DETERMINATION OF $p\mathrm{K_a}$ AND $p\mathrm{K_b}$ FROM ELECTRONIC PROPERTIES DERIVED FROM CONCEPTUAL DENSITY FUNCTIONAL THEORY (CDFT)

JOSÉ MUÑOZ-ESPINOZA [1,2], GERMÁN BARRIGA-GONZÁLEZ [2], AND SERGIO JIMÉNEZ-MEZA [2]*

[1]*Theoretical Chemistry Laboratory, Department of Chemistry, Faculty of Sciences, University of Chile, Santiago Chile*
[2]*MACEDONIA Laboratory, Department of Chemistry, Basic Sciences Faculty, Metropolitan University of Educational Sciences, Santiago Chile.*

ABSTRACT

In the present work we apply a semiempirical method capable of calculating the $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values of a series of organic acids and bases through their electronic properties. This multilinear model, analogous to the one introduced by Kamlet-Taft, relates acid-base properties such as HBA and HBD to the regional electrophilicity and nucleophilicity derived from cDFT. To test the model, it was applied to a series of mono and dicarboxylic acids, as well as aliphatic/cyclic and heterocyclic amines, showing us that the model is only functional when the series of compounds present similar characteristics such as functional groups. Based on the regionalization of the electrophilicity and nucleophilicity, it can be observed that the acid-base characteristics are not exclusive to limited regions of the molecules, but that these present a bifunctional character that shifts to the relative characteristics of acid and/or base according to the global distribution of the electronic density

Keywords: Acidity; Basicity; Nucleophilicity; Electrophilicity; Multilinear Correlation.

1. INTRODUCTION

Acid-base theory is one of the foundations of chemistry, since its understanding has made it possible to explain elementary phenomena ranging from the composition of the acid present in the defense mechanism of red ants (*Formica rufa*) [1], to the structure of coordination compounds [2]. This theory has varied over time, as its formulation has been approached from different approaches. One of the first approaches to the concepts of *acid* and *base* dates back to the works of Lémary (1717) [3], François Rouelle (1754) [4] and von Liebig (1838) [5], whose definitions and description of an acid-base reaction were the foundations of the well-known theory of Arrhenius (1884) [6]. From this theory new models were developed that corrected the errors underlying the predecessor theory (e.g. basic nature of ammonia), as does the model of Brønsted [7] and Lowry [8], which introduce the concepts of conjugate acids and bases, as well as the amphoteric nature of water. The current acid-base theory was developed in parallel to the Brønsted and Lowry model and is based on the electronic characteristics of atoms and molecules. The Lewis theory [9] describes the acid-base nature of atoms and molecules as species that either accept an electron pair (acid, electrophile) or donate an electron pair (base, nucleophile). This theory, together with the work of Koopmans [10] and Fukui et. al [11], encompasses the theory of frontier molecular orbitals (FMO theory), which describes chemical reactions as an interaction and effective overlap of the HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital) orbitals.

One of the most important properties derived from the acid-base theory are the acidic (K_a) and basic (K_b) dissociation constants, which according to the conjugate pair theory of Brønsted and Lowry are associated with a chemical equilibrium between the acid/base and its conjugate pairs (eq.1)

$$\mathrm{HA} \underset{\mathrm{K}_a}{\rightleftarrows} \mathrm{H^+ + A^-}\ ,\ \mathrm{K}_a = \frac{c_{\mathrm{H^+}}\, c_{\mathrm{A^-}}}{c_{\mathrm{HA}}}\ ;\ \mathrm{B} \underset{\mathrm{K}_b}{\rightleftarrows} \mathrm{BH^+ + OH^-}\ ,\ \mathrm{K}_b = \frac{c_{\mathrm{BH^+}}\, c_{\mathrm{OH^-}}}{c_{\mathrm{B}}} \qquad (1)$$

Generally, its expression and interpretation is associated with the negative logarithm of the dissociation constant, defined as $p\mathrm{K}_a$ and $p\mathrm{K}_b$. For the case of acids, a small $p\mathrm{K}_a$ is associated with a high relative strength of acidity (mineral acids [12]), while a large $p\mathrm{K}_a$ is associated with a low relative strength of acidity (organic acids [13]) [14]. For the case of the relative strength of bases, the same reasoning is followed (small $p\mathrm{K}_b$ for strong bases and high $p\mathrm{K}_b$ for weak bases).

The $p\mathrm{K}_a$ is a fundamental parameter since it influences different physicochemical, biochemical and physiological phenomena, such as chemical reactions [15], drug solubility [16], ligand-receptor interactions [17] and renal absorption/elimination of drugs [18]. That is why their experimental or theoretical determination is of great help for the understanding of different phenomena. Experimentally, $p\mathrm{K}_a$ values can be obtained from acid-base titrations [19] to spectroscopic methods [20,21]. From a theoretical and computational point of view, $p\mathrm{K}_a$ values are determined from thermodynamic cycles that include the abstraction of acidic protons and their subsequent solvation [12,22-24]. On the other hand, Sandoval-Lira et. al [25] proposed a method to determine $p\mathrm{K}_b$ values from the electrostatic surface potential of amines, obtaining significant values.

Kutt et. al [14] posit that the acid-base characteristics of organic compounds (mainly solvents) depend on HBD (Hydrogen Bond Donor) and HBA (Hydrogen Bond Acceptor) properties, which are related to Lewis theory, where HBD is associated with electrophilicity and HBA with nucleophilicity [26,27]. Given this relationship, $p\mathrm{K}_a$ and $p\mathrm{K}_b$ can be modeled as a function of electrophilicity and nucleophilicity, which can be described by descriptors derived from conceptual density functional theory (cDFT).

To correlate the HBA and HBD parameters with the $p\mathrm{K}_a$ and $p\mathrm{K}_b$, one can consider a multi-linear model analogous to that of Kamlet-Taft [28], which sets the parameter β as the HBA capacity of the composite [29], while the parameter α corresponds to the HBD capacity [30]. The general formulation in that empirical approach is to express the property $p\mathrm{K}_i$ $(i = a, b)$ as a multiparametric function as ($p\mathrm{K}_i = f(\alpha, \beta, \pi^*)$):

$$p\mathrm{K}_i = f(\alpha, \beta, \pi^*) = (p\mathrm{K}_i)_0 + a\alpha + b\beta + c\pi^* \qquad (2)$$

Where $(p\mathrm{K}_i)_0$ represents the intrinsic property of the isolated compound, α the HBD capacity, β the HBA capacity and π^* the polarizability of the compound; and a, b and c, the coefficients of the multilinear relationship defined by eq. (2).

The aim of this work is to describe the influence of the HBA and HBD capacity of organic acids and bases on the $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values, and to present a semi-empirical model to calculate $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values from the electronic properties of the compounds.

2. THEORY AND COMPUTATIONAL DETAIL

To describe the electrophilic (HBD) and nucleophilic (HBA) characteristics of organic acids and bases and their influence on $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values, use will be made of DFT-derived global and local descriptors such as electronic chemical potential (μ) [31], hardness (η), softness (S) [32], electrophilicity (ω^+) [33] and Fukui functions (FF) [34].

According to the Hohenberg - Kohn theorem [35], the energy is a functional that depends on the electron density $\rho_{(r)}$ of a system, and this in turn depends on the number of electrons N and the external potential $v_{(r)}$. In the context of chemical reactivity proposed by Parr and Yang [31], the energy can be written as a second order Taylor series:

$$\Delta \mathrm{E} = \left(\frac{\partial E}{\partial N}\right)_{v_{(r)}} \Delta N + \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_{v_{(r)}} \Delta N^2 = \mu \Delta N + \frac{1}{2}\eta \Delta N^2 \qquad (3)$$

Where the concepts of electronic chemical potential (μ) and chemical hardness (η) are introduced. The electronic chemical potential corresponds to the first

*Corresponding author email: jose.munoz.e@ug.uchile.cl

functional derivative of the energy with respect to the number of electrons N, while the chemical hardness corresponds to the second functional derivative of the energy. Both expressions can be written in terms of the HOMO (ε_{H}) and LUMO (ε_{L}) energies from the Janak-Koopmans theorem [10,36]:

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v_{(r)}} = \left(\frac{\varepsilon_{\mathrm{L}} + \varepsilon_{\mathrm{H}}}{2}\right) \qquad (4)$$

$$\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{v_{(r)}} = \left(\frac{\varepsilon_{\mathrm{L}} - \varepsilon_{\mathrm{H}}}{2}\right) \qquad (5)$$

According to the Parr and Yang parabola model [34], an electrophile immersed in a sea of free electrons of chemical potential will lead to a maximum charge transfer (ΔN_{max}) that minimizes energy, so that from eq. (3) it follows that:

$$\Delta \mathrm{E} = \mu \Delta \mathrm{N} + \frac{1}{2}\eta \Delta \mathrm{N}^2 \;\rightarrow\; \frac{\partial \Delta \mathrm{E}}{\partial \Delta \mathrm{N}} = \mu + \eta \Delta \mathrm{N} = 0 \qquad (6)$$

Which leads us to:

$$\Delta \mathrm{N}_{max} = -\frac{\mu}{\eta} \qquad (7)$$

Therefore, by replacing eq. (7) in (3) we obtain

$$-\Delta \mathrm{E}^{+} = \frac{\mu^2}{2\eta} \equiv \omega^{+} \qquad (8)$$

Where ω^{+} is known as the global electrophilicity index and can be understood as the stabilization energy of a system when it is saturated with electrons from the environment.

Since nucleophilicity cannot be described by a model analogous to equation (8), Contreras et. al [37] proposed a regional nucleophilicity model derived from the Politzer density-energy relation [38], which is associated with the first ionization energy, which in the context of the frontier molecular orbital theory (FMO theory) can be described as the HOMO energy:

$$\omega_k^{-} = -f_k^{-} I \approx f_k^{-} \varepsilon_{\mathrm{H}} \qquad (9)$$

So far, correlations between the HBA and HBD parameters with nucleophilicity and electrophilicity derived from cDFT have been performed, so it remains for us to find a relationship between the polarizability π^{*} with some index of reactivity. Recently, Weiß et. al [39] propose that the π^{*} parameter can be directly related to the polarizability of the molecular electron density, so that it can be associated with the work presented by Vela & Gazquez [40] and Simón-Manso & Fuentealba [41], which linearly correlate molecular polarizability with global softness (S) ($S \sim \langle\alpha\rangle^{1/3}$).

To describe the electrophilic and nucleophilic capacity of an organic acid and/or base, it is necessary to introduce the regional reactivity defined by the Fukui functions (FF), which show us the sites with electrophilic (f_k^{+}) and nucleophilic (f_k^{+}) characteristics and are associated with the Lewis acid (HBD) and Lewis base (HBA) characteristics. Given the discontinuity of N, the condensed Fukui functions can be written from finite differences [42].

$$f_{(r)}^{+} = \left(\frac{\partial \rho_{(r)}}{\partial N}\right)_{v_{(r)}}^{+} = \rho_{(r)\,N+1} - \rho_{(r)\,N} \qquad (10)$$

$$f_{(r)}^{-} = \left(\frac{\partial \rho_{(r)}}{\partial N}\right)_{v_{(r)}}^{-} = \rho_{(r)\,N} - \rho_{(r)\,N-1} \qquad (11)$$

To regionalize the electrophilicity and nucleophilicity defined by eqs. (8) and (9), the concept of molecular basin ($\mathrm{N}_{\Omega}^{\pm}$) is introduced [43], which represent the maximum regions of electrophilicity (N_{Ω}^{+}) and nucleophilicity (N_{Ω}^{-}).

$$\mathrm{N}_{\Omega}^{\pm} = \sum_{k \,\in\, \Omega_k^{\pm}} f_k^{\pm} \qquad (12)$$

For the values of N_{Ω}^{+} and N_{Ω}^{-} to be maximal, only those values of $f_k^{\pm} \geq 0.1$ are considered.

The molecular structures considered in this study were optimized using the Gaussian '09 revision E.01 package [44], at the B3LYP/6-31G(d,p) theoretical level. The global reactivity indices were calculated from the HOMO and LUMO energies of each molecule, while the local reactivity indices (Fukui function, FF) were determined using Fukui 4.1 software compatible with Gaussian '09. The electronic properties and reactivity indices of acids and bases can be found in Table 1.

3. RESULTS AND DISCUSSION

From the global and regional reactivity indices presented, the theoretical model of $p\mathrm{K}_a$ and $p\mathrm{K}_b$ from the electronic properties of organic acids and bases, constructed with the multilinear model defined by eq. (2), can be established as follows

$$p\mathrm{K}_i = a_0 + a_1 \mathrm{N}_{\Omega}^{+}\omega^{+} + a_2 \mathrm{N}_{\Omega}^{-}\varepsilon_{\mathrm{H}} + a_3 \mathrm{S} \qquad (13)$$

The electronic properties and reactivity indices of acids and bases can be seen in Table 1

Table 1. Electronic properties of organic acids and bases considered in this study (DFT, B3LYP/6-31G(d,p) level of theory)

Acids/bases	ε_{H} / eV	ε_{L} / eV	N_{Ω}^{+}	N_{Ω}^{-}	ω^{+} / eV	μ / eV	η / eV	S / eV^{-1}
Acetoacetic acid	-4.962	-0.390	0.836	0.761	1.005	-3.433	2.933	0.341
Acetopyruvic acid	-5.666	-1.943	0.926	0.726	2.525	-4.472	1.980	0.505
Formic acid	-5.571	-1.995	0.915	0.660	2.566	-4.853	2.294	0.436
Glyceric acid	-5.302	-0.543	0.522	0.764	1.151	-3.749	3.026	0.331
Glycollic acid	-5.485	-0.853	0.913	0.868	1.391	-4.066	2.971	0.337
Glyoxylic acid	-4.605	-1.342	0.860	0.871	1.738	-3.815	2.093	0.478
Lactic acid	-7.879	3.533	0.544	0.653	0.265	-2.787	7.320	0.137
Pyruvic acid	-5.890	-0.073	0.994	0.884	0.980	-3.825	3.732	0.268
Vinylacetic acid	-5.809	-1.992	0.936	0.787	2.557	-5.004	2.448	0.408
Cronotic acid	-5.485	-0.853	0.913	0.868	1.391	-4.066	2.971	0.337
Oxalic acid	-5.796	-1.759	0.892	0.801	2.268	-4.847	2.589	0.386
Malonic acid	-5.242	0.078	0.901	0.795	0.804	-3.312	3.412	0.293
methyl-malonic acid	-5.115	0.094	0.775	0.784	0.776	-3.220	3.341	0.299
ethyl-malonic acid	-5.107	0.082	0.792	0.782	0.780	-3.224	3.329	0.300
propyl-malonic acid	-5.093	0.110	0.796	0.783	0.765	-3.196	3.338	0.300
isopropyl-malonic acid	-5.078	0.043	0.592	0.784	0.794	-3.229	3.285	0.304
Glutaric acid	-5.568	0.114	0.773	0.755	0.840	-3.499	3.645	0.274
D-tartaric acid	-5.440	-0.270	0.722	0.789	1.011	-3.662	3.313	0.302
Succinic acid	-5.557	0.074	0.753	0.777	0.856	-3.517	3.612	0.277
Adipic acid	-5.584	0.224	0.785	0.762	0.793	-3.438	3.726	0.268
Pimelic acid	-5.547	0.257	0.796	0.759	0.773	-3.393	3.723	0.269
Suberic acid	-5.542	0.270	0.792	0.756	0.767	-3.381	3.728	0.268
Azelaic acid	-5.530	0.288	0.793	0.757	0.758	-3.363	3.732	0.268
methylamine	-4.743	1.652	0.900	0.819	0.240	-1.983	4.102	0.244
ethylamine	-4.714	1.614	0.839	0.804	0.244	-1.988	4.059	0.246
propylamine	-4.660	1.652	0.843	0.784	0.230	-1.929	4.049	0.247
isopropylamine	-4.724	1.534	0.715	0.797	0.261	-2.046	4.014	0.249
cyclohexylamine	-4.767	1.634	0.878	0.792	0.246	-2.010	4.106	0.244
benzylamine	-4.808	0.021	0.891	0.575	0.761	-3.071	3.097	0.323
pyrrole	-4.152	1.170	0.834	0.990	0.268	-1.913	3.413	0.293
pyrrolidine	-4.612	1.567	0.805	0.728	0.241	-1.953	3.963	0.252
imidazole	-4.578	0.723	0.963	0.928	0.458	-2.510	3.438	0.291
pyridine	-5.140	-0.513	0.843	0.708	1.108	-3.627	2.968	0.337
piperidine	-4.570	1.675	0.707	0.738	0.215	-1.857	4.006	0.250

To validate eq. (13), the calculated $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values were contrasted with experimental $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values of a series of acids (carboxylic acids) and organic bases (amine derivatives) described in the literature [45,46,47].

Tables 2 and 3 show the experimental and calculated $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values of a series of organic acids (aliphatic and di-carboxylic acids) and bases

(aliphatic/cyclic and heterocyclic amines). These $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values were calculated from the model described by eq. (13).

Table 2. Experimental and calculated pK_a of series of aliphatic carboxylic and dicarboxylic acids

Serie	Compound	$p\mathrm{K}_\mathrm{a}^{exp}$	$p\mathrm{K}_\mathrm{a}^{cal}$	Error (%)
Aliphatic acids [a]	Acetoacetic acid	3.58	3.685 (0.07)	2.85
	Acetopyruvic acid	2.61	2.753 (0.10)	5.19
	Formic acid	3.77	4.099 (0.23)	8.03
	Glyceric acid	3.55	3.951 (0.28)	10.15
	Glycollic acid	3.82	4.056 (0.17)	5.82
	Glyoxylic acid	3.32	2.990 (0.23)	11.04
	Lactic acid	3.86	3.994 (0.09)	3.36
	Pyruvic acid	2.50	3.042 (0.38)	17.82
	Vinylacetic acid	4.42	3.832 (0.42)	15.34
	Cronotic acid	4.69	3.719 (0.69)	26.11
Aliphatic acids [b]	Acetoacetic acid	3.58	3.548 (0.02)	0.90
	Acetopyruvic acid	2.61	2.613 (0.00)	0.11
	Glyceric acid	3.55	3.697 (0.10)	3.98
	Glycollic acid	3.82	3.742 (0.06)	2.08
	Lactic acid	3.86	3.826 (0.02)	0.89
	Pyruvic acid	2.50	2.494 (0.02)	0.24
dicarboxylic acids	Oxalic acid	1.23	1.142 (0.10)	7.71
	Malonic acid	2.83	3.490 (0.50)	18.91
	methyl-malonic acid	3.05	3.040 (0.01)	0.33
	ethyl-malonic acid	2.99	3.006 (0.01)	0.53
	propyl-malonic acid	3.00	3.027 (0.02)	0.89
	isopropyl-malonic acid	2.94	2.663 (0.20)	10.40
	Glutaric acid	4.34	4.210 (0.10)	3.09
	D-tartaric acid	3.03	3.158 (0.10)	4.05
	Succinic acid	4.19	4.129 (0.04)	1.48
	Adipic acid	4.42	4.447 (0.02)	0.61
	Pimelic acid	4.48	4.417 (0.04)	1.43
	Suberic acid	4.52	4.419 (0.10)	2.29
	Azelaic acid	4.55	4.422 (0.10)	2.89

a: multilinear regression considering all $p\mathrm{K}_a$ values, b: Multilinear regression considering $p\mathrm{K}_a$ in acids with similar chemical characteristics

Table 3. Experimental and calculated pK_b of aliphatic and heterocyclic amine series

Serie	Compound	$p\mathrm{K}_\mathrm{b}^{exp}$	$p\mathrm{K}_\mathrm{b}^{cal}$	Error (%)
Aliphatic/Cyclic amines	methylamine	3.36	3.363 (0.00)	0.09
	ethylamine	3.36	3.367 (0.00)	0.21
	propylamine	3.32	3.325 (0.00)	0.15
	isopropylamine	3.40	3.394 (0.00)	0.18
	cyclohexylamine	3.33	3.322 (0.01)	0.24
	benzylamine	4.67	4.670 (0.00)	0.00
Heterocyclic amines	pyrrole	15.00	14.994 (0.00)	0.04
	pyrrolidine	2.73	2.632 (0.07)	3.72
	imidazole	7.05	7.052 (0.00)	0.03
	pyridine	8.75	8.755 (0.00)	0.06
	piperidine	2.88	2.977 (0.07)	3.26

Graphically, the linear correlations between experimental and calculated $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values presented in Tables 2 and 3 can be seen in the following figures:

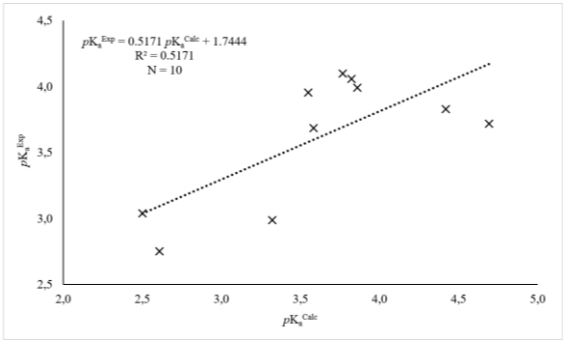

Figure. 1. Linear correlation between experimental and calculated pK_a of the carboxylic acid series: a) pK_a (×) model regression: pK_a^{Exp} = 0.5171 pK_a^{Calc} + 1.7444 with R² = 0.5171. b) pK_a (⁕) model regression: pK_a^{Exp} = 0.9837 pK_a^{Calc} + 0.054 with R² = 0.9837.

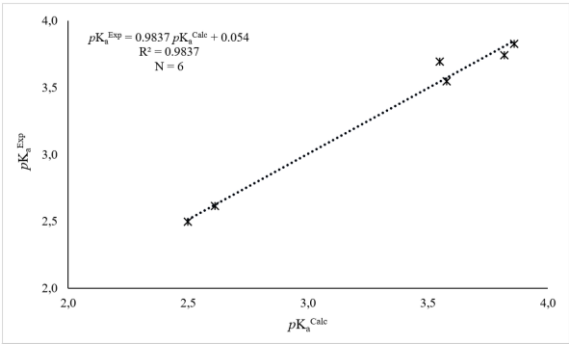

Figure. 2. Linear correlation between experimental and calculated pK_a of the dicarboxylic acid series. pK_a (×) model regression: pK_a^{Exp} = 0.9508 pK_a^{Calc} + 0.1724 with R² = 0.9508.

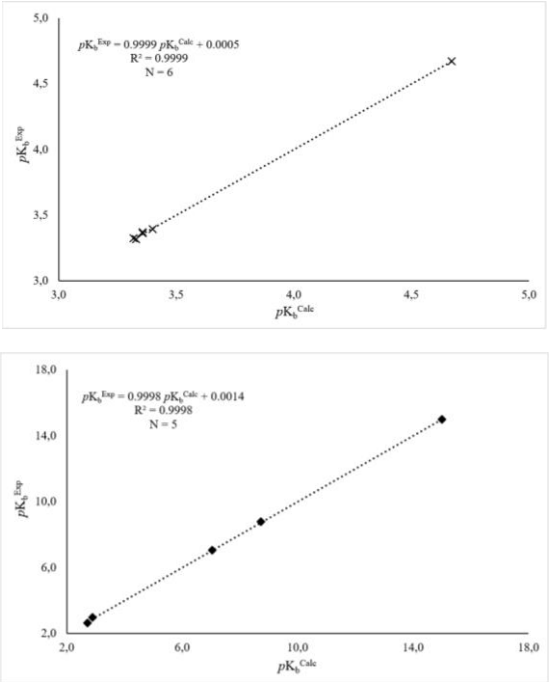

Figure. 3. Linear correlation between experimental and calculated pK_b of aliphatic/cyclic and heterocyclic amine series: a) pK_b (×) model regression: pK_b^{Exp} = 0.9999 pK_b^{Calc} + 0.0005 with R² = 0.9999. b) pK_b (♦) model regression: pK_b^{Exp} = 0.9998 pK_b^{Calc} + 0.0014 with R² = 0.9998.

The graphical representations of the calculated $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values show that the best correlations are generated when the species present similar characteristics or functional groups. This can be seen in Figure 1b, where when extracting compounds with aldehyde (Glyoxylic acid) and alkene (Vinylacetic acid and Cronotic acid) functional groups, the linear regression increases considerably (R^2 = 0.5171 to R^2 = 0.9837). This phenomenon is best evidenced by analyzing the linear correlation of the dicarboxylic acid series (Fig. 2), where, presenting all the same functional groups, the linear regression coefficient is considerably good (R^2 > 0.95). In the same way as the correlation of the dicarboxylic acid series, the amine series shows a very good correlation since both aliphatic/cyclic and heterocyclic amines belong to the same family and do not present any other type of functional group that differentiates them from each other, they only present differences in chain length and ring insaturations.

To validate the model presented, the statistical parameters play a fundamental role, so their values are presented in Table 4:

Table 4. Multilinear regression statistics and coefficients

Serie	R^2	Critical F	a_0	a_1	a_2	a_3
Aliphatic acids [a]	0.5171	0.196	4.613 ± 4.332	– 0.263 ± 0.792	– 0.133± 0.649	– 3.550 ± 8.116
Aliphatic acids [b]	0.9837	0.024	1.704 ± 1.857	– 1.082 ± 0.309	– 0.430 ± 0.331	3.307 ± 2.785
dicarboxylic acids	0.9508	3.287×10^{-6}	13.958 ± 5.818	1.142 ± 0.894	– 0.282 ± 1.101	– 42.558 ± 7.415
Aliphatic amines	0.9999	1.992×10^{-4}	– 2.289 ± 0.592	0.698 ± 0.184	– 0.336 ± 0.070	17.206 ± 1.612
Heterocyclic amines	0.9998	0.017	– 75.937 ± 1.184	– 31.982 ± 0.705	1.635 ± 0.261	357.705 ± 7.079

a: multilinear regression considering all $p\mathrm{K}_a$ values, b: Multilinear regression considering $p\mathrm{K}_a$ in acids with similar chemical characteristics.

The coefficients (in absolute value) of the multilinear correlation associated with the calculated values of $p\mathrm{K}_a$ and $p\mathrm{K}_b$ allow us to establish the weight of electrophilicity(HBD, a_1) and nucleophilicity (HBA, a_2) and their influence on the values of $p\mathrm{K}_a$ and $p\mathrm{K}_b$. As expected, regional electrophilicity (a_1) plays a fundamental role in the acidic characteristics of organic acids (aliphatic and diacid), as reported by Chakkamalayath et. al [48] in their study of reactivity of para-substituted organic acids. Likewise, a considerable influence of polarizability (a_3) can be observed, which is associated with the inductive effects of organic acids and their relative strength [48,49]. For the case of amines, it can be noted that not only nucleophilicity (a_2) plays a role in the $p\mathrm{K}_b$ value, but also electrophilicity (to a greater extent in heterocyclic amines), which can contribute to the $p\mathrm{K}_a$H of their conjugated species. The high value of electrophilicity may be due to its regionalization, since when calculating it, sites are considered that have a greater influence than nucleophilicity, which is mainly associated with nitrogen atoms.

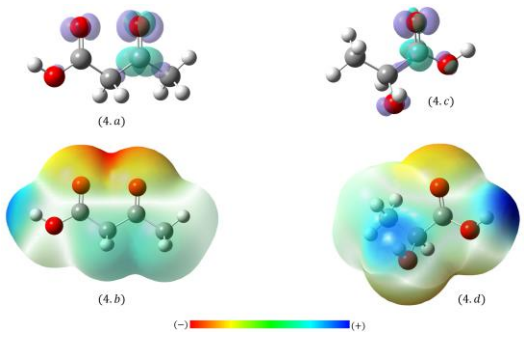

Figure. 4. Electrophilic (cyan) and nucleophilic (purple) basins (fig. 4.*a* and 4.*c*) of monocarboxylic acids (acetoacetic acid/lactic acid). Electrostatic potential map (EPM) showing the electrophilic (low electron density) and nucleophilic (high electron density) zones (fig. 4.*b* and 4.*d*).

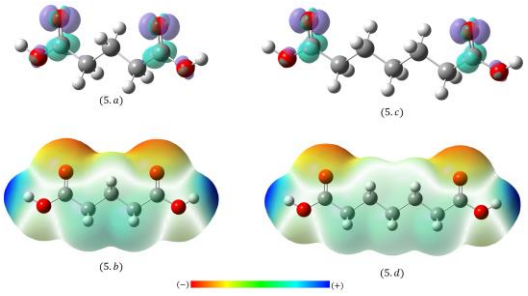

Figure. 5. Electrophilic (cyan) and nucleophilic (purple) basins (fig. 5.*a* and 5.*c*) of dicarboxylic acids (glutaric acid/pimelic acid). Electrostatic potential map (EPM) showing the electrophilic (low electron density) and nucleophilic (high electron density) zones (fig. 5.*b* and 5.*d*).

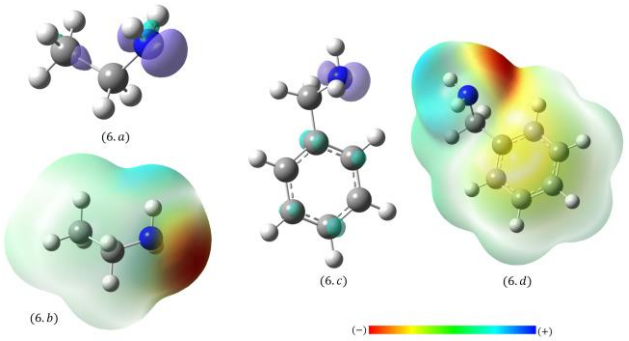

Figure. 6. Electrophilic (cyan) and nucleophilic (purple) basins (fig. 6.*a* and 6.*c*) of aliphatic and cyclic amines (ethylamine/benzylamine). Electrostatic potential map (EPM) showing the electrophilic (low electron density) and nucleophilic (high electron density) zones (fig. 6.*b* and 6.*d*).

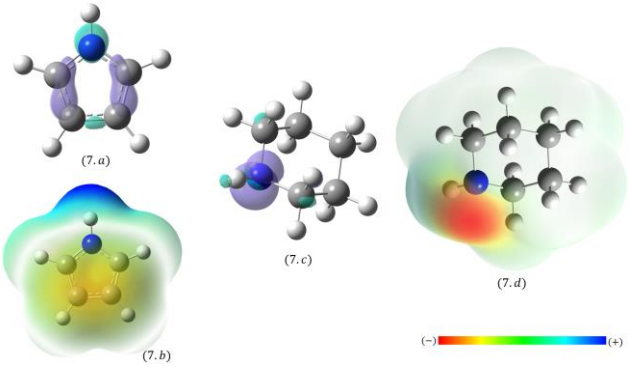

Figure. 7. Electrophilic (cyan) and nucleophilic (purple) basins (fig. 7.*a* and 7.*c*) of heterocyclic amines (pyrrole/piperidine). Electrostatic potential map (EPM) showing the electrophilic (low electron density) and nucleophilic (high electron density) zones (fig. 7.*b* and 7.*d*).

From Figures 4 to 7 (*a* and *c* specifically) one can observe the electrophilic and nucleophilic basins, which show the maximum values of the condensed Fukui. It can be noticed that these are not located on a single atom, but are largely distributed over the adjacent atoms which are believed to concentrate the HBA and HBD features, which shows us that the compounds present bifunctional features with respect to acid-base properties, only that the acidic and basic features are predominant according to the tendency to distribute the electron density over the atoms in a molecule, as can be seen in figures *b* and *d*, where it is evident that in acids there is a zone of minimum electron density on the acidic Hydrogens, while in bases there is a zone of maximum electron density on the sites that present unpaired electronic pairs such as Nitrogen, as described in the Lewis theory.

From the statistical point of view, the models that consider families with similar chemical characteristics present a good linear correlation, which is evident when observing the coefficient of determination (R^2) whose value is equal to or higher than 0.95 ($R^2 \geq 0.95$), which describes the reliability and significance of the model. In the same way, a critical F value equal to or lower than 0.05 represents the significance and predictive capacity of the model [52]. It can be observed that the theoretical models for calculating the $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values of carboxylic and dicarboxylic acids, as well as aliphatic and heterocyclic amines fit this statistical parameter, which indicates the significance of the data considered and their predictive ability. Therefore, the following theoretical $p\mathrm{K}_a$ expressions are presented:

$$p\mathrm{K}_a^{\mathrm{Exp}\,(1)} = 0.9837\, p\mathrm{K}_a^{\mathrm{Calc}\,(1)} + 0.054 \;\; ;$$

$$\text{With: } p\mathrm{K}_a^{\mathrm{Calc}\,(1)} = 1.704 - 1.082\, \mathrm{N}_{\Omega}^{+}\omega^{+} - 0.430\, \mathrm{N}_{\Omega}^{-}\varepsilon_{\mathrm{H}} + 3.307\, \mathrm{S} \qquad (14)$$

$$p\mathrm{K}_a^{\mathrm{Exp}\,(2)} = 0.9508\, p\mathrm{K}_a^{\mathrm{Calc}\,(2)} + 0.1724 \;\; ;$$

$$\text{With: } p\mathrm{K}_a^{\mathrm{Calc}\,(2)} = 13.958 + 1.142\, \mathrm{N}_{\Omega}^{+}\omega^{+} - 0.282\, \mathrm{N}_{\Omega}^{-}\varepsilon_{\mathrm{H}} - 42.558\, \mathrm{S} \qquad (15)$$

Where $p\mathrm{K}_a^{\mathrm{Calc}\,(1)}$ corresponds to the series of carboxylic acids that preferentially present alcohols and ketones in their structures, while $p\mathrm{K}_a^{\mathrm{Calc}\,(2)}$ corresponds to the series of dicarboxylic acids. On the other hand, for the case of theoretical $p\mathrm{K}_b$:

$$p\mathrm{K}_b^{\mathrm{Exp}\,(1)} = 0.9999\, p\mathrm{K}_b^{\mathrm{Calc}\,(1)} + 0.0005 \;\; ;$$

$$\text{With: } p\mathrm{K}_b^{\mathrm{Calc}\,(1)} = -2.289 + 0.698\, \mathrm{N}_{\Omega}^{+}\omega^{+} - 0.336\, \mathrm{N}_{\Omega}^{-}\varepsilon_{\mathrm{H}} + 17.206\, \mathrm{S} \qquad (16)$$

$$p\mathrm{K}_b^{\mathrm{Exp}\,(2)} = 0.9998\, p\mathrm{K}_b^{\mathrm{Calc}\,(2)} + 0.0014 \;\; ;$$

$$\text{With: } p\mathrm{K}_b^{\mathrm{Calc}\,(2)} = -75.937 - 31.982\, \mathrm{N}_{\Omega}^{+}\omega^{+} + 1.635\, \mathrm{N}_{\Omega}^{-}\varepsilon_{\mathrm{H}} + 357.705\, \mathrm{S} \qquad (17)$$

Where $p\mathrm{K}_b^{\mathrm{Calc}\,(1)}$ corresponds to the aliphatic and cyclic amine series, while $p\mathrm{K}_b^{\mathrm{Calc}\,(2)}$ corresponds to the heterocyclic amine series.

It should be emphasized that the limitation of the model is that it is only applicable to series of acids and bases with similar structural characteristics (mono and diprotic acids, and aliphatic and heterocyclic bases). Likewise, the model is applicable for the first value of $p\mathrm{K}_a$ and $p\mathrm{K}_b$. For species with more than one $p\mathrm{K}_a$ and $p\mathrm{K}_b$ value, the same methodology should be performed, but with cationic and/or anionic species.

CONCLUDING REMARKS

Based on what has been presented, it is concluded that the semi-empirical expression correlating $p\mathrm{K}_a$ and $p\mathrm{K}_b$ with the electronic properties of organic acids and bases shows a good relationship between the character HBD with the regional electrophilicity and HBA with the regional nucleophilicity. In the same way, the polarizability described by the global softness plays a weighty role in the acidic and basic properties of the modeled compounds. Based on the statistics, it can be observed that the proposed multilinear correlation model (eq. 13) is able to calculate and predict the $p\mathrm{K}_a$ and $p\mathrm{K}_b$ values of organic acids and bases, specifically if they correspond to series of compounds that present similar chemical properties such as functional groups. Finally, based on the basins and regionalization of electrophilicity and nucleophilicity, it can be noted that the acid-base properties are not specific to a molecular region, but this is of a bifunctional character, only that zones that present a high and low electron density confer the acid-base characteristic and its relative strength to the compounds.

ACKNOWLEDGMENTS

Grateful thanks for the PIDi - UTEM high performance computing system (SCC - PIDi - UTEM CONICYT - FONDEQUIP - EQM180180) of the Universidad Tecnologica Metropolitana, Santiago, Chile, for providing the computational support computational support to perform the computational calculations. We thank Dr. Eduardo Chamorro (UNAB) for providing the Fukui 4.1 software to determine the reactivity indices

REFERENCES

1. Löfqvist, J. *Journal of Insect Physiology*, **22(10)**, 331–1346 (1976)
2. Tutusaus, O., Ni, Ch., Szymczak, N. K. *Journal of the American Chemical Society*, **135(9)**, 3403–3406 (2013)
3. Lémery, L. *Mém. Acad. R. Sci.* (Paris), 1717 ; pp. 31–51, 122–146
4. Rouelle, G. F. *Mémoire sur les sels neutres. Mém. Acad. R. Sci.* (Paris), 1754 ; pp. 572–588
5. Bell, R.P. *Acids, Bases, and the Nature of the Hydrogen Ion. In: The Proton in Chemistry.* Springer, Boston, MA.
6. Arrhenius, S. *Philosophical Magazine Series 5*, **41(251)**, 237–276 (1896)
7. Brønsted, J. N. *Einige Bemerkungen über den Begriff der Säuren und Basen*, **42(8)**, 718–728 (1923)
8. Lowry, T.M. *chemistry and industry*, **42(3)**, 43–47 (1923)
9. Lewis, G.N. *Valence and the Structure of Atoms and Molecules.* American Chemical Monograph Series, the Chemical Catalog Co., Inc., New York, 1923 ; pp. 141-142
10. Koopmans, T. *Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms.* **1(1-6)**, 104–113 (1934).
11. Fukui, K., Yonezawa, T., Shingu, H. *The Journal of Chemical Physics*, **20(4)**, 722–725 (1952)
12. Trummal, A., Lipping, L., Kaljurand, I., Koppel, I. A., Leito, I. *The Journal of Physical Chemistry A*, **120(20)**, 3663-3669, (2016)
13. Paresh Chandra, R., Das Kumar, P. *The Journal of Physical Chemistry*, 99(51), 17891–17895, (1995)
14. Kütt, A., Selberg, S., Kaljurand, I., Tshepelevitsh, S., Heering, A., Darnell, A., Kaupmees, K., Piirsalu, M., Leito, I. *Tetrahedron Letters*, **59(42)**, 3738-3748, (2018)
15. Bernhardsen, Ida M., Krokvik, I. R.T., Perinu, C., Pinto, D., Jens, K. J., Knuutila, H. K. *International Journal of Greenhouse Gas Control*, **68**, 68–76, (2018)
16. Manallack, David T. *Perspectives in Medicinal Chemistry*, **1**, 25-38, (2007)
17. Petukh, M., Stefl, S. Alexov, E. *Current Pharmaceutical Design*, **19(23)**, 4182–4190, (2013)
18. Ullrich, K, J. *Specificity of transporters for 'organic anions' and 'organic cations' in the kidney,* **1197(1)**, 45–62, (1994)
19. Reijenga, J., van Hoof, A., van Loon, A., Teunissen, B. *Analytical Chemistry Insights*, **8**, 53–71, (2013)
20. Romero, R., Salgado, P. R., Soto, C., Contreras, D., Melin, V. *Frontiers in Chemistry*, **6,** 208, (2018)
21. Gift, A. D., Stewart, S. M., Kwete Bokashanga, P. *Journal of Chemical Education*, **89(11)**, 1458–1460, (2012)
22. Jia, Z., Du, D., Zhou, Z., Zhang, A., & Hou, R. *Chemical Physics Letters*, **439(4-6)**, 374–380, (2007)
23. Thapa, B. & Schlegel, H. B. *The Journal of Physical Chemistry A*, **120(28)**, 5726–5735, (2016)
24. Dutra, F. R., Silva, C. de S., & Custodio, R. *The Journal of Physical Chemistry A*, **125(1)**, 65–73, (2020)
25. Sandoval-Lira, J., Mondragón-Solórzano, G., Lugo-Fuentes, L. I., & Barroso-Flores, J. *Journal of Chemical Information and Modeling*, **60**, 1445-1452, (2020)
26. Dahmani, R., Manachou, M., Belaidi, S., Chtita, S., & Boughdiri, S. *New Journal of Chemistry*, **45(3)**, 1253–1261, (2021)
27. Panayiotou, C., Mastrogeorgopoulos, S., & Hatzimanikatis, V. *The Journal of Chemical Thermodynamics*, **110**, 3–15, (2017)
28. Kamlet, M. J., Doherty, R. M., Abraham, M. H., Marcus, Y., & Taft, R. W. *The Journal of Physical Chemistry*, **92(18)**, 5244–5255, (1988)
29. Kamlet, M. J., & Taft, R. W. *Journal of the American Chemical Society*, **98(2)**, 377–383, (1976)
30. Taft, R. W., & Kamlet, M. J. *Journal of the American Chemical Society*, 98(10), 2886–2894, (1976)
31. Parr, R. G., Weitao, Y. *Density – Functional Theory of Atoms and Molecules*. Oxford University Press, USA Chap. 4 – 5, 1994 ; pp. 70 – 104
32. Pearson, R. G. *Journal of Chemical Sciences*, **117(5)**, 369–377, (2005)
33. Parr, R. G., Szentpály, L. v., & Liu, S. *Journal of the American Chemical Society*, **121(9)**, 1922–1924, (1999)
34. Parr, R. G., & Yang, W. *Journal of the American Chemical Society*, **106(14)**, 4049–4050, (1984)
35. Hohenberg, P., & Kohn, W. *Physical Review*, **136(3B)**, B864–B871, (1964)
36. Janak, J. F. *Physical Review B*, **18(12)**, 7165–7168, (1978)

37. Contreras, R., Andres, J., Safont, V. S., Campodonico, P., Santos, J. G. *The Journal of Physical Chemistry A*, **107(29)**, 5588–5593, (2003)
38. Politzer, P., Murray, J. S., Bulat, F. A. *A review*, **16(11)**, 1731–1742, (2010)
39. Weiß, N., Schmidt, C. H., Thielemann, G., Heid, E., Schröder, C., & Spange, S. *Physical Chemistry Chemical Physics*, **23(2)**, 1616–1626, (2021)
40. Vela, A., & Gazquez, J. L. *Journal of the American Chemical Society*, **112(4)**, 1490–1492, (1990)
41. Simón-Manso, Y., & Fuentealba, P. *The Journal of Physical Chemistry A*, **102(11)**, 2029–2032, (1998)
42. Contreras, R. R., Fuentealba, P., Galván, M., & Pérez, P. *Chemical Physics Letters*, **304(5-6)**, 405–413, (1999)
43. Cerda-Monje, A., Ormazábal-Toledo, R., Cárdenas, C., Fuentealba, P., & Contreras, R. *The Journal of Physical Chemistry B*, **118(13)**, 3696–3701, (2014)
44. 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., et al. Gaussian 09, Revision E.01; Wallingford, CT, 2013
45. Dawson, R.M.C., Elliot, D.C., Elliot, W.H. and Jones, K. M. Data for biochemical research. Clarendon Press, Oxford, (1986)
46. Brown, H. C., McDaniel, D. H., & Hafiiger, O. *Determination of Organic Structures by Physical Methods," ed. by Braude E. A. and Nachod, F. C., Academic Press, New York*, 1959
47. Lide, D. R. *Handbook of Chemistry and Physics. Edition 84°, CRC press*, 2004
48. Chakkamalayath, J., Kuttay, C. R. S., & Ghosh, S. K. *The Journal of Physical Chemistry A*, **124**, 3770-3777, (2020)
49. Barret, R. *Therapeutical Chemistry*, 21–51 (2018)
50. Siegel, A. F. & Wagner, M. R. Chapter 12 - Multiple Regression: Predicting One Variable from Several Others, Editor(s): Andrew F. Siegel, Michael R. Wagner, Practical Business Statistics (Eighth Edition), Academic Press, 2022 ; pp. 371-431.