PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

ARTICLE TYPE www. rsc.org/xxxxxx | XXXXXXXX

Predicting Bond Strength from a Single Hartree-Fock Ground State Using the Localized Pair Model †

Dylan C. Hennessey, Brendan J. H. Sheppard, Dalton E. C. K. Mackenzie, Jason K. Pearson*

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

We present an application of the recently introduced Localized Pair Model (LPM) [Z. A. Zielinksi and J. K. Pearson, *Comput. Theor. Chem.*, 2013, **1003**, 7990.] to characterize and quantify properties of the chemical bond in a series of substituted benzoic acid molecules. By computing interelectronic distribution functions for doubly-occupied Edmiston-Ruedenberg localized molecular orbitals (LMOs), we show that chemically intuitive electron pairs may be uniquely classified and bond strength may be predicted with remarkable accuracy. Specifically, the HF/*u*6-311G(d,p) level (where *u* denotes a complete uncontraction of the basis set) is used to generate the relevant LMOs and their respective interelectronic distribution functions can be linearly correlated to the well-known Hammett σ_p or σ_m parameters with near-unity correlation coefficients.

1 Introduction

Predicting chemical properties and reactivity is an important challenge in computational quantum chemistry and to do so from first principles generally requires detailed knowledge of potential energy surfaces (PESs) including minima and often first-order saddle points. An accurate and detailed PES yields invaluable knowledge about the reactive pathways accessible to a given molecular structure, and the precise mechanism of how chemical bonds may be broken and formed throughout a reactive process. This information can subsequently be utilized for the purposes of explaining chemical behaviour and the rational design of novel chemical structures with tuned chemical properties. Though one can often employ so-called "chemical intuition" to assess the strength of a network of bonded atoms in a molecular structure, bond strength is more rigorously defined in terms of the bond dissociation energy (BDE). The heterolytic BDE refers to the enthalpy (per mole) required to heterolytically break a given bond of a specific molecular structure.[1]

$$\mathrm{BDE}(A-B) = \Delta_f H^0_{298}(A^-) + \Delta_f H^0_{298}(B^+) - \Delta_f H^0_{298}(A-B) \quad (1)$$

It is an important thermodynamic quantity accessible by computational quantum chemical methods if one can obtain sufficiently accurate thermochemical data on the relevant stationary states. The required accuracy depends on the desired application but generally one is forced to employ sophisticated electronic structure theories[2–7] that can scale prohibitively for large systems. It would therefore be highly advantageous if one could reliably assess the strength of chemical bonds from a single inexpensive ground state electronic structure model, though it is not obvious how one should proceed in doing so. It is tempting, however, to approach the problem from the perspective of electron pair models because since the seminal work of Lewis[8] and later by others,[9] the concept of the electron pair has been a central theme in chemistry that is used to predict chemical structure and reactivity among many other properties of interest. The question is then how to unite intuitive and practical concepts of electron pairs in bonding with rigorous electronic structure theory.

The high dimensionality of the electronic wave function, $\Psi(x_1,\ldots,x_N)$, where $x_i = (r_i, s_i)$ denotes the combined spatial and spin coordinates of electron i, prevents it from being readily interpretable in terms of *practical* chemical concepts. Even within the Born-Oppenheimer approximation,[10] the wave function will explicitly depend on $4N$ coordinates (three spatial and one spin coordinate for each of the N electrons) and quickly becomes intractable for chemical systems of even modest size. The electron density

$$\rho(r) = N \int |\Psi(x_1,\ldots,x_N)|^2 \, ds_1 \, dx_2 \ldots dx_N \quad (2)$$

is a common alternative quantity from which all observable ground state properties are available[11] and is attractive due to its comparatively modest complexity. Interestingly though, neither of these fundamental quantities in quantum chemistry explicitly makes use of the ubiquitous concept of the electron pair, even despite the bielectronic character of the molecular Hamiltonian.

To explicitly consider electron pairs from an *ab initio* perspective en route to predicting accurate ground state properties would require the second order reduced density matrix, or 2-

matrix as it is commonly known, given by [12]

$$\rho_2(r_1,r_2) = \frac{N(N-1)}{2}\int |\Psi(x_1,\ldots,x_N)|^2 ds_1\, ds_2\, dx_3 \ldots dx_N \quad (3)$$

and there exists several emerging techniques with $\rho_2(r_1,r_2)$ as the fundamental variable [13–15]. Perhaps then, employing $\rho_2(r_1,r_2)$ and explicitly considering the electronic structure of a chemical system in terms of its electron pairs may provide a fruitful alternative model for chemical bond properties. [16]

Though one might intuitively harbour ideas regarding the relative distribution of electron pairs within chemical bonds or lone pairs, one is required to first *define* such entities from a quantum mechanical standpoint and then distill relevant interelectronic information from some fundamental quantity, for example $\rho_2(r_1,r_2)$, to properly model such distributions. We have recently introduced the Localized Pair Model (LPM), [17] which accomplishes this by determining interelectronic distribution functions for individual localized molecular orbitals (LMOs). [18–20] LMOs are the so-called "chemically intuitive" orbitals, which capture the notion of localized electron pairs in quantum mechanics by adhering to a unitary transformation of the canonical molecular orbitals (or Kohn-Sham orbitals). The localized pair model seeks to produce qualitative and quantitative data regarding electron pair distributions within the intuitive chemical constructs given by LMOs. One could consider geminals as a natural alternative choice for electron-pair density analyses, however most electronic structure codes utilize one-electron orbitals, making LMOs somewhat more accessible and these are sufficiently general for our purposes. One could also employ Kohn-Sham LMOs to make LPM predictions using density functional theory and this could, in principle, offer an exact treatment of the electronic structure for doubly-occupied LMOs.

This procedure was first proposed by Thakkar and Moore [21] and has since been applied to understanding the nature of bonding in p-block hydrides, saturated main group compounds, fluorinated species, N→B dative structures, and small cyclic molecules. [17] In our original work, we demonstrated that experimental bond dissociation energies (BDEs) among p-block hydrides could be reproduced with very good accuracy by considering any one of several properties of the interelectronic distribution function for the LMO representing that bond type. Notably, this could be accomplished utilizing nothing more than the Hartree-Fock LMOs of a single stationary point representing the minimum energy ground state along the reaction coordinate of the bond dissociation.

In the current work, we extend the application of the LPM to an alternative range of compounds to more generally assess its ability to characterize covalent bonds from the ground state electronic structure. Specifically, we turned our attention to Hammett systems because they represent one of the most widely used scaffolds for studying quantitative aspects of organic reaction mechanisms and there is a wealth of experimental data [22–25] and complementary theoretical work [26–30] available for comparison.

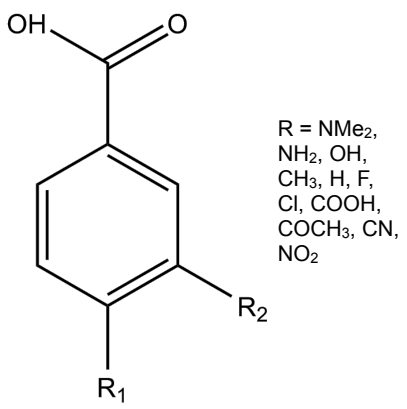

Fig. 1 The series of substituted benzoic acid derivatives studied in the current work. In the text, all molecules are referred to as either p-*R* or m-*R*, where *R* is the substituent at the para (p) or meta (m) position.

We refer to the Hammett parameter, σ_p or σ_m, as a straightforward indicator of bond strength but strictly speaking, they are of course a direct measure of the relative free energies of (in our case) para- or meta-substituted benzoic acid derivatives and their deprotonated analogues as compared to that of the unsubstituted benzoic acid, all in water at 25^{o}C.

$$\log\left(\frac{K_{p/m}}{K_H}\right) = \sigma_{p/m} \quad (4)$$

$K_{p/m}$ is the acid dissociation constant for the appropriate substituted benzoic acid and K_H is the corresponding ionization constant for the unsubstituted benzoic acid. We chose the set of para- and meta-substituted benzoic acids so as to avoid the conflation of steric and electronic effects as would be the case with the ortho-substituted derivatives. In the following sections, we provide details regarding the LPM as well as a comparison between it and other methods of electronic structure analysis over a series of $11 \times 2 = 22$ benzoic acid derivatives (see Figure 1). We discuss the features of the LPM that make it an intuitive and useful interpretive tool in quantum chemistry and show that it can reliably predict $\sigma_{p/m}$. Atomic units are used throughout unless otherwise indicated.

2 Methods

The geometries of a series of substituted benzoic acid derivatives (shown in Figure 1) were optimized at the Hartree-Fock

(HF) level and were confirmed as minima on their respective potential energy surfaces by a harmonic frequency analysis. Subsequently, the canonical molecular orbitals were localized using the Edmiston-Ruedenberg scheme,[18] which we have found is particularly suitable for application to the LPM.[17] The basis sets employed were completely uncontracted Pople 6-311G(d,p) and 6-311G(2d,2p) basis sets (referred to herein as *u*6-311G(d,p) and *u*6-311G(2d,2p)), which have been benchmarked for our purposes.[17] The latter has been used only to assess the level of convergence of our calculated properties with respect to the number of basis functions. All optimizations, frequency calculations, and orbital localizations were performed using either the General Atomic and Molecular Electronic Structure Software (GAMESS) package[31] or Q-Chem.[32] The canonical and localized orbitals were first identified and characterized using the *MacMolPlt*[33] program before they were submitted to our own in-house code for the LPM analysis (*vide infra*).

The topology of the electron density was studied with the AIMALL program[34] using the HF/*u*6-311G** wave functions and the interelectronic probability distributions were plotted and characterized using the *Mathematica* package.[35]

2.1 Localized Pair Model (LPM): Theory

Interelectronic distribution functions, first popularized by Coulson and Neilson,[36] describe the relative distribution of a pair of electrons in space and are often referred to as *intracules*. A *position* intracule, $P(u)$, is the probability distribution for $u = |r_2 - r_1|$, and therefore yields the likelihood of a pair of electrons being separated by a distance u

$$P(u) = \int \rho_2(r_1, r_2)\delta(r_{12} - u)\, dr_1\, dr_2\, d\Omega_u \qquad (5)$$

where δ is the one-dimensional Dirac delta function and $d\Omega_u$ indicates integration over the angular components of the u vector. As such, $P(u)$ is often referred to as the "spherically averaged" intracule density. In the case of a restricted single-determinant wave function, $\rho_2(r_1, r_2)$ may be calculated for two electrons within the same spatial orbital (say, orbital k; ψ_k) by

$$\rho_2^k(r_1, r_2) = |\psi_k(r_1)|^2 |\psi_k(r_2)|^2. \qquad (6)$$

If ψ_k is expanded in a linear combination of one-electron Gaussian basis functions, then substituting (6) into (5) yields

$$P^k(u) = \sum_{\mu\nu\lambda\sigma} \Gamma^k_{\mu\nu\lambda\sigma} (\mu\nu\lambda\sigma)_P \qquad (7)$$

where $\Gamma^k_{\mu\nu\lambda\sigma}$ is the HF two-particle density matrix (which in this case reduces to $C_{\mu k}C_{\nu k}C_{\lambda k}C_{\sigma k}$, where $C_{\mu k}$ are the LCAO coefficients for the kth localized orbital) and $(\mu\nu\lambda\sigma)_P$ are the resultant integrals over the Gaussian primitives indexed by μ, ν, λ and σ. $P^k(u)$ is therefore an intracule describing a single pair of electrons (which are represented by orbital k) and thus the full $\rho_2(r_1, r_2)$ need not be used.

The integrals $(\mu\nu\lambda\sigma)_P$ have been solved analytically[37] with Gaussian primitives of general angular momentum using an Obara-Saika type recurrence relation.[38–41]

Interestingly, equation (6) also allows us to relate the LMO intracule density directly to the electron density by the so-called Coulomb component, defined as[42]

$$J(u) = \frac{1}{2}\int\int \rho(r)\rho(r+u)\, \mathrm{d}r\, \mathrm{d}\Omega_u. \qquad (8)$$

This affords one the opportunity to apply the LPM analysis within the framework of Kohn-Sham density functional theory (which can, in principle, provide exact electronic structures) and this is the subject of ongoing research.

Equation (7) indicates that the generation of $P^k(u)$ formally scales as K^4, where K is the number of basis functions; however, given that we work in a space of LMOs, increasing the size of the molecule does not significantly increase the computational cost since many of the LCAO coefficients will be zero or near zero in regions far from the LMO of interest.

In addition to the qualitative features of $P^k(u)$, we also extract quantitive information from the interelectronic distribution. Generally, $P^k(u)$ for a single LMO is a unimodal distribution with a global maximum near 1 or 2 a.u. As such, it is appropriate to expect that metrics describing where the intracular densities peak and the breadth of their distribution are reasonably universal tools to characterize the distribution of electron pairs. The peak positions (i.e. most probable interelectronic separations) are given by u_{max} but we also consider the gradient of $P^k(u)$ near the origin, $\nabla P^k(0.1)$.

The degree to which the electron pair is likely to deviate from u_{max} (i.e. the breadth of the density) is given by the curvature of the intracule about its maximum, which is explicitly measured by the value of the Laplacian of the intracule at that point

$$\delta_P = \nabla^2 P^k(u_{max}) \qquad (9)$$

A relatively *high* value of δ_P (since we are generally concerned with maxima, δ_P will always be negative and therefore a value close to 0 is considered high) indicates a *high* degree of fluctuation in the interelectronic distance between the electrons. Conversely, a relatively low value indicates a low degree of fluctuation in the interelectronic distance.

Additionally, several moments of the intracules are useful characterization metrics and we employ

$$P_m = \int_0^\infty P^k(u)\, u^m du \qquad (10)$$

where $m = \{-1, 0, 1\}$. P_{-1} is directly equated to the Coulomb repulsion energy of an electron pair[14,42] (E_J), while P_1 corresponds to the *average* interelectronic separation. Finally,

$P_0 = \frac{N(N-1)}{2} = 1$ is the total number of electron pairs described by $P^k(u)$ and is a necessary condition to ensure our densities are properly normalized.

Additionally, the breadth of an intracule may also be measured by u_{95}, which is defined as the limit of integration such that

$$\int_0^{u_{95}} P^k(u)\,du = 0.95\frac{N(N-1)}{2} \tag{11}$$

The value of u_{95} provides a numerical measure of the maximum interelectronic separation one is *likely* to observe for the particular electron pair (i.e. only 5% of the probability distribution exists beyond u_{95}).

Our code uses LMO coefficients $\{C_{\mu k}\}$ to generate $\Gamma_{\mu\nu\lambda\sigma}$ and subsequently contract it with $(\mu\nu\lambda\sigma)_P$ integrals according to (7). The portion of our code that calculates $(\mu\nu\lambda\sigma)_P$ was modified from an earlier version written by Hollett and Gill.[37] The output is $P^k(u)$ calculated over a series of u points on a radial quadrature grid defined by Mura and Knowles.[43] With this grid, a total of i_{max} grid points are chosen such that

$$u_i = -R\log\left(1 - \left[\frac{i}{i_{max}+1}\right]^3\right) \tag{12}$$

where R is a scaling parameter that allows the user to control the span of u space. The resultant data is then interpolated for further analysis using the *Mathematica* package.[35] R and i_{max} must therefore be chosen such that the interpolated $P(u)$ have converged and we find that $R = 3.0$ and $i_{max} = 250$ affords a suitable grid density.

3 Results and Discussion

3.1 An LPM description of benzoic acid

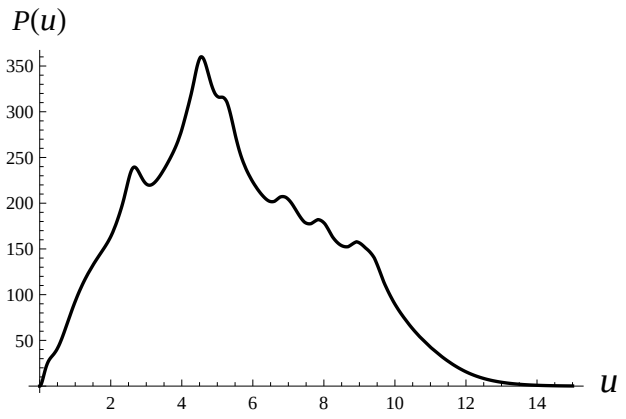

Fig. 2 The total molecular intracule, $P(u)$ for the benzoic acid molecule.

We begin with an example of the total molecular intracule (given by equation 5) for the prototypical benzoic acid system, shown in Figure 2, indicating the probability of observing *any* pair of electrons separated by a distance u. Its rich topology is related to the complexity of the electronic structure[44] of benzoic acid in terms of electron pairs. It is difficult to glean much information about *individual* pairs of electrons from this figure because it represents the probability distribution for all $N(N-1)/2 = 2016$ unique pairs, since benzoic acid has $N = 64$ electrons. The LPM replaces Ψ in equation 3 with a specific LMO, effectively reducing the resultant $P(u)$ to a distribution of a *single* electron pair, localized to the region of the LMO. The LPM is therefore an orbital decomposition of intracule functional theory, introduced by Gill and coworkers.[15] While one is free to choose any orbital representation scheme they like, we find that it is particularly instructive to focus on the localized set of Edmiston and Reudenberg (ER)[17,18] and illustrated in Figure 3 are a subset of valence ER LMOs for benzoic acid.

Borrowing terms from Valence Bond Theory, we may identify a σ LMO as one formed predominantly by head-on overlap of neighbouring atomic orbitals, while a π LMO is generally formed by a bimodal side-on overlap of it's constituent atomic orbitals. Furthermore, there exist non-bonding LMOs (which we label n), formed primarily from atomic orbitals on a single atomic center. Depicted in the red area of Figure 3 are the σ LMOs. The aromatic ring gives rise to π LMOs and a representative example is shown in the blue area of Figure 3. There are three roughly equivalent analogues of such orbitals in all of our phenyl rings, as expected. The carbonyl group orbitals are also shown in the blue area of the figure and are neither strictly characteristic of σ nor π. Though formally a σ *and* a π bond from the perspective of valence bond theory, each of the two carbonyl bonds in our localized set is equivalent (equal but opposite in sign as defined by their respective molecular orbital coefficients) and thus neither are purely σ nor π in nature having significant side-on p orbital contributions as well as head-on overlap of s and p atomic orbitals. The nonbonding orbitals are presented in the green area of the figure and one example of each of the O atom lone pairs is shown.

The intracules for each of these orbitals are plotted in the figure for the purposes of comparison and some associated data is presented in Table 1. All of our LPM distributions exhibit a unimodal peak and are (by definition) normalized to $N(N-1)/2 = 1$ for a single electron pair. Generally, $P^k(u)$ peaks near 1 atomic unit, with the obvious exception of π_{CC}, whose relatively delocalized distribution results in an average electronic separation that is significantly larger (e.g. $u_{max} = 2.901$). The π_{CC} intracule is a significant outlier in all of the listed metrics in Table 1. This case exhibits about half of the repulsion energy between electrons and a much higher δ_P, in-

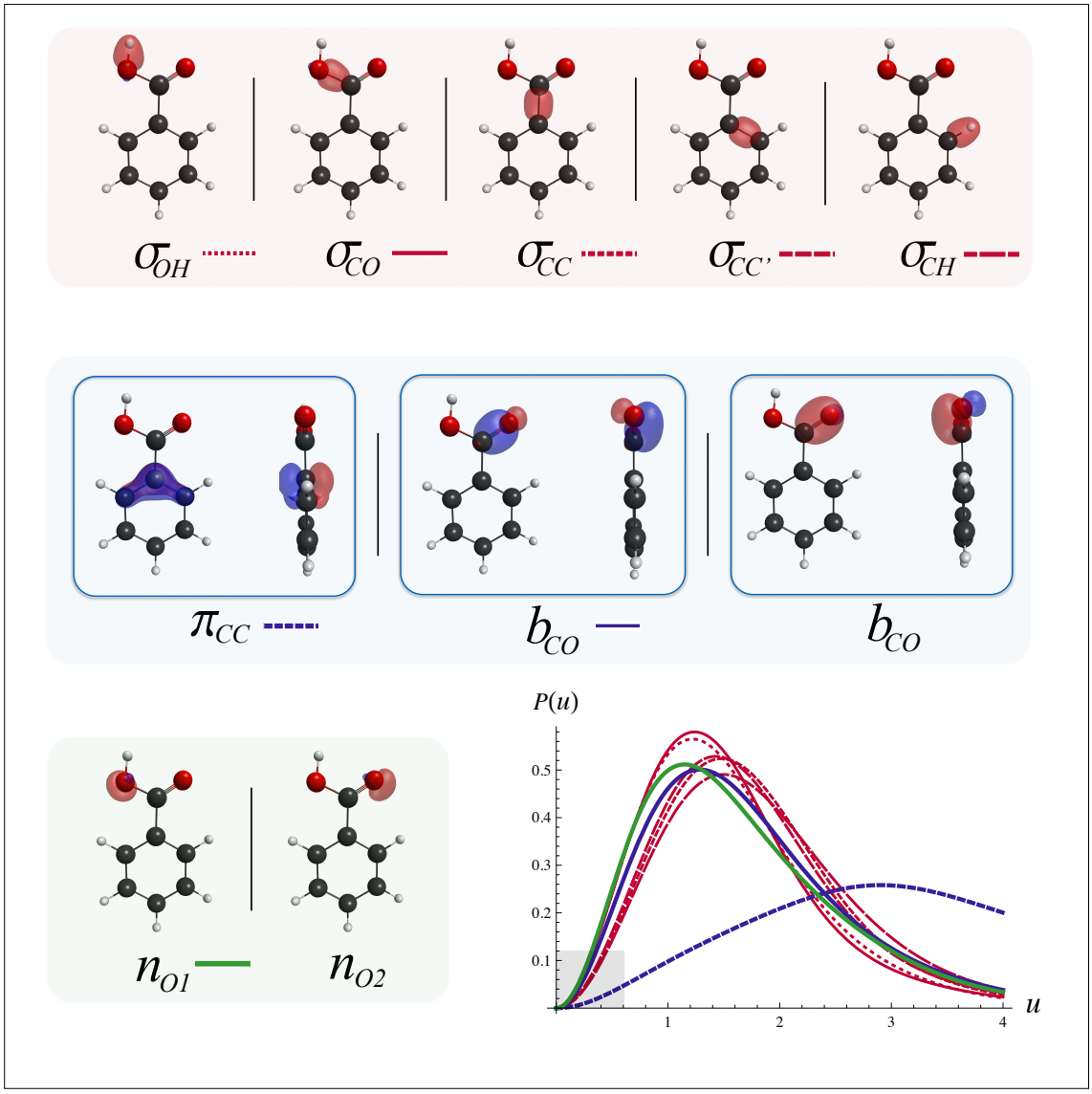

Fig. 3 The set of ER localized orbitals and their associated intracules, $P^k(u)$.

Table 1 Calculated properties of valence ER LMOs in benzoic acid (see text for a description of each parameter). Values in parenthesis are the absolute differences between each parameter obtained with the *u*6-311G(d,p) and *u*6-311G(2d,2p) basis sets. All values are in atomic units unless otherwise stated.

LMO	u_{max}	$\nabla P^k(0.1)$	P_{-1}	P_1	δ_P	u_{95}	r (Å)
σ_{OH}	1.2198 (0.0009)	0.2464 (0.0010)	0.8583 (0.0007)	1.5982 (0.0006)	-1.2365 (0.0023)	3.16	0.95
σ_{CO}	1.2373 (0.0016)	0.2430 (0.0005)	0.8609 (0.0002)	1.5900 (0.0013)	-1.2987 (0.0010)	3.20	1.33
σ_{CC}	1.4775 (0.0013)	0.1382 (0.0005)	0.7366 (0.0004)	1.8060 (0.0016)	-0.9322 (0.0026)	3.43	1.49
$\sigma_{CC'}$	1.4266 (0.0028)	0.1485 (0.0013)	0.7504 (0.0013)	1.7902 (0.0024)	-1.0069 (0.0008)	3.49	1.39
σ_{CH}	1.4955 (0.0030)	0.1308 (0.0010)	0.7137 (0.0009)	1.8763 (0.0010)	-0.7913 (0.0043)	3.59	1.07
π_{CC}	2.9010 (0.0029)	0.0321 (0.0000)	0.3954 (0.0002)	3.4324 (0.0002)	-0.1394 (0.0004)	6.34	1.39
b_{CO}	1.2839 (0.0016)	0.2129 (0.0005)	0.7936 (0.0015)	1.7766 (0.0057)	-0.9014 (0.0099)	3.62	1.18
n_{O1}	1.1481 (0.0003)	0.2714 (0.0008)	0.8440 (0.0018)	1.7075 (0.0042)	-1.0563 (0.0105)	3.54	n/a
n_{O2}	1.1386 (0.0001)	0.2795 (0.0003)	0.8633 (0.0003)	1.6491 (0.0004)	-1.1487 (0.0015)	3.36	n/a

dicating a wide range of probable interelectronic distances.

With respect to the σ intracules, there appear to be three qualitatively distinct types. The σ_{OH} and σ_{CO} intracules show similar characteristics, peaking at small u and having the narrowest distributions in the set ($u_{95} = 3.2$). This is due to the similar electronegativities of C and H, and consequently the similar *differences* in electronegativity between the bonded atoms in each of the σ_{OH} and σ_{CO} LMOs.[45] This causes the pair of bonding electrons to be contracted toward the more electronegative O atom, causing a narrow distribution of interelectonic distances with small u_{max}. The second type of σ intracule observed is exhibited by σ_{CC} and $\sigma_{CC'}$, whose intracules are also nearly identical. Apparently, the presence of additional π electrons has a negligible effect on the electronic distribution within $\sigma_{CC'}$ and the small difference that is observed between it and σ_{CC} can be attributed to the small difference in their respective bond lengths. The $\sigma_{CC'}$ bond, being marginally shorter, exhibits a smaller u_{max} and consequently a steeper initial gradient, more Coulomb repulsion, and a lower δ_P than σ_{CC}, though each of these differences is minor. Finally, the σ_{CH} intracule exhibits one of the broadest densities of the entire set, despite having one of the shortest bonds. This is a result of the relatively equal electronegativities of C and H coupled with the fact that σ_{CH} is a *terminal* bond, allowing the electron pair to be more evenly distributed throughout the space of the LMO.

Similarly broad intracule densities are observed for b_{CO} and n_O, although these are for qualitatively different reasons. In the case of b_{CO}, the interorbital repulsion between the two equivalent bonding LMOs forces the spatial distribution out perpendicularly from the bond axis. This has the effect of increasing the available space for each electron pair because the orbital is effectively less "bounded" by the adjoining atoms as in the case of σ LMOs. Likewise, each of the n_O LMOs is also "unbounded" from the perspective that the electrons they describe are not confined to an interatomic region and this results in broad intracular densities. Though broad, each of these intracules exhibits a somewhat contradictory early maximum and steep initial gradient in $P^k(u)$, characteristic of all cases regarding LMOs containing a significant amount of O character.

Comparison between the u6-311G(d,p) and u6-311G(2d,2p) data indicate that our LPM parameters are generally converged to 10^{-3} but are often converged beyond that level.

3.2 Using the LPM to predict Hammett parameters

If one is to make use of the LPM in quantifying chemical bond strength for these systems (Figure 1), then it becomes necessary to know how the aforementioned properties should change with respect to chemical substitution in the para position. The relative thermodynamic stability of the ionized species is directly related to the electron withdrawing or donating character of the substituent on the aromatic ring. As the electron density migrates toward or away from the carboxyl group, we should expect to observe a concomitant change in electron-electron interactions throughout the system.

Though the site of substitution is well removed from the dissociating O-H bond and localization procedures are largely meant to isolate local electronic structure to retain transferability, some measurable changes must be sustained in σ_{OH} (and/or neighbouring LMOs) for the LPM to be of use. In general, we do indeed observe a small but significant change in the properties of the relevant LMOs upon para or meta substitution and note that the magnitude of this change increases in the order $\sigma_{OH} < \sigma_{CO} < \sigma_{CC}$. For example, in the para series of molecules, the average absolute change in u_{max} for σ_{CC}, σ_{CO}, and σ_{OH} are 2.24×10^{-3}, 7.29×10^{-4}, and 4.04×10^{-4}, respectively over the entire set. Likewise, the average absolute change in P_{-1} is 8.57×10^{-4}, 6.27×10^{-4}, and 3.58×10^{-4}, for the same respective LMOs. We chose to investigate σ_{CO} and σ_{CC} in addition to σ_{OH} because we were interested in quantifying the extent to which electronic structure changes would be observable from distant regions of a molecule.

The differences between the various bond intracules are illustrated in Figure 4, which shows plots of the intracules for each of the σ_{OH} LMOs for every molecule in our para-substituted test set (the analogous data for the meta-substituted systems shows a qualitatively similar trend). The amount by which these densities change after a para substitution is small but not negligible and the magnitude of the change is resolved within the inset plot in the figure.

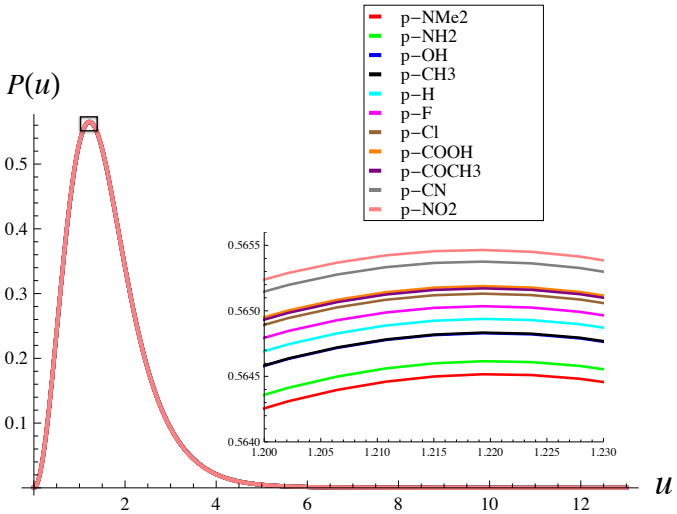

Fig. 4 $P^k(u)$ of the σ_{OH} LMOs for each of the para-substituted molecules with an inset to highlight their differences near the maximum.

Table 2 lists selected LPM metrics for each of the σ_{CC}, σ_{CO}, and σ_{OH} LMOs. Upon inspection it can be seen that the data approximately varies linearly throughout. We have therefore included linear correlation coefficients corresponding to the empirical relationship between each of the LPM metrics and the appropriate Hammett σ parameter.

Such a relationship may be rationalized by considering the definition of the Hammett σ parameter[46] given in equation (4). The more stable the dissociated acidic O-H bond is upon substitution at the para or meta position, the higher its associated $K_{p/m}$ will be relative to K_H and the greater the value of $\sigma_{p/m}$ will be. One usually generalizes this effect according to the electron *withdrawing* or electron *donating* ability of the substituent. Electron withdrawing substituents, such as the p-NO_2 case will withdraw electron density from the carboxylic acid moiety through the aromatic ring and the net result within the terminal O-H bond would be that the σ_{OH} pair of electrons would be contracted toward the oxygen atom. This would *decrease* u_{max} and consequently *increase* P_{-1}. The opposite would be true for electron donating substituents and this is indeed observed in Table 2. The trend is less obvious for σ_{CO} and σ_{CC} but still present.

One can see that these data correlate very well, having no R^2 correlation coefficient less than 0.92 (with the obvious exception of σ_{CC} data, which isn't correlated at all for the meta derivatives) and in one case achieving $R^2 = 0.9773$. We find that Hammett parameters may be predicted on average to within 0.05 considering only the average interelectronic separation within the terminal OH bond. The supposition that electron pair properties are related to bond strengths is apparently well justified. The various LPM metrics for σ_{CC} and σ_{CO} show a different relationship to $\sigma_{p/m}$, though both are also strongly correlated as shown in Table 2 and Figure 5. In fact, the best quantum mechanical predictor of Hammett $\sigma_{p/m}$ parameters for benzoic acid derivatives that we are aware of is P_1 for the σ_{OH} bond.

As an additional comparator, we have also collected data for the one-electron density at the bond critical point (BCP) between the bonded atoms in σ_{OH}. This value of the density, ρ_{BCP}, is a popular Quantum Theory of Atoms in Molecules (QTAIM)[47,48] bond descriptor and has been applied to our systems previously.[26,30] Figure 6 illustrates an example of our results for ρ_{BCP} for the σ_{OH}, σ_{CO}, and σ_{CC} bonds in each of our substituted benzoic acids. We find that ρ_{BCP} for the σ_{OH} and σ_{CO} bonds are also well correlated to the Hammett σ parameters, in good agreement with earlier work by Mandado et al.,[26] Kégl[30] et al., and Popelier.[28]

Interestingly, the ρ_{BCP} data varies *with* u_{max} in the case of the σ_{OH} bond but in the *opposite* direction of u_{max} for the σ_{CO} and σ_{CC} bonds. This can be explained by considering the relative location of the BCPs between the atomic maxima in $\rho(r)$. In the case of σ_{OH}, the BCP is near the H atom and the vast

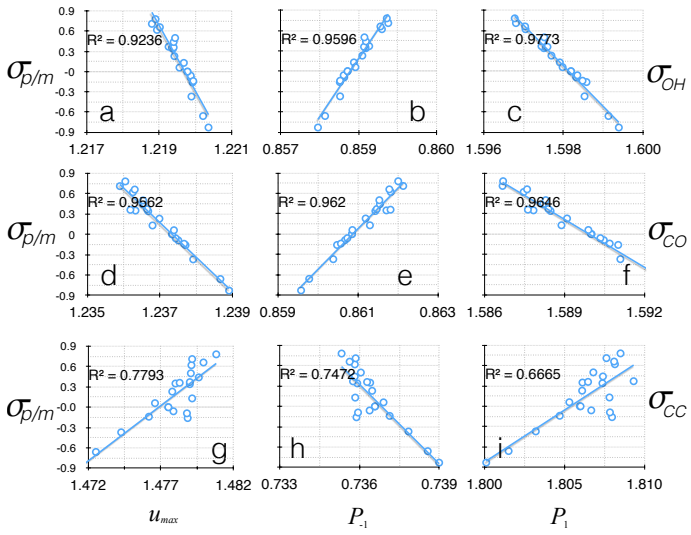

Fig. 5 Linear correlations between Hammett $\sigma_{p/m}$ and u_{max}, P_{-1}, and P_1 for a-c) σ_{OH}, d-f) σ_{CO}, and g-i) σ_{CC}.

majority of the electron density is associated with the O atomic basin according to QTAIM. So an additional shift in electron density toward the O atom would only further rob the H atom and the value of ρ_{BCP} would decrease, consequently decreasing u_{max} because the electron pair is effectively in a smaller space (i.e. near the O atom). The opposite is true for σ_{CO} and σ_{CC} because these are bonds between two heavy atoms whose BCPs are closer to the centre of the internuclear axis and having larger maxima in ρ at each atomic coordinate. As such, there is significant electron density on "either side" of the BCP and this results in a relatively high probability of observing u near the internuclear distance. An increase in ρ_{BCP} would signify a greater probability of finding electrons in the internuclear region and therefore a greater probability of observing smaller u_{max}. A smaller ρ_{BCP} would indicate a lower probability of observing electrons in the internuclear region and thus the electron density would remain relatively isolated on each nucleus and u_{max} would increase. Such intuitive observations stem from the connection between the LPM and the electron density (*vide supra*), therefore the most probable interelectronic distances are directly related to the distances between regions of high electron density.

The connection of the LPM to *experimental* bond properties is a particularly promising result considering that HF is incapable of accurately capturing the free energy change associated with a heterolytic cleavage of the O-H bond. The LPM essentially bypasses the shortcomings of HF by deducing the relationship between the ground state electronic distribution and the BDE. Though our test set of molecules span

Table 2 The u_{max}, P_{-1}, and P_1 values for σ_{CC}, σ_{CO} and σ_{OH} of the substituted benzoic acid derivatives. Also listed are the linear correlation coefficients.

		σ_{CC}			σ_{CO}			σ_{OH}		
Molecule	$\sigma_{p/m}$	u_{max}	P_{-1}	P_1	u_{max}	P_{-1}	P_1	u_{max}	P_{-1}	P_1
p-NMe_2	-0.83	1.4717	0.7390	1.8001	1.2389	0.8596	1.5931	1.2204	0.8577	1.5994
p-NH_2	-0.66	1.4725	0.7386	1.8015	1.2387	0.8598	1.5926	1.2202	0.8579	1.5991
p-OH	-0.37	1.4743	0.7378	1.8032	1.2379	0.8604	1.5911	1.2199	0.8581	1.5985
p-CH_3	-0.14	1.4762	0.7371	1.8047	1.2377	0.8606	1.5907	1.2199	0.8581	1.5985
p-H	0	1.4775	0.7366	1.8060	1.2373	0.8609	1.5900	1.2198	0.8583	1.5982
p-F	0.06	1.4766	0.7369	1.8053	1.2374	0.8609	1.5899	1.2196	0.8584	1.5980
p-Cl	0.23	1.4778	0.7365	1.8061	1.2370	0.8612	1.5891	1.2194	0.8586	1.5977
p-COOH	0.44	1.4796	0.7358	1.8074	1.2365	0.8616	1.5882	1.2194	0.8586	1.5975
p-$COCH_3$	0.50	1.4791	0.7361	1.8068	1.2365	0.8615	1.5883	1.2194	0.8586	1.5975
p-CN	0.66	1.4799	0.7356	1.8081	1.2363	0.8618	1.5876	1.2190	0.8589	1.5970
p-NO_2	0.78	1.4808	0.7353	1.8085	1.2361	0.8620	1.5867	1.2189	0.8590	1.5968
m-NMe_2	-0.16	1.4789	0.7359	1.8080	1.2377	0.8605	1.5910	1.2199	0.8581	1.5986
m-NH_2	-0.09	1.4788	0.7359	1.8078	1.2375	0.8607	1.5905	1.2199	0.8582	1.5983
m-OH	0.13	1.4792	0.7358	1.8078	1.2368	0.8613	1.5889	1.2197	0.8584	1.5979
m-CH_3	-0.06	1.4779	0.7364	1.8067	1.2374	0.8607	1.5904	1.2199	0.8582	1.5984
m-H	0	1.4775	0.7366	1.8060	1.2373	0.8609	1.5900	1.2198	0.8583	1.5982
m-F	0.34	1.4790	0.7359	1.8074	1.2367	0.8614	1.5885	1.2194	0.8586	1.5975
m-Cl	0.37	1.4790	0.7358	1.8093	1.2367	0.8615	1.5884	1.2193	0.8587	1.5975
m-COOH	0.35	1.4780	0.7364	1.8061	1.2363	0.8617	1.5878	1.2194	0.8586	1.5976
m-$COCH_3$	0.36	1.4783	0.7363	1.8065	1.2362	0.8618	1.5876	1.2194	0.8586	1.5976
m-CN	0.62	1.4791	0.7358	1.8082	1.2363	0.8618	1.5875	1.2189	0.8589	1.5971
m-NO_2	0.71	1.4792	0.7359	1.8076	1.2359	0.8621	1.5867	1.2188	0.8591	1.5968
R^2 (para only)		0.9791	0.9780	0.9659	0.9936	0.9955	0.9920	0.9439	0.9691	0.9835
R^2 (meta only)		0.1679	0.0979	0.0326	0.9112	0.9299	0.9421	0.9820	0.9915	0.9920
R^2		0.7793	0.7472	0.6665	0.9562	0.9620	0.9646	0.9236	0.9596	0.9773

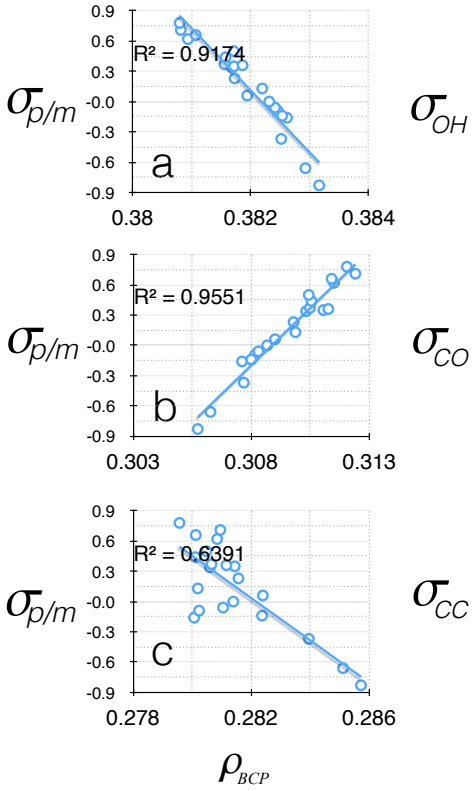

Fig. 6 Linear correlations between Hammett $\sigma_{p/m}$ and ρ_{BCP} for a) σ_{OH}, b) σ_{CO}, and c) σ_{CC}.

a large range in the chemical space characterized by σ_p and σ_m, they represent a very narrow band of quantum mechanical free energy changes for the ionization of benzoic acid and its derivatives, and therein lies the significant challenge of calculating such bond properties *ab initio* using methods such as HF theory. Though unable to adequately describe the physics of the bond-breaking process itself (ironically, due to the bielectronic character of the Hamiltonian), HF theory is still able to retain important differences in the ground state electronic structure of each species in the current study. By exploiting the intuitive relationships between ground state electronic structure and bond strength descriptors (in this case, σ_p and σ_m) one is able to fortuitously bypass these shortcomings of a so-called "non-correlated" model.

4 Conclusions

In the current work, we explored the Localized Pair Model (LPM) as an interpretive tool in the analysis of the electronic structure of a series of para- and meta-substituted benzoic acid derivatives. Because the LPM predicts distributions of localized electron pairs representing chemically intuitive features of electronic structure, it is a natural analytic tool and potentially very useful for a variety of applications, namely the characterization of chemical bond properties. Specifically we sought to predict chemical bond strengths for the terminal O-H group, which is experimentally captured using the well-known Hammett parameter (σ_p or σ_m), by considering only a single ground state Hartree-Fock reference structure. Though HF suffers from an incorrect treatment of electron repulsion and is therefore incapable of accurately describing the bond-breaking process, we show that HF theory is still able to retain important differences in the ground state electronic structure of molecules that is sufficient to predict bond strength for the chosen set of chemical systems. Of course, application of the LPM is not restricted to HF electronic structures but it is an illustrative example of the utility of intracular analysis.

By measuring properties of the σ_{OH} orbital such as the average interelectronic separation or the Coulomb repulsion energy, the bond strength (characterized by the Hammett parameter σ_p or σ_m) could be empirically predicted to within 0.05 on average. This is impressive considering the fact that an explicit evaluation of bond strength using the underlying HF/*u*6-311G** wave function is in error by almost three times that much. Additionally, the σ_{CO} LMOs proved useful in the prediction of bond strength but the σ_{CC} intracules did not due to . We have also compared our analysis to that using the Quantum Theory of Atoms in Molecules and note that their respective performance is poorer, with ρ_{BCP} of the σ_{CO} bond predicting $\sigma_{p/m}$ on average to within 0.08.

The prediction of bond strengths or kinetic labilities[49] *a priori*, without complete knowledge of the electronic structure of the relevant critical points along a particular reaction coordinate is of obvious utility. In the current work we have shown that one can bypass the shortcomings of HF theory by making use of characteristic ground state properties to achieve relatively high accuracy via the LPM, which is an analogous strategy to that employed within the broader intracule functional models.[15]

5 Acknowledgements

The authors gratefully acknowledge the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Canada Foundation for Innovation (CFI) for funding that supported this work. Computational facilities are provided by the Atlantic Computational Excellence Network (ACEnet), the regional high-performance computing consortium for universities in Atlantic Canada. ACEnet is funded by CFI, the Atlantic Canada Opportunities Agency (ACOA), and the provinces of

NewFoundland & Labrador, Nova Scotia, and New Brunswick.

References

1 Y.-R. Luo, *Handbook of Bond Dissociation Energies in Organic Compounds*, CRC, New York, 2003.
2 C. E. Check and T. Gilbert, *J. Org. Chem.*, 2005, **70**, 9828–9834.
3 K. C. Hunter and A. East, *J. Phys. Chem. A*, 2002, **106**, 1346–1356.
4 S. Gronert, *J. Org. Chem.*, 2006, **71**, 1209–1219.
5 J. Martell, R. Boyd and Z. Shi, *J. Phys. Chem. A*, 1993, **97**, 7208–7215.
6 L. Curtiss, K. Raghavachari, G. Trucks and J. Pople, *J. Chem Phys.*, 1991, **94**, 7221–7230.
7 L. Curtiss, K. Raghavachari, P. Redfern, V. A. Rassolov and J. Pople, *J. Chem Phys.*, 1998, **109**, 7764–7776.
8 G. N. Lewis, *J. Am. Chem. Soc.*, 1916, **38**, 762–785.
9 R. Gillespie and E. Robinson, *Angew. Chem. Int. Ed.*, 1996, **35**, 495–514.
10 M. Born and J. Oppenheimer, *Ann. Physik*, 1927, **389**, 457–484.
11 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
12 A. Coleman and V. Yukaloz, *Reduced Density Matrices: Coulson's Challenge*, Springer, 2000.
13 N. Shenvi, H. van Aggelen, Y. Yang, W. Yang, C. Schwerdtfeger and D. Mazziotti, *J. Chem Phys.*, 2013, **139**, 054110.
14 J. K. Pearson, D. L. Crittenden and P. M. W. Gill, *J. Chem. Phys.*, 2009, **130**, 164110.
15 P. M. W. Gill, *Annu. Rep. Prog. Chem., Sect. C*, 2011, **107**, 229–241.
16 M. Piris, X. Lopez and J. Ugalde, *J. Chem. Phys.*, 2008, **128**, 214105.
17 Z. A. Zielinksi and J. K. Pearson, *Comput. Theor. Chem.*, 2013, **1003**, 79–90.
18 C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.*, 1963, **35**, 457–464.
19 J. Pipek and P. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916–4926.
20 S. Boys, *Rev. Mod. Phys.*, 1960, **32**, 296–299.
21 A. J. Thakkar and N. Moore, *Int. J. Quantum Chem.*, 1981, **20**, 393–400.
22 F. Bordwell and P. Boutan, *J. Am. Chem. Soc.*, 1956, **78**, 854–860.
23 D. H. McDaniel and H. Brown, *J. Org. Chem.*, 1958, **23**, 420–427.
24 C. Hansch and A. Leo, *Substituents Constants for Correlation Analysis in Chemistry and Biology*, Wiley-Interscience, New York, 1979.
25 C. Hansch, A. Leo and R. Taft, *Chem. Rev.*, 1991, **91**, 165–195.
26 M. Mandado, R. Mosquera and A. Graña, *Chem. Phys. Lett.*, 2004, **386**, 454–459.
27 I. Fernández and G. Frenking, *J. Org. Chem.*, 2006, **71**, 2251–2256.
28 P. Popelier, *J. Phys. Chem. A*, 1999, **103**, 2883–2890.
29 W. Oziminski and J. C. Dobrowolski, *J. Phys. Org. Chem.*, 2009, **22**, 769–778.
30 T. Papp, L. Kollar and T. Kegl, *Chem. Phys. Lett.*, 2013, **588**, 51–56.
31 M. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordon, J. Jensen, S. Koseki, N. Matsunaga, K. Nguyen, S. Su, T. Windus, M. Dupuis and J. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347–1363.
32 Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. V. Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C. P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Z. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. W. III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. S. III, J. Kong, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3172–3191.
33 B. M. Bode and M. S. Gordon, *J. Mol. Graphics Modell.*, 1998, **16**, 133–138.
34 T. A. Keith, *AIMALL (Version 13.11.04)*, TK Gristmill Software, Overland Park KS, USA, 2013.
35 Wolfram Research, Inc., *Mathematica 7*, 2008.
36 C. A. Coulson and A. H. Neilson, *Proc. Phys. Soc. (London)*, 1961, **78**, 831–837.
37 J. W. Hollett and P. M. W. Gill, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2972–2978.
38 H. Schlegel, *J. Chem. Phys.*, 1982, **77**, 3676–3681.
39 S. Obara and A. Saika, *J. Chem. Phys.*, 1986, **84**, 3963–3974.
40 S. Obara and A. Saika, *J. Chem. Phys.*, 1988, **89**, 1540–1559.
41 R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3072–3077.
42 A. Proud, M. Walker and J. Pearson, *Int. J. Quant. Chem.*, 2013, **113**, 76–82.
43 M. Mura and P. Knowles, *J. Chem. Phys.*, 1996, **104**, 9848–9858.
44 P. M. W. Gill, A. M. Lee, N. Nair and R. D. Adamson, *J. Mol. Struct. (Theochem)*, 2000, **506**, 303–312.
45 W. B. Jensen, *J. Chem. Ed.*, 1996, **73**, 11–20.
46 I. Hammett, *Physical Organic Chemistry*, McGraw-Hill, New York, 2nd edn, 1970.
47 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, USA, 1994.
48 *The Quantum Theory of Atoms in Molecules*, ed. C. Matta and R. Boyd, Wiley-VCH, Weinheim, 2007.
49 G. Markopoulos and J. Grunenberg, *Angew. Chem. Int. Ed.*, 2013, **52**, 10648–10651.

TOC Text Entry:

We present electron pair distributions within chemical bonds and show that these characterize and quantify chemical bond strength.