, W.M. (Ed.). *CRC Handbook of Chemistry and Physics*. 95 th Edition. CRC Press LLC, Boca Raton: FL 2014-2015, p. 5-100
- [46]. Riddick, J.A., W.B. Bunger, Sakano T.K. *Techniques of Chemistry* 4th ed., Volume II. Organic Solvents. New York, NY: John Wiley and Sons. 1985. p. 376
- [47]. Haynes, W.M. (ed.) *CRC Handbook of Chemistry and Physics*. 91 st ed. Boca Raton, FL: CRC Press Inc., 2010-2011, p. 8-47
- [48]. Serjeant, E.P., Dempsey B.; *Ionisation Constants of Organic Acids in Aqueous Solution*. International Union of Pure and Applied Chemistry (IUPAC).
- [49]. IUPAC Chemical Data Series No. 23, 1979. New York, New York: Pergamon Press, Inc., p. 21
- [50]. O’Neil, M.J. (Ed.). *The Merck Index - An Encyclopedia of Chemicals, Drugs, and Biologicals*. 13th Edition, Whitehouse Station, NJ: Merck and Co., Inc., 2001. p. 1725
- [51]. Dean, J. A. (Ed) *Lange’s Handbook of Chemistry* (15th Ed) McGraw-Hill Inc. 1999
- [52]. Dean, J. A. (Ed) *Lange’s Handbook of Chemistry* (15th Ed) McGraw-Hill Inc. 2000
- [53]. Dean, J. A. (Ed) *Lange’s Handbook of Chemistry* (15th Ed) McGraw-Hill Inc. 2001
- [54]. Brown, H. C.; McDaniel, H.; Hafliger, O. In “determination of organic structure by physical methods”
- [55]. Braude, E. A.; Nachod, F. C. Eds. Academic press Inc, New York, 1955, pp567-662
- [56]. Lide, DR (Ed.). *CRC Handbook of Chemistry and Physics*. 8 first Edition. CRC Press LLC, Boca Raton: FL 2000, p. 8-46.
- [57]. Colvin, E. Butterworth, *Silicon in Organic Synthesis*: London 1981.

**APPENDIX**

**Table 1.** Calculated pKa's at HF and B3LYP level using 6-31++G\*\* (B1) and LANL2DZ (B2) basis set with SMD Solvation model and Experimental data of the phosphonic acids.

Compounds		Calculated values (this work)								Mean values		Literature				
N	R-PO(OH) <sub>2</sub>	HF/B1		HF/B2		B3LYP/B1		B3LYP/B2		pKa <sub>1</sub> Cal±SD	pKa <sub>2</sub> Cal±SD	[1]		[2]		
1	Hydrogene												2,00	6,58	1,30	6,70
2	Methyl	3,15	8,40	3,84	7,65	2,73	8,66	2,48	6,95	3.05±0.51	7.91±0.61	2,35	7,20	2,68	6,32	
3	Ethyl	3,04	8,69	2,66	8,50	3,81	7,86	2,22	7,24	2.93±0.58	8.08±0.64	2,43	7,85	2,43	8,05	
4	n-propyl	3,50	8,35	3,62	8,42	3,86	6,87	3,03	7,20	3.50±0.30	7.71±0.36	2,45	8,06	2,49	8,18	
5	n-Butyl	3,61	8,37	2,90	8,37	4,18	7,53	2,16	6,99	3.21±0.76	7.82±0.63	2,59	8,19			
6	n-Hexyl	3,68	8,35	1,28	9,24	4,21	7,89	2,19	7,31	2.84±1.17	8.20±0.69	2,60	7,90			
7	isoPropyl	3,56	8,71	3,37	8,71	3,97	7,72	2,94	6,91	3.46±0.37	8.01±0.82	2,55	7,75	2,79	8,88	
8	isoButyl	3,44	8,67	2,68	10,6 9	3,89	9,88	1,72	8,66	2.93±0.82	9.47±0.92	2,70	8,43			
9	s-butyl	3,61	8,84	2,60	8,64	4,16	7,91	2,62	6,63	3.25±0.66	8.00±0.96	2,74	8,48			
10	t-butyl	3,68	8,81	3,70	8,76	4,15	7,39	3,58	7,48	3.77±0.22	8.11±0.67	2,79	8,88			
11	neopentyl	4,48	8,74	3,23	8,88	4,74	8,13	2,67	8,55	3.78±0.86	8.57±0.19	2,84	8,65			
12	1,1-dimethylpropyl	3,54	8,79	3,16	8,87	3,97	7,34	2,63	6,69	3.32±0.49	7.92±0.99	2,88	8,96			
13	Benzyl	2,87	7,66	2,34	7,55	3,20	6,98	1,69	5,72	2.52±0.57	6.98±0.84	2,30	7,55	2,38	5,98	
14	Ethenyl	2,53	7,58	1,41	7,54	2,72	7,19	1,09	5,77	1.94±0.70	7.02±0.78			3,48	8,55	
15	2-phenylethenyl	2,48	7,70	1,85	7,63	2,59	6,58	1,34	5,86	2.06±0.51	6.94±0.84	2,00	7,10			
16	phenyl	1,42	7,12	1,58	8,17	2,27	7,54	1,54	3,94	1.70±0.33	6.69±1.59			2,05	5,51	
17	methoxy	1,54	7,07	1,92	7,18	2,01	6,14	1,00	5,50	1.62±0.40	6.47±0.73	1,91	7,15			
18	trifluoromethyl	-5,24	1,16	-7,46	1,46	-4,30	2,68	7,98	2,45	6.25±1.52	-0.63±1.88	1,16	3,93	1,46	2,43	
19	chloromethyl	-0,20	5,41	-1,29	5,85	0,13	4,10	1,51	3,61	0.72±0.70	4.74±0.92	1,40	6,30	1,70	5,22	
20	dichloromethyl	-3,04	3,76	-3,71	4,05	-2,96	1,28	4,08	0,86	3.45±0.47	2.49±1.44	1,14	5,61			
21	trichloromethyl	-4,91	1,29	-6,79	1,68	-4,33	1,69	7,82	1,40	5.96±1.41	-0.03±1.44	1,63	4,81			
22	bromomethyl	0,55	6,38	-0,52	6,48	0,80	5,21	0,42	6,17	0.10±0.58	6.06±0.31	1,14	6,52			
23	iodomethyl	***	***	1,31	5,31	***	***	0,67	10,9 5	0.99±0.32	8.13±2.82	1,30	6,72			
24	trimethylsilano methyl	***	***	2,40	8,92	***	***	1,18	8,20	1.79±0.61	8.56±0.36	3,22	8,60			

**Table 2.** Errors ( $\Delta pK_a$ ) mean signed errors (ME), mean unsigned errors and standard deviation (SD) in  $pK_a$ 's of phosphonic acids calculated at SMD-HF and SMD-B3LYP.

Compounds		$\Delta pK_{a1}$ and $\Delta pK_{a2}$												
N	R-P(O)(OH) <sub>2</sub>	HF/B1		HF/B2		B3LYP/B1		B3LYP/B2		Mean value		Exp		Ref
2	Methyl	0.80	1,20	1,49	0,45	0,38	1,46	0,13	-0,25	0,70	0,71	2,35	7,20	[1]
3	Ethyl	0.61	0,84	0,23	0,65	1,38	0,01	-0,21	-0,61	0,50	0,23	2,43	7,85	[1]
4	n-propyl	1.50	0,29	1,17	0,36	1,41	-1,19	0,58	-0,86	1,05	-0,35	2,45	8,06	[1]
5	n-Butyl	1.02	0,18	0,31	0,18	1,59	-0,66	-0,43	-1,20	0,62	-0,37	2,59	8,19	[1]
6	n-Hexyl	1.08	0,45	-1,32	1,34	1,61	-0,01	-0,41	-0,59	0,24	0,30	2,60	7,90	[1]
7	isoPropyl	1.01	0,96	0,82	0,96	1,42	-0,03	0,39	-0,84	0,91	0,26	2,55	7,75	[1]
8	isoButyl	0.47	0,24	-0,02	2,26	1,19	1,45	-0,98	0,23	0,23	1,04	2,70	8,43	[1]
9	s-butyl	0.87	0,36	-0,14	0,16	1,42	-0,57	-0,12	-1,85	0,51	-0,48	2,74	8,48	[1]
10	t-butyl	0.89	-0,07	0,91	-0,12	1,36	-1,49	0,79	-1,40	0,98	-0,77	2,79	8,88	[1]
11	néopentyl	1.64	0,09	0,39	0,23	1,90	-0,52	-0,17	-0,10	0,94	-0,08	2,84	8,65	[1]
12	1,1 dimethylpropyl	0.66	-0,17	0,28	-0,09	1,09	-1,62	-0,25	-2,27	0,44	-1,04	2,88	8,96	[1]
13	Benzyl	0.57	0,11	0,04	0,00	0,90	-0,57	-0,61	-1,83	0,22	-0,57	2,30	7,55	[1]
14	Ethenyl	-0.95	-0,97	-2,07	-1,01	-0,76	-1,36	-2,39	-2,78	-1,54	-1,53	3,48	8,55	[2]
15	2-phenylethenyl	0.48	0,60	-0,15	0,53	0,59	-0,52	-0,66	-1,24	0,06	-0,16	2,00	7,10	[1]
16	phenyl	-0.63	1,61	-0,47	2,66	0,22	2,03	-0,51	-1,57	-0,35	1,18	2,05	5,51	[2]
17	methoxy	-0.37	-0,08	0,01	0,03	0,10	-1,01	-0,91	-1,65	-0,29	-0,68	1,91	7,15	[1]
18	trifluoromethyl	-6.40	-2,77	-8,62	-2,47	-5,46	-6,61	-9,14	-6,38	-7,41	-4,56	1,16	3,93	[1]
19	chloromethyl	-1.60	-0,89	-2,69	-0,45	-1,27	-2,20	-2,91	-2,69	-2,12	-1,56	1,40	6,30	[1]
20	dichloromethyl	-4.18	-1,85	-4,85	-1,56	-4,10	-4,33	-5,22	-4,75	-4,59	-3,12	1,14	5,61	[1]
21	trichloromethyl	-6.54	-3,52	-8,42	-3,13	-5,96	-6,50	-9,45	-6,21	-7,59	-4,84	1,63	4,81	[1]
22	bromomethyl	-0.59	-0,14	-1,66	-0,04	-0,34	-1,31	-1,56	-0,35	-1,04	-0,46	1,14	6,52	[1]
23	iodomethyl			0,01	-1,41			-0,63	2,50	-0,31	1,41	1,30	6,72	[1]
24	trimethylsilano méthyl			-0,82	0,32			-2,04	-0,40	-1,43	-0,04	3,22	8,60	[1]

**Table 3:** Relative difference ( $\Delta pK_a$ ), mean relative difference (MRD), standard deviation (SD) of  $pK_{a1}$  values of phosphonic acids compared to  $pK_a$  of carboxylic acids.  $\Delta pK_a(\text{Exp-Exp}) = pK_a(\text{RCOOH}) - pK_{a1}(\text{RPO}_3\text{H}_2)$ , and  $\Delta pK_a(\text{Scal-Calc}) = pK_{a1}(\text{scaled}) - pK_{a1}(\text{calc.})$ , and errors ( $\Delta pK_{a1}$ ), mean signed errors (MSE), and standard deviation (SD) in calculated  $pK_a$ 's of phosphonic acids at SMD-HF and SMD-B3LYP

N°	Substituent		pKaexperimental		$\Delta pK_a$ (Exp-Exp)	pKa <sub>1</sub> (RPO <sub>3</sub> H <sub>2</sub> )		$\Delta pK_a$ (Scal-Calc)
	Name	R	RCOOH (ref)	RPO <sub>3</sub> H <sub>2</sub>		Scaled	Calculated	
1	Hydrogene	H	3,77 (Riddick J.A,1985)	1,30	2,47	1,84	1,3	0,54
2	Methyl	CH <sub>3</sub>	4,76 (Sørjeant, E.P, 1979)	2,35	2,41	2,54	3,05	-0,51
3	Ethyl	C <sub>2</sub> H <sub>5</sub>	4,88 (Sørjeant, E.P, 1979)	2,43	2,45	2,62	2,93	-0,31
4	n-propyl	C <sub>3</sub> H <sub>7</sub>	4,82 (Riddick J.A,1985)	2,45	2,37	2,58	3,5	-0,92
5	n-Butyl	C <sub>4</sub> H <sub>9</sub>	4,84 (Dean, J.A, 1987)	2,59	2,25	2,61	3,21	-0,6
6	n-Hexyl	C <sub>6</sub> H <sub>13</sub>	4,89 (Dean, J.A, 1987)	2,60	2,29	2,63	2,84	-0,21
7	isoPropyl	(CH <sub>3</sub> ) <sub>2</sub> CH	4,84 (Kortum G, 1961)	2,55	2,29	2,6	3,46	-0,86
8	isoButyl	(CH <sub>3</sub> ) <sub>2</sub> CHCH <sub>2</sub>	4,77 (Kortum G, 1961)	2,70	2,07	2,55	2,93	-0,39
9	s-butyl	CH <sub>3</sub> -CH <sub>2</sub> CH(CH <sub>3</sub> )	4,97 (Dean, J.A, 1987)	2,74	2,23	2,69	3,25	-0,56
10	t-butyl	(CH <sub>3</sub> ) <sub>3</sub> -C	5,03 (Dean, J.A, 1987)	2,79	2,24	2,73	3,77	-1,05
11	neopentyl	(CH <sub>3</sub> ) <sub>3</sub> -C-CH <sub>2</sub>	4,79 [6]	2,84	1,95	2,56	3,78	-1,22
12	t-amyl	C <sub>2</sub> H <sub>5</sub> (CH <sub>3</sub> ) <sub>2</sub> C	NA	2,88	----	2,88	3,32	-0,44
13	Benzyl	C <sub>6</sub> H <sub>5</sub> -CH <sub>2</sub>	4,31 (Haynes, W.M 2015)	2,30	2,01	2,22	2,52	-0,3
14	Ethenyl or vinyl	CH <sub>2</sub> =CH	4,25 (Riddick J.A, 1985)	3,48	0,77	2,18	1,94	0,24
15	2-phenylethenyl	C <sub>6</sub> H <sub>5</sub> -CH=CH	4,44 [8]	2,00	2,44	2,31	2,06	0,25
16	phenyl	C <sub>6</sub> H <sub>5</sub>	4,2 (Haynes, W.M 2011)	2,05	2,15	2,14	1,7	0,44
17	methoxy	HOCH <sub>2</sub>	3,83 (Sørjeant, E.P, 1979)	1,91	1,92	1,79	1,62	0,17
18	trifluoromethyl	CF <sub>3</sub>	0,3 (O'Neil, M.J, 2001)	1,16	-0,86	-0,47	-6,25	5,78
19	chloromethyl	CH <sub>2</sub> Cl	2,86 (Dean, J.A, 1999)	1,40	1,46	1,19	-0,72	1,91
20	dichloromethyl	CHCl <sub>2</sub>	1,26 (Dean, J.A, 2000)	1,14	0,12	0,08	-3,45	3,53
21	trichloromethyl	CCl <sub>3</sub>	0,52 (Dean, J.A, 2001)	1,63	-1,11	-0,38	-5,96	5,58
22	bromomethyl	CH <sub>2</sub> Br	2,86 (Brown, H. C, 1955)	1,14	1,72	1,19	0,1	1,09
23	iodomethyl	CH <sub>2</sub> I	3,18 (Lide, DR, 2000)	1,30	1,88	1,38	0,99	0,39
24	trimethylsilanomethyl	(CH <sub>3</sub> ) <sub>3</sub> -Si-CH <sub>2</sub>	5,22 (Colvin, E, 1981)	3,22	2,00	2,87	1,79	1,08
mean relative difference (MRD)					1,73	MSE		0,57
standard deviation (SD)					0,95	SD		1,85
Without 14, 18, 19, 20, 21, 22 and 23.							Without 18, 20 and 21.	
mean relative difference (MRD)					2,21	MSE		-0,06
standard deviation (SD)					0,53	SD		1,03