Understanding the Highly Varying p K a of Arylamines. A Perspective from the Average Local Ionization Condensed-to-Atom Framework

2 AUTHORS, INCLUDING:

Eduardo Chamorro
Universidad Andrés Bello
**79** PUBLICATIONS   **1,694** CITATIONS

Article

Understanding the Highly Varying pKa of Arylamines. A Perspective from the Average Local Ionization Condensed-to-Atom Framework

Eduardo Chamorro* and Mario Duque-Noreña

Departamento de Ciencias Quimicas. Facultad de Ciencias Exactas. Nucleus Millennium of Chemical Processes and Catalysis, Universidad Andres Bello, Avenida República 275, 8370146 Santiago, Chile

S *Supporting Information*

ABSTRACT: The highly varying experimental pK_a values for 36 arylamines spanning 7 orders of magnitude is carefully examined. Within this framework, a valence condensed-to-atom model for the average ionization energy is introduced and tested. The theoretical approach is connected to orbital Fukui functions directly mapped into semilocal or regional site-specific responses. It is revealed that the average local ionization energies associated with the amino nitrogen atom is linearly correlated to the basicity of the substituted arylamines, properly reproducing the experimental ordering of basicity. The condensed-to-atom descriptor exhibits a high predictive power, providing a new direct reactivity evaluation of significant value.
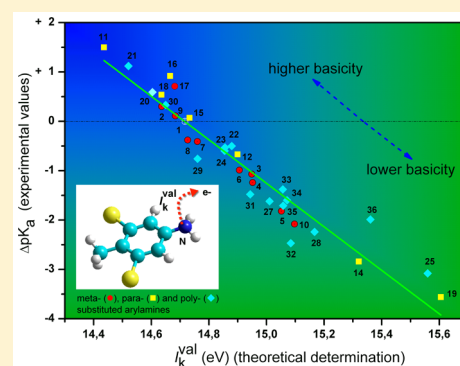
■ INTRODUCTION

A formal approximation to an average per-electron local ionization energy $\overline{I}(r)$ at molecular spatial point r was introduced by Politzer and co-workers[1,2] in terms of the single-particle energy density of orbital i, $\varepsilon_i\rho_i(r)$, and the total electron density $\rho(r)$, eq 1,

$$\overline{I}(\mathbf{r}) = \sum_{i}^{occ} \frac{|\varepsilon_i|\rho_i(\mathbf{r})}{\rho(\mathbf{r})} \tag{1}$$

For a given representation of the electronic molecular structure in terms of Hartree−Fock (HF) or Kohn−Sham (KS) molecular orbitals (MOs) $\psi_i(r)$ with occupation numbers n_i, the electron density $\rho(r)$ is evaluated as a superposition of state densities, namely, $\rho(r) = \Sigma_i^{occ}\rho_i(r) = \Sigma_i^{occ} n_i|\psi_i(r)|^2$. Within such a framework of validity/suitability of using orbital energies[3−5] for describing ionization energies in a molecular system (i.e., $I_i \approx |\varepsilon_i|$), $\overline{I}(r)$ emerges as a useful energetical index of local chemical reactivity[2,6−9] revealing the most labile or reactive electronic sites (i.e., those with the lowest values of ionization energy) to participate in processes involving interactions with electrophile or radical species.[1,2] The outcomes of using orbital energies (both HF and KS) as approximate vertical ionization potentials is well-known[1−5] Certainly, ionization energy is a ubiquitous key quantity entering and driving the behavior of several global (i.e., thermodynamic) and local (i.e., selectivity) density functional theory (DFT)-based chemical reactivity descriptors.[10−14] The average local ionization energy $\overline{I}(r)$ is related to concepts such as local electronic temperatures (i.e., kinetic energy densities), atomic shell structures, electronegativity, local polarizability, hardness, and electrostatic potential.[7,15−18] A proper connection of the $\overline{I}(r)$ index and

the Fukui function,[10,13,19] the key descriptor of local reactivity in DFT, is also known.[20] Recently, the idea of the average local ionization energy[2,7] has been reinterpreted and generalized to correlated wave functions.[21] In studying reactivity trends, $\overline{I}(r)$ is usually computed on a chosen surface (S) defined to be the 0.001−0.002 au (electrons per bohr) contour of $\rho(r)$, that is, $\overline{I}_S(r)$.[2,6] However, though the analysis of local properties on molecular surfaces is appealing,[6] it loses a direct connection with typical concepts associated with the analysis/rationalizing of chemistry in terms of bonding and nonbonding regions.[22−24] Recently, we presented an extension of the applicability of such ideas within the context of the division of molecular regions in basins of attractors provided by the topological analysis of the electron localization function (ELF).[22−26] Such a condensed-to-basin model approximation probed to be useful in the examination of vertical ionization potentials associated with lone pairs and bonding regions.[26] Following ongoing interests in the development and application of condensed or regional reactivity descriptors within an energetical framework,[26−34] we herein introduce and computationally test a model for condensed-to-atom k valence average vertical ionization energy, $\overline{I}_k^{val}$. Our aim is to provide an alternative simplified and direct method to the evaluation of local ionization energies, of complementary value to the traditional mapping[2,6,7] of the local pattern on a molecular surface[6] or even the recently proposed topological-based 3D-grid representations.[26−28,35,36]

THEORETICAL MODEL

Although the definition of $\bar{I}(r)$ given by eq 1 is not amenable to being directly reduced to a condensed-to-atom resolution, we will take an alternative route by considering only valence MOs. In such a context, it is possible to assume that the key averaging kernel $\chi_i(r) \equiv (\rho_i(r)/\rho(r))$ can be approximated as $\chi_k \cong (N_{i,k}/N_k)$, where $N_{i,k} \equiv \int_k \rho_i(r)dr$, and $\bar{N}_k = \Sigma_{i \in \text{val}}^{\text{occ}} \bar{N}_{i,k}$. Several schemes of "condensation" (i.e., integration of the local r-dependent response to a given atomic center) will be thereafter feasible,[37−41,27,28] and for the sake of simplicity we will keep the discussion on the simplest approaches.[42−47,29,28] Within these considerations, a condensed-to-atom k valence average vertical ionization energy $\bar{I}_k^{\text{val}}$ can be drawn from valence MOs $\psi_i(r)$ expressed in terms of a linear combination of N_b basis functions $\{\chi_\mu(r)\}$:

$$\bar{I}_k^{\text{val}} \equiv \sum_{i \in \text{val}}^{\text{occ}} \frac{|\varepsilon_i| \bar{N}_{i,k}}{\bar{N}_k} \qquad (2)$$

where

$$\bar{N}_{i,k} = n_i \sum_{\mu \in k}^{N_b} \sum_{\nu}^{N_b} c_\mu^i c_\nu^i S_{\mu\nu} \qquad (3)$$

The coefficients c_μ^i correspond to the eigenvectors of MO $\psi_i(r)$, and $S_{\mu\nu}$ represents the overall overlap matrix elements, that is, $S_{\mu\nu} \equiv \int \chi_\mu(r)\chi_\nu(r)dr$. Note immediately from eq 3 that the MO contribution population to atom k, $\bar{N}_{i,k}$ reduces in this simple case to the condensed-to-atom Fukui shape function[46,47] scaled by the occupation of the ith valence MO, namely, $\bar{N}_{i,k} = n_i f_k^{(i)}$ with $f_k^{(i)} \equiv \Sigma_{\mu \in k}^{N_b} \Sigma_\nu^{N_b} c_\mu^i c_\nu^i S_{\mu\nu}$. It is important to remark that, as it is also the case for the integral of $\bar{I}(r)$ over all space,[2,6,7,15,48] the approximated condensed-to-atom descriptors in eq 2 obeys no normalization condition when summing on all atoms. Inspired in ref 49, the classical problem of correlating the highly varying pK_a values of a set of 36 substituted anilines offers an example of wide chemical interest for which several experimental[50−52] and theoretically related[53−57] evidence is available. Gross and co-workers[49] have emphasized that the prediction of experimental pK_a[51,52] values of this particular set of anilines can be achieved by using parameters derived from quantum mechanical calculations, including natural charges, relative proton transfer enthalpies, and the minimum values on the molecular surface of both electrostatic potentials $\bar{V}(r)$ and average local ionization energies $\bar{I}(r)$. It was shown that these quantities perform as well as the traditional (but empirical) use of Hammett substituent parameters. It was also remarked that $\bar{I}_{S,\text{min}}(r)$ correlated better than the other theoretical descriptors.[49]

COMPUTATIONAL DETAILS

Optimized structures for the set of substituted anilines included in Table 1 were calculated using the Gaussian 09 (G09)[58] package of programs at the B3LYP/6-311G(d,p) level of theory. Characterization as a minimum on the potential energy surfaces was verified by analytical frequency calculations. Water solvent effects were incorporated using the simplest polarizable continuum model (PCM). Simple computational routines for the evaluation of $\bar{I}_k^{\text{val}}$ were implemented for processing results from single-point G09 calculations. These routines are available from the authors upon request.

Table 1. Theoretical Average Ionization Energies $\bar{I}_k^{\text{val}}$ and $\bar{I}_{S,\text{min}}(r)$ and Experimental pK_a Values for Aniline (01) and Some Substituted Derivatives (02−36)

	substituent	$\bar{\bar{I}}_k^{\text{val}\,a}$	$\bar{\bar{I}}_{S,\text{min}}(r)^b$	$pK_a^{\,c}$
01	H (aniline)	14.717	13.33	4.58
02	m-NH$_2$	14.636	13.32	4.88
03	m-Br	14.949	13.71	3.51
04	m-Cl	14.952	13.71	3.34
05	m-CN	15.052	13.91	2.76
06	m-F	14.906	13.67	3.59
07	m-OH	14.760	13.52	4.17
08	m-OCH$_3$	14.726	13.41	4.20
09	m-CH$_3$	14.684	13.31	4.69
10	m-NO$_2$	15.098	14.01	2.50
11	p-NH$_2$	14.436	12.96	6.08
12	p-Br	14.899	13.69	3.91
13	p-Cl	14.887	13.66	3.98
14	p-CN	15.320	14.26	1.74
15	p-F	14.733	13.35	4.65
16	p-OH	14.666	13.08	5.50
17	p-OCH$_3$	14.682	13.06	5.29
18	p-CH$_3$	14.635	13.22	5.12
19	p-NO$_2$	15.605	14.53	1.02
20	3-CH$_3$, 4-CH$_3$	14.604	13.17	5.17
21	3-NH$_2$, 4-OH	14.521	13.27	5.70
22	3-Br, 4-OCH$_3$	14.879	13.33	4.08
23	3-Br, 4-CH$_3$	14.855	13.55	3.98
24	3-Cl, 4-CH$_3$	14.856	13.54	4.05
25	3-CH$_3$, 4-NO$_2$	15.559	14.49	1.50
26	4-Cl, 3-NO$_2$	15.250	14.28	1.90
27	4-CH$_3$, 3-NO$_2$	15.011	13.84	2.96
28	3-Br, 5-Br	15.167	14.03	2.34
29	3-OCH$_3$, 5-OCH$_3$	14.761	13.42	3.82
30	3-CH$_3$, 5-CH$_3$	14.650	13.24	4.91
31	3-Cl, 5-OCH$_3$	14.943	13.73	3.10
32	3-OCH$_3$, 5-NO$_2$	15.084	14.03	2.11
33	3-Br, 5-Br, 4-OH	15.057	13.64	3.20
34	3-Br, 5-Br, 4-OCH$_3$	15.072	13.85	2.98
35	3-Br, 5-Br, 4-CH$_3$	15.058	13.87	2.87
36	3-CH$_3$, 5-CH$_3$, 4-NO$_2$	15.361	14.23	2.59

^{a}This work. ^{b}Taken from ref 49 evaluated at the HF/STO-5G(d)//HF/6-311G(d,p) level of theory. Minimum value is reported to be associated with the amino nitrogen in all cases along the series. ^{c}Experimental data51 and 52 in water at 298.15 K. Ionization energies in electronvolts.

RESULTS AND DISCUSSION

It is well-known that amines are weak bases, and the relative basicity is commonly discussed in terms of the pK_a of its conjugate acid (i.e., the more basic the amine, the higher the pK_a of its conjugate acid). Arylamines are weaker bases than the corresponding nonaromatic analogues. The nitrogen lone pair is less available for protonation given its delocalization to the ortho and para positions of the ring. Therefore, higher values of ionization energy will be associated with the amino nitrogen center in the case of weaker bases. For each aniline derivative, Table 1 reports the experimental pK_a values in water at 298 K[51,52] and the theoretical average ionization energies $\bar{I}_k^{\text{val}}$ (from eqs 2 and 3) as well as $\bar{I}_{S,\text{min}}(r)$[49] evaluated at the nitrogen atom of the basic amino group. To explore the validity of the proposed condensed-to-atom model on both large and restricted sets, analysis was also performed on chemically
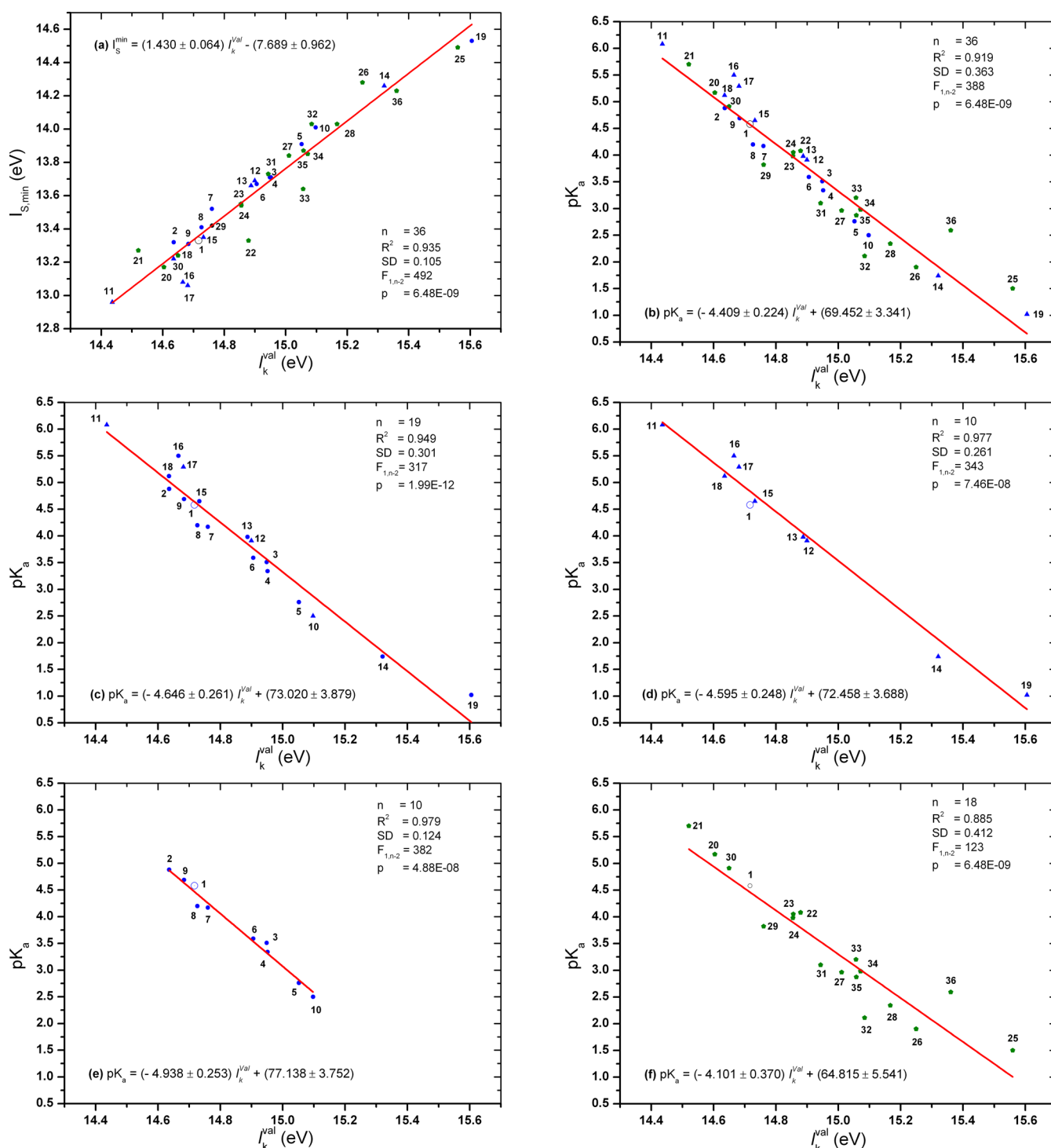
Figure 1. Linear correlation plots for (a) $\overline{I}_{S,min}(r)$ vs $\overline{I}_k^{val}$ and (b–f) experimental pK_a values vs $\overline{I}_k^{val}$. R^2, SD, $F_{1,n-2}$, and p are the correlation coefficient, standard deviation, ANOVA F-test, and p-value statistics, respectively, for n compounds. The reported linear relationship is valid for arylamines as follows: (a, b): 01–36; (c): 01–19; (d): 01, 11–19; (e): 01–10; and (f): 01, 20–36. Symbols ●, ▲, and ◆ represent the anilines with meta-, para-, and polysubstitution patterns, respectively.

significant groups of these systems, including subsets integrated by only mono-, para-, meta-, and polysubstituted anilines. Figure 1a–f shows plot results for linear correlation analysis of (a) $\overline{I}_{S,min}(r)$ versus $\overline{I}_k^{val}$ and (b–f) experimental pK_a values versus $\overline{I}_k^{val}$, including main statistics parameters, that is, R^2, SD, $F_{1,n-2}$, and p, which are the correlation coefficient, standard deviation, ANOVA F-test statistic, and the corresponding p-value, respectively. n is the number of compounds in the set of

correlation. Note immediately from the R^2 values that the level of confidence for the model correlations ranges from 88.5 to 97.9% with p-values (associated with the $F_{1,n-2}$ statistics) much less than 0.05 in all cases. This implies that there is a significant linear correlation between $\overline{I}_k^{val}$ and $\overline{I}_{S,min}(r)$ and between $\overline{I}_k^{val}$ and pK_a. Low values of the local ionization energy associated with the amino nitrogen atom in the aniline derivative become properly and significantly associated with less basic arylamines.

Results reveal that the simple energetical reactivity descriptor $\overline{I}_k^{\mathrm{val}}$ correctly reproduces the experimental ordering of basicity along this series spanning 7 orders of magnitude in reactivity, bounded by p-NO$_2$-Ar (19, pK$_a$ = 1.02) and p-NH$_2$-Ar (11, pK$_a$ = 6.08). The predicted $\overline{I}_k^{\mathrm{val}}$ values reveals also proper reactivity relationships both on the side of less basic aniline derivatives (i.e., ΔpK$_a$ < 0, $\Delta\overline{I}_k^{\mathrm{val}}$ > 0) as well as on the side of more basic compounds (i.e., ΔpK$_a$ > 0, $\Delta\overline{I}_k^{\mathrm{val}}$ < 0) with respect to the aniline (01) system, taken as a reference for comparison. For instance, note that decreased basicity introduced by p-Br- (ΔpK$_a$ = −0.67, $\Delta\overline{I}_k^{\mathrm{val}}$ = +0.36 eV) or p-Cl- (ΔpK$_a$ = −0.60, $\Delta\overline{I}_k^{\mathrm{val}}$ = +0.33 eV) substituents is appropriately predicted to be lower than those associated with p-CN- (ΔpK$_a$ = −2.84, $\Delta\overline{I}_k^{\mathrm{val}}$ = +0.93 eV) or p-NO$_2$- (ΔpK$_a$ = −3.56, $\Delta\overline{I}_k^{\mathrm{val}}$ = +1.20 eV). Note also that cooperative electronic effects of substituents can be properly assessed by our simple model as revealed for instance in the cases of decreasing basicity in polysubstituted derivatives (25), (32), or (33) for which ΔpK$_a$ = −3.08, −2.47, and −1.38, respectively, with $\Delta\overline{I}_k^{\mathrm{val}}$ = +1.16 eV, +0.70, and +0.31 eV, respectively. Correspondingly, in the case of increasing basicity, the pattern of substituent is again suitably incorporated by the simple model given in eq 2. Note for instance that the observed experimental ordering of basicity in monosubstituted systems, namely, m-CH$_3$ < p-CH$_3$ < p-OCH$_3$ < p-NH$_2$, is correctly predicted, that is, (ΔpK$_a$ = +0.11, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.02 eV), (ΔpK$_a$ = +0.54, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.11 eV), (ΔpK$_a$ = +0.70, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.27 eV), (ΔpK$_a$ = +1.50, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.37 eV), respectively. As before, the cooperative effects of multiple substituents can be also properly incorporated and properly assessed in the global electronic response.

Compare for instance the cases of systems (30) and (20) for which (ΔpK$_a$ = +0.33, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.09 eV) and (ΔpK$_a$ = +0.59, $\Delta\overline{I}_k^{\mathrm{val}}$ = −0.16 eV), respectively. It can therefore be concluded that the simple $\overline{I}_k^{\mathrm{val}}$ descriptor features in good order the well-known inductive and resonance effects of common chemical substituents along this series spanning a so long-range of order of magnitude in pK$_a$. Naturally, and as exemplified in Figure 1c−e, these effects become better represented when shorter series including only monosubstituted derivatives are correlated. In these cases, the variation of pK$_a$ can be confidently represented by $\overline{I}_k^{\mathrm{val}}$ in a 97.7% and 97.9% SD for meta- and para-substituted series, respectively. However, and as it was also the case for $\overline{I}_{S,\mathrm{min}}(r)$,[49] Figure 1f reveals that the significance of such a representation is less satisfactory (ca. 88.5%) for a series of polysubstituted derivatives. These results contribute to validate the condensed-to-atom model $\overline{I}_k^{\mathrm{val}}$ as a useful tool correlating the aqueous acidity trends of substituted anilines with essentially the same predictive power of the traditional approach via $\overline{I}_{S,\mathrm{min}}(r)$.[49] Note that within an orbital-based framework, the partition presented in eq 3 represents the simplest way to interpret and computationally implement electronic population analysis.[42−47,29,28]

Equations eqs 2 and 3 constitute a simple representation of the average ionization energy within a valence condensed-to-atom resolution (i.e., semilocal or regional site-specific) providing a tool of high predictive power. For the selected 36 substituted aniline derivatives, such an approach for $\overline{I}_k^{\mathrm{val}}$, as it is for $\overline{I}_{S,\mathrm{min}}(r)$, linearly correlates with the experimental pK$_a$ values (water, 298 K) associated with the conjugated acids of this series of anilines. These results further extend and broaden the relevance of the regional average ionization potential as a useful index for chemical reactivity, providing a direct computational tool of complementary value to $\overline{I}_{S,\mathrm{min}}(r)$[2,6,7,15,27,48] within a simplified orbital-based but chemically significant condensed-to-atom framework.

Cross-Validation. We must note that the R-squared value measures the goodness of fit on existing data. However, to measure if the linear regression models can be confidently applied to predict future values in each case, cross-validation (CV) techniques are needed.[59−67] Such techniques provide a measure of the prediction power of the fitted models. Within this context, we evaluated the predicted residual sum of squares (PRESS) statistic associated with the *leave-one-out* CV method. In that case, for a given fitted model, each of the predictors $\{x_i\}_{i=1...n}$ is removed (i.e., testing set), and the model is refitted to the $n-1$ remaining points (i.e., the training set), allowing the prediction of the excluded points $\hat{y}_{[i,-i]}$. This is performed iteratively until all points have served as the test case. The PRESS statistic simply corresponds to PRESS $= (1/n)\sum_{i=1}^{n}[y_i - \hat{y}_{[i,-i]}]^2$. For samples of small size (as our present case), this index constitutes indeed a measure of internal validity. Small values of PRESS indicate that the model is not overly sensitive to any single data point and, hence, provides a higher predictive power. Complementarily, we also performed *k-fold CV* on the different models. In that case the CV procedure randomly divides the data into k partitions or folds for the testing set, modeling all the remaining (training) data. Hence, the mean squared error (MSE), based on the testing-set sum of errors (i.e., residuals) for all folds provides a measure of the prediction power for the fitted model.[61,65−69] The CV analyses were performed using the R software environment for statistical computing and graphics.[68,69] Main results are reported in Table 2.

Relevant information can be highlighted from Table 2. Note for instance that, given the RMSE has the same units of the predicted quantity (i.e., the quantity plotted on the vertical axis), it can be observed that the proposed condensed-to-atom model $\overline{I}_k^{\mathrm{val}}$ and the $\overline{I}_{S,\mathrm{min}}(r)$ shares vis-à-vis essentially a higher predictive power in all cases. Certainly, the $\overline{I}_{S,\mathrm{min}}(r)$ exhibits a

Table 2. Predicted Residual Sum of Squares (PRESS)[a] and Overall Mean Squared Error (MSE)[b] for Linear Models Correlating the Experimental pK$_a$ Values of Arylamines to Theoretical Models of the Average Ionization Energy ($\overline{I}_k^{\mathrm{val}}$ and $\overline{I}_{S,\mathrm{min}}(r)$)

data set[c]	pK$_a$ vs $\overline{I}_k^{\mathrm{val}}$		pK$_a$ vs $\overline{I}_{S,\mathrm{min}}(r)$		$\overline{I}_{S,\mathrm{min}}(r)$ vs $\overline{I}_k^{\mathrm{val}}$	
	PRESS	MSE	PRESS	MSE	PRESS	MSE
I	5.260	0.1470	3.570	0.0974	0.148	0.0116
		0.3340		0.2843		0.0839
II	2.430	0.1400	0.605	0.0324	0.229	0.0131
		0.2862		0.1530		0.0538
III	1.150	0.1150	0.450	0.0450	0.172	0.0172
		0.2675		0.1694		0.1038
IV	0.181	0.0181	0.172	0.0172	0.025	0.0025
		0.1071		0.0931		0.0300
V	3.740	0.2010	3.250	0.1880	0.262	0.0148
		0.4141		0.4051		0.0848

[a]Associated to leave-one-out cross validation statistics. [b]Associated to the k-fold cross validation statistics. The commonly accepted $k = 10$ value has been applied.[67] The root mean squared error (RMSE) is reported as second entry in the MSE column. [c]I: the entire set of studied anilines ($n = 36$); II: the subset of monosubstituted anilines ($n = 19$), III: the subset of para-substituted anilines ($n = 10$), IV: the subset of meta-substituted anilines ($n = 10$), and V: the subset of polysubstituted anilines.

slightly higher predictive power than the condensed-to-atom proposal $\overline{I}_k^{val}$, but differences in prediction will be always within the same order of magnitude in all cases. This is clear by comparing the low values of PRESS or equivalently the MSE, or the very low standard deviations for the model prediction error (RMSE) values for pK_a versus $\overline{I}_k^{val}$ and pK_a versus $\overline{I}_{S,min}(r)$ along the different family of series. Indeed it can be learned that the predictive power for each index is the same in regard to the type of substitution pattern being examined, namely, meta- < para- < mono- < poly- substituted anilines. The analysis of the variation of CV results along the series emphasizes that both $\overline{I}_k^{val}$ and the $\overline{I}_{S,min}(r)$ are particularly valuable in predicting pK_a on meta-substituted anilines in all cases (RMSE ≈ 0.1). Note that in the case of examination of meta- and para-substituted series of anilines, the k-fold CV corresponds in fact to a leave-one-out CV procedure.[67−69] Clearly, both $\overline{I}_k^{val}$ and the $\overline{I}_{S,min}(r)$ properly assess better the inductive effects of substituents on the reactive basic center. Finally, note that a direct comparison of the excellent performance of $\overline{I}_k^{val}$ in predicting the $\overline{I}_{S,min}$ values (see last two columns in Table 2) provides certainly confidence in the use of the approximated condensed-to-atom model of eqs 2 and 3 in proper estimation of *local* ionization energies within a chemical reactivity framework.

■ CONCLUSION

We would like finally to comment that the applicability of a general condensed-to-atom resolution is certainly not free of additional scrutiny.[37−41] Different approaches are certainly feasible, including Hirshfeld,[70,38,71−73] Löwdin,[74] and NPA[73,75] partitions, for instance. Hence, despite gross approximations behind eqs 2 and 3, the predictive value of $\overline{I}_k^{val}$ for the experimental values pK_a along this series is certainly remarkable! The performance of the predictive models in each case has been properly characterized to be associated with small using both exhaustive leave-one-out CV method as well as the nonexhaustive k-fold CV statistic ($k = 10$). These findings are relevant because the condensed-to-atom model represents a simplification in the evaluation/interpretation of the chemical response as compared to more complex tasks associated with the evaluation of $\overline{I}_{S,min}(r)$,[2,6,7,15,27] that is, mapping and minimum search for the local pattern on a given surface of electron density. Certainly, it is expected that the applicability of the proposed model $\overline{I}_k^{val}$ as a reactivity tool be readily extendable to other problems of general interest in chemistry. An example of current interest in such a context is provided by the definition and application of *nucleophilicity* descriptors where ionization energy, both global and local, plays a relevant role.[32,34,76−83]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Plots for each fold's predictor values (x axis) against the actual response variable (y axis). Each fold is represented in a different color, for the k-fold CV statistics for linear regression model of pK_a versus $\overline{I}_k^{val}$, pK_a versus $\overline{I}_{S,min}(r)$, and $\overline{I}_{S,min}(r)$ versus $\overline{I}_k^{val}$, for the studied series of substituted anilines. Cartesian coordinates (ready for Gaussian 09 input) for aniline (01) and substituted derivatives (02−36) characterized in this work at the B3LYP/6-311G(d,p) level of theory in water solvent using PC. This material is available free of charge via the Internet at http://pubs.acs.org. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.5b03252.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: echamorro@unab.cl. Phone: +56 2 661 82 29.

Author Contributions
The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Murray, J. S.; Seminario, J. M.; Politzer, P.; Sjoberg, P. Average Local Ionization Energies Computed on the Surfaces of Some Strained Molecules. *Int. J. Quantum Chem.* **1990**, 645−653.

(2) Politzer, P.; Murray, J. S.; Bulat, F. A. Average Local Ionization Energy: A Review. *J. Mol. Model.* **2010**, *16*, 1731−1742.

(3) Chong, D. P.; Gritsenko, O. V.; Baerends, E. J. Interpretation of the Kohn-Sham Orbital Energies as Approximate Vertical Ionization Potentials. *J. Chem. Phys.* **2002**, *116*, 1760−1772.

(4) Zhan, C. G.; Nichols, J. A.; Dixon, D. A. Ionization Potential, Electron Affinity, Electronegativity, Hardness, and Electron Excitation Energy: Molecular Properties from Density Functional Theory Orbital Energies. *J. Phys. Chem. A* **2003**, *107*, 4184−4195.

(5) Zhang, G.; Musgrave, C. B. Comparison of Dft Methods for Molecular Orbital Eigenvalue Calculations. *J. Phys. Chem. A* **2007**, *111*, 1554−1561.

(6) Bulat, F. A.; Toro-Labbe, A.; Brinck, T.; Murray, J. S.; Politzer, P. Quantitative Analysis of Molecular Surfaces: Areas, Volumes, Electrostatic Potentials and Average Local Ionization Energies. *J. Mol. Model.* **2010**, *16*, 1679−1691.

(7) Bulat, F. A.; Levy, M.; Politzer, P. Average Local Ionization Energies in the Hartree-Fock and Kohn-Sham Theories. *J. Phys. Chem. A* **2009**, *113*, 1384−1389.

(8) Politzer, P.; Murray, J. S.; Concha, M. C. The Complementary Roles of Molecular Surface Electrostatic Potentials and Average Local Ionization Energies with Respect to Electrophilic Processes. *Int. J. Quantum Chem.* **2002**, *88*, 19−27.

(9) Murray, J. S.; Brinck, T.; Grice, M. E.; Politzer, P. Correlations between Molecular Electrostatic Potentials and Some Experimentally-Based Indexes of Reactivity. *J. Mol. Struct.: THEOCHEM* **1992**, *88*, 29−45.

(10) Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual Density Functional Theory. *Chem. Rev.* **2003**, *103*, 1793−1873.

(11) Chattaraj, P. K.; Sarkar, U.; Roy, D. R. Electrophilicity Index. *Chem. Rev.* **2006**, *106*, 2065−2091.

(12) Chamorro, E.; Chattaraj, P. K.; Fuentealba, P. Variation of the Electrophilicity Index Along the Reaction Path. *J. Phys. Chem. A* **2003**, *107*, 7068−7072.

(13) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. Perturbative Perspectives on the Chemical Reaction Prediction Problem. *Int. J. Quantum Chem.* **2005**, *101*, 520−534.

(14) Cohen, M. H.; Wasserman, A. On the Foundations of Chemical Reactivity Theory. *J. Phys. Chem. A* **2007**, *111*, 2229−2242.

(15) Politzer, P.; Shields, Z. P.-I.; Bulat, F. A.; Murray, J. S. Average Local Ionization Energies as a Route to Intrinsic Atomic Electronegativities. *J. Chem. Theory Comput.* **2011**, *7*, 377−384.

(16) Jin, P.; Murray, J. S.; Politzer, P. Computational Determination of the Relative Polarizabilities of Molecular Components. *Int. J. Quantum Chem.* **2006**, *106*, 2347−2355.

(17) Politzer, P.; Murray, J. S.; Grice, M. E. Electronegativity and Average Local Ionization Energy. *Collect. Czech. Chem. Commun.* **2005**, *70*, 550−558.

(18) Chattaraj, P. K.; Chamorro, E.; Fuentealba, P. Chemical Bonding and Reactivity: A Local Thermodynamic Viewpoint. *Chem. Phys. Lett.* **1999**, *314*, 114−121.

(19) Ayers, P. W.; Parr, R. G. Variational Principles for Describing Chemical Reactions: The Fukui Function and Chemical Hardness Revisited. *J. Am. Chem. Soc.* **2000**, *122*, 2010−2018.

(20) Toro-Labbe, A.; Jaque, P.; Murray, J. S.; Politzer, P. Connection between the Average Local Ionization Energy and the Fukui Function. *Chem. Phys. Lett.* **2005**, *407*, 143−146.

(21) Ryabinkin, I. G.; Staroverov, V. N. Average Local Ionization Energy Generalized to Correlated Wavefunctions. *J. Chem. Phys.* **2014**, *141*, 084107.

(22) Silvi, B.; Savin, A. Classification of Chemical-Bonds Based on Topological Analysis of Electron Localization Functions. *Nature* **1994**, *371*, 683−686.

(23) Savin, A.; Nesper, R.; Wengert, S.; Fassler, T. F. Elf: The Electron Localization Function. *Angew. Chem., Int. Ed.* **1997**, *36*, 1809−1832.

(24) Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; Vonschnering, H. G. A New Look at Electron Localization. *Angew. Chem., Int. Ed.* **1991**, *30*, 409−412.

(25) Becke, A. D.; Edgecombe, K. E. A Simple Measure of Electron Localization in Atomic and Molecular-Systems. *J. Chem. Phys.* **1990**, *92*, 5397−5403.

(26) Chamorro, E.; Perez, P.; Duque-Norena, M.; Paul Romagnoli, P.; Pons, D.; Gonzalez, M. Average Local Ionization Potential within the Framework of the Electron Localization Function. *Indian J. Chem., Sect. A* **2014**, *53*, 958−964.

(27) Bulat, F. A.; Chamorro, E.; Fuentealba, P.; Toro-Labbe, A. Condensation of Frontier Molecular Orbital Fukui Functions. *J. Phys. Chem. A* **2004**, *108*, 342−349.

(28) Tiznado, W.; Chamorro, E.; Contreras, R.; Fuentealba, P. Comparison among Four Different Ways to Condense the Fukui Function. *J. Phys. Chem. A* **2005**, *109*, 3220−3224.

(29) Chamorro, E.; Perez, P. Condensed-to-Atoms Electronic Fukui Functions within the Framework of Spin-Polarized Density-Functional Theory. *J. Chem. Phys.* **2005**, *123*, 114107.

(30) Chamorro, E.; Perez, P.; De Proft, F.; Geerlings, P. Philicity Indices within the Spin-Polarized Density-Functional Theory Framework. *J. Chem. Phys.* **2006**, *124*, 044105.

(31) Chamorro, E.; Perez, P.; Duque, M.; De Proft, F.; Geerlings, P. Dual Descriptors within the Framework of Spin-Polarized Density Functional Theory. *J. Chem. Phys.* **2008**, *129*, 064117.

(32) Perez, P.; Domingo, L. R.; Duque-Norena, M.; Chamorro, E. A Condensed-to-Atom Nucleophilicity Index. An Application to the Director Effects on the Electrophilic Aromatic Substitutions. *J. Mol. Struct.: THEOCHEM* **2009**, *895*, 86−91.

(33) Chamorro, E.; Perez, P.; Domingo, L. R. On the Nature of Parr Functions to Predict the Most Reactive Sites Along Organic Polar Reactions. *Chem. Phys. Lett.* **2013**, *582*, 141−143.

(34) Alcoba, D. R.; Lain, L.; Torre, A.; Ona, O. B.; Chamorro, E. Fukui and Dual-Descriptor Matrices within the Framework of Spin-Polarized Density Functional Theory. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9594−9604.

(35) Soler, P.; Fuster, F.; Chevreau, H. Fast Topological Analysis of 2d and 3d Grids of Data: Application to the Atoms in Molecule (Aim) and the Electron Localization Function (Elf). *J. Comput. Chem.* **2004**, *25*, 1920−1925.

(36) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. Computational Tools for the Electron Localization Function Topological Analysis. *Comput. Chem.* **1999**, *23*, 597−604.

(37) Nalewajski, R. F.; Parr, R. G. Information Theory, Atoms in Molecules, and Molecular Similarity. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8879−8882.

(38) Ayers, P. W. Atoms in Molecules, an Axiomatic Approach. I. Maximum Transferability. *J. Chem. Phys.* **2000**, *113*, 10886−10898.

(39) Ayers, P. W.; Morrison, R. C.; Roy, R. K. Variational Principles for Describing Chemical Reactions: Condensed Reactivity Indices. *J. Chem. Phys.* **2002**, *116*, 8731−8744.

(40) Bultinck, P.; Fias, S.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. Critical Thoughts on Computing Atom Condensed Fukui Functions. *J. Chem. Phys.* **2007**, *127*, 034102.

(41) Zielinski, F.; Tognetti, V.; Joubert, L. Condensed Descriptors for Reactivity: A Methodological Study. *Chem. Phys. Lett.* **2012**, *527*, 67−72.

(42) Mulliken, R. S. Electronic Population Analysis on Lcao-Mo Molecular Wave Functions. Iv. Bonding and Antibonding in Lcao and Valence-Bond Theories. *J. Chem. Phys.* **1955**, *23*, 2343−2346.

(43) Mulliken, R. S. Electronic Population Analysis on Lcao-Mo Molecular Wave Functions. Ii. Overlap Populations, Bond Orders, and Covalent Bond Energies. *J. Chem. Phys.* **1955**, *23*, 1841−1846.

(44) Mulliken, R. S. Electronic Population Analysis on Lcao-Mo Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(45) Yang, W.; Mortier, W. J. The Use of Global and Local Molecular Parameters for the Analysis of the Gas-Phase Basicity of Amines. *J. Am. Chem. Soc.* **1986**, *108*, 5708−5711.

(46) Contreras, R. R.; Fuentealba, P.; Galvan, M.; Perez, P. A Direct Evaluation of Regional Fukui Functions in Molecules. *Chem. Phys. Lett.* **1999**, *304*, 405−413.

(47) Fuentealba, P.; Perez, P.; Contreras, R. On the Condensed Fukui Function. *J. Chem. Phys.* **2000**, *113*, 2544−2551.

(48) Murray, J. S.; Shields, Z. P.-I.; Lane, P.; Macaveiu, L.; Bulat, F. A. The Average Local Ionization Energy as a Tool for Identifying Reactive Sites on Defect-Containing Model Graphene Systems. *J. Mol. Model.* **2013**, *19*, 2825−2833.

(49) Gross, K. C.; Seybold, P. G.; Peralta-Inga, Z.; Murray, J. S.; Politzer, P. Comparison of Quantum Chemical Parameters and Hammett Constants in Correlating Pk(a) Values of Substituted Anilines. *J. Org. Chem.* **2001**, *66*, 6919−6925.

(50) Bryson, A. The Effects of M-Substituents on the Pk, Values of Anilines, and on the Stretching Frequencies of the N-H Bonds. *J. Am. Chem. Soc.* **1960**, *82*, 4859−4862.

(51) Haynes, W. M., Ed., *CRC Handbook of Chemistry and Physics*, 94th ed.; CRC Press Taylor & Francis Group: 2013.

(52) Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *Pka Prediction for Organic Acids and Bases*, 1st ed.; Springer: Netherlands, 1981; p 108.

(53) Seybold, P. G. Analysis of the Pk(a)S of Aliphatic Amines Using Quantum Chemical Descriptors. *Int. J. Quantum Chem.* **2008**, *108*, 2849−2855.

(54) Gross, K. C.; Seybold, P. G.; Hadad, C. M. Comparison of Different Atomic Charge Schemes for Predicting Pk(a) Variations in Substituted Anilines and Phenols. *Int. J. Quantum Chem.* **2002**, *90*, 445−458.

(55) Kreye, W. C.; Seybold, P. G. Correlations between Quantum Chemical Indices and the Pk(a)S of a Diverse Set of Organic Phenols. *Int. J. Quantum Chem.* **2009**, *109*, 3679−3684.

(56) Zhang, S. A Reliable and Efficient First Principles-Based Method for Predicting Pka Values. 4. Organic Bases. *J. Comput. Chem.* **2012**, *33*, 2469−2482.

(57) Liu, S.; Pedersen, L. G. Estimation of Molecular Acidity Via Electrostatic Potential at the Nucleus and Valence Natural Atomic Orbitals. *J. Phys. Chem. A* **2009**, *113*, 3648−3655.

(58) 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 C.01*; Gaussian, Inc.: Wallingford, CT, 2010.

(59) Allen, D. M. Relationship between Variable Selection and Data Augmentation and a Method for Prediction. *Technometrics* **1974**, *16*, 125−127.

(60) Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. Roy. Stat. Soc., Ser. B* **1974**, *36*, 111−147.

(61) Snee, R. D. Validation of Regression-Models - Methods and Examples. *Technometrics* **1977**, *19*, 415−428.

(62) Efron, B. Estimating the Error Rate of a Prediction Rule - Improvement on Cross-Validation. *J. Am. Stat. Assoc.* **1983**, *78*, 316−331.

(63) Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and Regression Trees*; Wadsworth: Belmont, CA, 1984.

(64) Breiman, L.; Spector, P. Submodel Selection and Evaluation in Regression - the X-Random Case. *Int. Stat. Rev.* **1992**, *60*, 291−319.

(65) Efron, B.; Tibshirani, R. Improvements on Cross-Validation: The 0.632+ Bootstrap Method. *J. Am. Stat. Assoc.* **1997**, *92*, 548−560.

(66) Myers, R. H. *Classical and Modern Regression with Applications*, 2nd ed.; Duxbury Press (PWS-KENT Publishing Company), 2000.

(67) Harrell, F. E. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, 1st ed.; Springer-Verlag: New York, 2001.

(68) Sheather, S. *A Modern Approach to Regression with R*; Springer: New York, 2008.

(69) Bloomfield, V. A. *Using R for Numerical Analysis in Science and Engineering*; Chapman & Hall/CRC the R series, 2014.

(70) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta.* **1977**, *44*, 129−136.

(71) Saha, S.; Roy, R. K.; Ayers, P. W. Are the Hirshfeld and Mulliken Population Analysis Schemes Consistent with Chemical Intuition? *Int. J. Quantum Chem.* **2009**, *109*, 1790−1806.

(72) De Proft, F.; Van Alsenoy, C.; Peeters, A.; Langenaeker, W.; Geerlings, P. Atomic Charges, Dipole Moments, and Fukui Functions Using the Hirshfeld Partitioning of the Electron Density. *J. Comput. Chem.* **2002**, *23*, 1198−1209.

(73) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. Critical Analysis and Extension of the Hirshfeld Atoms in Molecules. *J. Chem. Phys.* **2007**, *126*, 144111.

(74) Löwdin, P. O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. Chem. Phys.* **1950**, *18*, 365−375.

(75) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735−746.

(76) Contreras, R.; Andres, J.; Safont, V. S.; Campodonico, P.; Santos, J. G. A Theoretical Study on the Relationship between Nucleophilicity and Ionization Potentials in Solution Phase. *J. Phys. Chem. A* **2003**, *107*, 5588−5593.

(77) Jaramillo, P.; Fuentealba, P.; Perez, P. Nucleophilicity Scale for N- and Pi-Nucleophiles. *Chem. Phys. Lett.* **2006**, *427*, 421−425.

(78) Jaramillo, P.; Domingo, L. R.; Chamorro, E.; Perez, P. A Further Exploration of a Nucleophilicity Index Based on the Gas-Phase Ionization Potentials. *J. Mol. Struct.: THEOCHEM* **2008**, *865*, 68−72.

(79) Chamorro, E.; Duque-Norena, M.; Perez, P. A Comparison between Theoretical and Experimental Models of Electrophilicity and Nucleophilicity. *J. Mol. Struct.: THEOCHEM* **2009**, *896*, 73−79.

(80) Domingo, L. R.; Perez, P. The Nucleophilicity N Index in Organic Chemistry. *Org. Biomol. Chem.* **2011**, *9*, 7168−7175.

(81) Chamorro, E.; Duque-Norena, M.; Notario, R.; Perez, P. Intrinsic Relative Scales of Electrophilicity and Nucleophilicity. *J. Phys. Chem. A* **2013**, *117*, 2636−2643.

(82) Szczepanik, D. W.; Mrozek, J. Nucleophilicity Index Based on Atomic Natural Orbitals. *J. Chem.* **2013**, DOI: 10.1155/2013/684134.

(83) Morell, C.; Gazquez, J. L.; Vela, A.; Guegan, F.; Chermette, H. Revisiting Electroaccepting and Electrodonating Powers: Proposals for Local Electrophilicity and Local Nucleophilicity Descriptors. *Phys. Chem. Chem. Phys.* **2014**, *16*, 26832−26842.