

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Density Functional Reactivity Theory Study of $S_N2$ Reactions from the Information-Theoretic Perspective

Zemin Wu,<sup>a</sup> Chunying Rong,<sup>a,b\*</sup> Tian Lu,<sup>c</sup> Paul W. Ayers,<sup>b</sup> and Shubin Liu<sup>a,d\*</sup>

<sup>a</sup> *Key Laboratory of Chemical Biology and Traditional Chinese Medicine Research (Ministry of Education of China), College of Chemistry and Chemical Engineering, Hunan Normal University, Changsha Hunan 410081, China*

<sup>b</sup> *Department of Chemistry, McMaster University, Hamilton, Ontario, Canada L8S 4M1*

<sup>c</sup> *Beijing Kein Research Center for Natural Sciences, Beijing 100022, P. R. China*

<sup>d</sup> *Research Computing Center, University of North Carolina, Chapel Hill, North Carolina 27599-3420, USA*

## Abstract

As a continuation of our recent efforts to quantify chemical reactivity with quantities from the information-theoretic approach within the framework of density functional reactivity theory, the effectiveness of applying these quantities to quantify electrophilicity for the bimolecular nucleophilic substitution ( $S_N2$ ) reactions in both gas phase and aqueous solvent is presented in this work. We examined a total of 21 self-exchange  $S_N2$  reactions for the compound with the general chemical formula of  $R_1R_2R_3C-F$ , where  $R_1$ ,  $R_2$ , and  $R_3$  represent substituting alkyl groups such as  $-H$ ,  $-CH_3$ ,  $-C_2H_5$ ,  $-C_3H_7$ , and  $-C_4H_9$  in both gas and solvent phases. Our findings confirm that scaling properties for information-theoretic quantities found elsewhere are still valid. It has also been verified that the barrier height has the strongest correlation with the electrostatic interaction, but the contributions from the exchange-correlation and steric effects, though less significant, are indispensable. We additionally unveiled that the barrier height of these  $S_N2$  reactions can reliably be predicted not only by the Hirshfeld charge and information gain at the regioselective carbon atom, as have been previously reported by us for other systems, but also by other information-theoretic descriptors such as Shannon entropy, Fisher information, and Ghosh-Berkowitz-Parr entropy on the same atom. These new findings provide further insights for the better understanding of the factors impacting the chemical reactivity of this vastly important category of chemical transformations.

---

Correspondence authors. E-mail: [chunyingrong@gmail.com](mailto:chunyingrong@gmail.com) (C.Y.R.); [shubin@email.unc.edu](mailto:shubin@email.unc.edu) (S.B.L.)

## 1. Introduction

In principle, the best way to evaluate chemical reactivity for a chemical reaction is through examining its associated transition state structure to obtain the activation energy (or barrier height), intrinsic reaction pathway, rate constant, etc. These numerical results nevertheless do not provide with us any conceptual framework for the insightful understanding about the origin, tendency, pattern, and factors governing the chemical reactivity of the molecular system. According to density functional theory,<sup>1</sup> any property of a molecule in the ground state including chemical reactivity is dictated by its electron density. The endeavor to make use of this density and its associated quantities such as density gradient and Laplacian as reactivity descriptors to establish a conceptual framework to quantitatively predict chemical reactivity is called density functional reactivity theory (DFRT).<sup>1-4</sup> It has been our ongoing efforts in the past few years to develop such a reliable framework under the umbrella of DFRT, which is able to accurately predict reactivity properties, such as acidity,<sup>5-9</sup> basicity,<sup>5,10</sup> nucleophilicity,<sup>11-15</sup> electrophilicity,<sup>11,12,16</sup> regioselectivity,<sup>11-16</sup> among others.<sup>17-20</sup>

Information-theoretic quantities such as Shannon entropy,<sup>21</sup> Fisher information,<sup>22</sup> Ghosh-Berkowitz-Parr entropy<sup>23</sup> and information gain<sup>24,25</sup> are all functions of the electron density and its associated quantities (e.g., density gradient and Laplacian, etc.), so they should be natural choices as reactivity descriptors in DFRT. Earlier, we examined their scaling behaviors with respect to the number of electron populations under different partition schemes of atoms in molecules.<sup>26</sup> Also, based on the *information conservation principle* recently introduced by us,<sup>11</sup> we proposed to use the information gain and Hirshfeld charge as a quantitative measure of both nucleophilicity and electrophilicity. Its effectiveness has been verified in a series of carbon<sup>11</sup> and nitrogen<sup>12</sup> containing systems. This idea has also been successfully applied to investigate the reactivity and regioselectivity for electrophilic aromatic substitution reactions. A novel explanation was provided to understand the nature and origin of the so-called *ortho/para* and *meta* group directing phenomena for this important category of reactions.<sup>13-15</sup>

In this work, we apply the idea to study another extremely important category of chemical reactions, bimolecular nucleophilic substitution ( $S_N2$ ).<sup>27,28</sup> The main difference between this work and our previous studies<sup>13-15</sup> is that our previous focus was on quantifying nucleophilicity with the information gain or Hirshfeld charge, but in the present study we instead work on quantifying electrophilicity,<sup>16</sup> the capability of an electrophile to accept as many electrons as possible, for  $S_N2$  reactions. In a typical  $S_N2$  reaction, when an incoming nucleophile such as halides tackles the central atom of a substrate, which often is an electrophilic atom such as an aliphatic  $sp^3$  carbon center with an electronegative leaving group attached, one bond will be broken and another bond be formed synchronously with the inversion of the tetrahedral geometry at the central atom.<sup>27,28</sup> Kinetically, the rate of an  $S_N2$  reaction is second order, as the rate-determining step depends on both the nucleophile concentration and the concentration of the substrate. The factors impacting the reactivity of  $S_N2$  reactions include the incoming nucleophile, the size of the groups attached to the central atom, solvents, the nature of the leaving group, etc.<sup>28</sup> In an earlier study, we examined this kind of reactions and systematically looked into the question of which energy component dominates the barrier height of these reactions.<sup>29</sup> Here, we revisit these systems, trying to address the following issues from the information-theoretic perspective of DFRT. At first, we examine the behavior of Hirshfeld charge, information gain, and other DFRT quantities as descriptors to predict reaction barrier heights, and reveal strong linear correlations between the barrier height and these quantities. Secondly, with the tools we recently developed in quantifying chemical effects such as steric and electrostatic effects,<sup>30</sup> we decompose the barrier height and determine the energy components that play dominant and indispensable roles. Finally, we disclose scaling properties of information-theoretic quantities in these systems with respect to the electronic population at the entire molecular and atoms-in-molecules levels.<sup>26</sup> Put together, these novel results from this work should shed fresh light on further

understanding the factors governing the reactivity for this vastly important category of chemical transformations.

## 2. Theoretical and Methodological Considerations

Earlier, we investigated the origin of barrier heights for  $S_N2$  reactions from the perspective of the following two energy partition schemes.<sup>29</sup> One is the conventional approach in density functional theory (DFT), where the total energy difference  $\Delta E$  between the transition state and reactant complex of an  $S_N2$  reaction system is partitioned as follows,<sup>1,29</sup>

$$\Delta E[\rho] = \Delta T_s[\rho] + \Delta E_e[\rho] + \Delta E_{xc}[\rho] \quad (1)$$

where  $T_s$  stands for the non-interacting kinetic energy,  $E_e$  denotes the electrostatic energy, and  $E_{xc}$  represents the exchange-correlation energy component with  $\rho(\mathbf{r})$  as the ground state electron density of an  $N$ -electron system, which satisfies the normalization condition,  $\int \rho(\mathbf{r}) d\mathbf{r} = N$ . The electrostatic term consists of the nuclear-electron attraction,  $V_{ne}$ , classical electron-electron Coulombic repulsion,  $J$ , and nuclear-nuclear repulsion,  $V_{nn}$ ,

$$\Delta E_e[\rho] = \Delta V_{ne}[\rho] + \Delta J[\rho] + \Delta V_{nn}. \quad (2)$$

Alternatively, a new partition scheme was recently proposed by one of us,<sup>30</sup>

$$\Delta E[\rho] = \Delta E_s[\rho] + \Delta E_e[\rho] + \Delta E_q[\rho] \quad (3)$$

where the total energy difference comes from the contribution of three distinct but independent physiochemical effects, steric  $\Delta E_s$ , electrostatic  $\Delta E_e$ , and fermionic quantum due to exchange and correlation effects  $\Delta E_q$ . The common term in these two schemes is the electrostatic contribution  $\Delta E_e$ . The benefit of applying these formulas to examine energy differences is that contributions from energy components are directly related to the barrier height, so one will have a clear picture of which components contribute positively or negatively and dominantly or negligibly to  $\Delta E$ , and whether or not there is any strong correlation of the barrier height to any of these energy components. Studies for a

number of systems have shown that it is the electrostatic term  $\Delta E_e$  that plays the dominant role in governing the total energy difference, but contributions from other terms, though less significant, are also indispensable.<sup>31</sup>

The second energy partition scheme, Eq. (3), was based on a solid physiochemical foundation, as has been previously showed.<sup>30</sup> The physical nature of this new quantification of the steric effect is based on the introduction of a new reference state, where electrons in atoms and molecules are assumed to behave like bosons. If the density of the hypothetical boson state is the same as that of the fermionic state,  $\rho(\mathbf{r})$ , the total wave function of the hypothetical state will be  $\sqrt{\rho(\mathbf{r})/N}$ , where  $N$  is the number of electrons. It measures the intrinsic dimensions peculiar to the system, when both the quantum and electrostatic effects are completely excluded. The total kinetic energy of the hypothetical state is simply the Weizsäcker kinetic energy  $T_w[\rho]$ ,<sup>32</sup>

$$E_s[\rho] \equiv T_w[\rho] = \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r}. \quad (4)$$

with  $\nabla \rho(\mathbf{r})$  being the density gradient. Notice that the definition of steric energy above is consistent with Weisskopf's original attribution of the steric effect to the "kinetic energy pressure" because steric energy in Eq. (4) is kinetic and its density has the dimension of pressure.<sup>33</sup> Further, it is exclusive because it is separate from the other effects, both quantum and electrostatic.  $E_s$  is non-negative everywhere and, therefore, repulsive in nature. In addition, it vanishes for the homogeneous electron gas. It is extensive, i.e., larger system the larger the steric energy, because it is homogeneous of degree one in density scaling.<sup>34-40</sup> This novel quantification of the steric effect has been applied to a number of systems<sup>31,41-52</sup>, such as conformational changes of small molecules,  $S_N2$  reactions, chained and branched alkanes, experimental electron densities of crystals, cis effect, the anomeric effect, water clusters, weak interactions, etc.

This quantification of steric energy in Eq. (4) is closely related to the Fisher information in information theory<sup>22</sup>

$$I_F = \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} = 8 E_S[\rho]. \quad (5)$$

Other quantities from the information-theoretic approach in DFRT are the Shannon entropy,<sup>21</sup>

$$S_S = -\int \rho(\mathbf{r}) \ln \rho(\mathbf{r}) d\mathbf{r} \quad (6)$$

and the Ghosh-Berkowitz-Parr (GBP) entropy,<sup>23</sup>

$$S_{GBP} = \int \frac{3}{2} k \rho(\mathbf{r}) \left[ c + \ln \frac{t(\mathbf{r}; \rho)}{t_{TF}(\mathbf{r}; \rho)} \right] d\mathbf{r} \quad (7)$$

where  $t(\mathbf{r}, \rho)$  is the kinetic energy density, which is related to the total kinetic energy  $T_S$  via

$$\int t(\mathbf{r}; \rho) d\mathbf{r} = T_S, \quad (8)$$

and  $t_{TF}(\mathbf{r}; \rho)$  is the Thomas-Fermi kinetic energy density given by

$$t_{TF}(\mathbf{r}; \rho) = c_K \rho^{5/3}(\mathbf{r}), \quad (9)$$

with  $k$  as the Boltzmann constant (set to be unity for convenience in this work),

$$c = \frac{5}{3} + \ln \frac{4\pi c_K}{3}, \text{ and } c_K = \frac{3}{10} (3\pi^2)^{2/3}. \quad (10)$$

Following the original literature,<sup>15,16</sup> we use

$$t(\mathbf{r}; \rho) = \sum_i \frac{1}{8} \frac{\nabla \rho_i \cdot \nabla \rho_i}{\rho_i} - \frac{1}{8} \nabla^2 \rho, \quad (11)$$

where  $\rho_i$  are the orbital densities. Shannon entropy measures the spatial delocalization of the electron density, whereas Fisher information is a gauge of the sharpness or concentration of the electron density distribution.<sup>53</sup> The GBP entropy was resulted from the reformulation the ground-state density-functional theory into a macroscopic local version of thermodynamics to describe the behavior of microscopic electrons with the introduction of the concepts of local temperature.<sup>23</sup>

These three quantities, which were defined by the electron density  $\rho(\mathbf{r})$ , can also be re-defined using the shape function  $\sigma(\mathbf{r})$ , defined as  $\sigma(\mathbf{r}) \equiv \rho(\mathbf{r})/N$ , leading to the new information-theoretic quantities denoted by  $S_\sigma$  and  $I_\sigma$ , respectively,<sup>26</sup>

$$S_\sigma = -\int \sigma(\mathbf{r}) \ln \sigma(\mathbf{r}) d\mathbf{r} \quad (11)$$

and

$$I_\sigma = \int \frac{|\nabla \sigma(\mathbf{r})|^2}{\sigma(\mathbf{r})} d\mathbf{r}. \quad (12)$$

With the relationship between the electron density and shape function, one can readily prove that<sup>18</sup>

$$S_\sigma = \frac{S_S}{N} + \ln N \quad (13)$$

and

$$I_\sigma = \frac{I_F}{N}. \quad (14)$$

Recently, information gain  $I_G$ , also called Kullback–Leibler divergence, relative entropy, or information divergence, of a molecular system due to its formation from composing ingredients,

$$I_G = \sum_A \int \rho_A \ln \frac{\rho_A}{\rho_A^0} d\mathbf{r}, \quad (15)$$

was found to be a quantitatively reliable measure of electrophilicity, nucleophilicity, and regioselectivity.<sup>11-15</sup> In Eq. (15),  $\rho_A$  is the electron density on Atom (or Group) A in a molecule, whose total electron density is  $\rho$ , and  $\rho_A^0$  is the counterpart of Atom (or Group) A in the reference state. This result was from the previous work by Nalewajski and Parr,<sup>54-57</sup> showing that minimizing Eq. (15) gives rise to the “stockholder partition” of the electron density proposed by Hirshfeld,<sup>58</sup>

$$\rho_A = \frac{\rho_A^0}{\sum_A \rho_A^0} \rho. \quad (16)$$

If a new variable,  $x = (\rho_A - \rho_A^0)/\rho_A$ , is introduced with the help that  $\ln \frac{1}{1-x} \approx x$  as the first-order approximation, Eq. (15) can be simplified to become<sup>11</sup>

$$I_G \approx \sum_A \int (\rho_A - \rho_A^0) dr = -\sum_A q_A \equiv 0 \quad (17)$$

where  $q_A$  is the Hirshfeld charge on Atom (or Group) A. This result suggests that under the first-order approximation, the minimal information gain not only yields the Hirshfeld charge distribution but it also simply vanishes. We called this result the *information conservation principle*. Its effectiveness has been verified in a series of carbon and nitrogen containing systems.<sup>11-12</sup> It has also been successfully applied to examine the reactivity and regioselectivity for electrophilic aromatic substitution reactions. A novel explanation was obtained about the nature and origin of the so-called *ortho/para* and *meta* group directing phenomena for those reactions.<sup>13-15</sup>

In this work, we apply these formulations to bimolecular nucleophilic substitution ( $S_N2$ ) reactions and examine their effectiveness of these quantities in predicting chemical reactivity for these systems. We are also interested in appraising the strong scaling properties for them in these systems and compare them with those unveiled from other systems reported elsewhere.

### 3. Computational Details

We considered, in this work, the transition state and reactant/product complex in both the gas phase and aqueous solvent for a total of 21 self-exchange  $S_N2$  reactions of  $R_1R_2R_3C-F$  reacting with the fluoride anion, where  $R_1$ ,  $R_2$ , and  $R_3$  represent substituting alkyl groups such as  $-H$ ,  $-CH_3$ ,  $-C_2H_5$ ,  $-C_3H_7$ , and  $-C_4H_9$  groups (Scheme 1). We denote these systems in terms of three substituted groups in the format of  $(R_1R_2R_3)$ . The  $S_N2$  reaction in gas and solvent phases is known to proceed via a pre-transition state association type of the molecular complex.<sup>27,28</sup> All calculations were performed at the DFT B3LYP/6-311+G(d) level of theory.<sup>59-61</sup> Calculations of the reactions paths were performed with Gaussian 09 package version D01<sup>62</sup> with tight SCF convergence and ultra-fine integration grids. For transition-state structure searches, the quadratic synchronous transit approach<sup>63</sup> was employed and single-point frequency calculation was followed to ensure that the final structure obtained has only one imaginary

frequency and that the vibration mode of the negative frequency corresponds to the bond formation and bond breaking process of a  $S_N2$  reaction. In addition, an intrinsic reaction coordinate (IRC)<sup>64</sup> run was ensued to verify the relevance of transition-state structures. As an example, Scheme 1 shows the IRC result for the system of  $R_1=R_2=R_3=H$ . To simulate the solvent effect, an implicit solvent model, the conductor-like polarizable continuum model (CPCM),<sup>65</sup> was utilized in this study.

The Multiwfn 3.2 program<sup>66</sup> developed by one of the present authors was used to calculate the information-theoretic quantities in Eqs. (4)-(17) using the checkpoint file from the above Gaussian calculations as the input file. To obtain the electron density for the isolated state, we employed the spherically-averaged electron density of the neutral atom at the same level of theory. Besides neutral molecules, we also examine the scaling properties of information-theoretic quantities for atoms in molecules. To that end, three schemes to perform the atomic partition, Becke's fuzzy atom approach,<sup>67</sup> Bader's zero-flux atoms-in-molecules criterion,<sup>68</sup> and Hirshfeld's stockholder approach,<sup>58</sup> are possible. As have been demonstrated earlier,<sup>26</sup> these three approaches yield similar results. In this work, we choose the Hirshfeld's stockholder approach<sup>58</sup> to partition atoms in molecules. The molecular and atomic values of the steric energy were obtained through Eq. (5). Other energetic components in Eqs (1) - (3) were gathered from the Gaussian calculation using the internal option keyword of IOp(5/33=1).

#### 4. Results and Discussion

Table 1 exhibits the molecular value of Shannon entropy, Fisher information, GBP entropy and information gain for the 21 reactants of  $S_N2$  reactions in aqueous solvent studied in this work. Also shown in the Table is the total number of electrons  $N$  in the systems as well as molecular values of the shape function counterpart for Shannon entropy and Fisher information,  $S_\sigma$  and  $I_\sigma$ . The correlation coefficient  $R^2$  between  $N$  and these information-theoretic quantities is displayed in the last row of the

Table. As can be seen from the Table, all quantities shown in the Table,  $S_S$ ,  $S_{\sigma}$ ,  $I_F$ ,  $I_{\sigma}$ ,  $S_{GBP}$ , and  $I_G$ , are positive. The correlation between  $N$  and  $S_S$ ,  $I_F$ , and  $S_{GBP}$  is perfect, each with  $R^2=1.000$ , indicating that for systems studied in this work these three information-theoretic quantities are linearly extensive with respect to the total population of electrons and the system size. Correlations between  $N$  and  $S_{\sigma}$ ,  $I_{\sigma}$ , and  $I_G$  are not perfect but they are still markedly strong. Similar linear strong relationships were found for reactions in the gas phase (results not shown).

Table 2 presents the atomic value of Shannon entropy, Fisher information, GBP entropy and information gain on the reactive carbon center for the 21 transition state structures of bimolecular nucleophilic substitution reactions in gas phase studied in this work. The atomic value was calculated using Hirshfeld's shareholder partition scheme.<sup>58</sup> Also shown in the Table is the Shannon entropy and Fisher information using the shape function as the local function,  $S_{\sigma}$  and  $I_{\sigma}$ , plus the Hirshfeld charge and the total electron population,  $N_C$ , on the reactive carbon atom. In addition, the barrier height (activation energy) of these reactions is tabulated as well. The purpose of this Table is to correlate atomic values of those information quantities on the reactive carbon center with the reaction barrier height. Shown in the last row of the Table are the correlation coefficients  $R^2$  of these relationships. One can see that all these quantities are remarkable indicators of the reaction barrier height, all with  $R^2$  larger than 0.95. Recall that in our previous work,<sup>14,15</sup> the information gain and Hirshfeld charge were employed to quantify reactivity with strong correlations between them and the barrier height disclosed. Our results from this work suggest that for  $S_N2$  reactions, at least for systems studied here, there are more descriptors of the reactivity from the information-theoretic approach. These results also indicate that atomic values of the quantities from the information-theoretic approach at the reaction center are strongly correlated. For example, the correlation coefficient between  $S_S$  and  $I_F$ , between  $I_F$  and  $S_{GBP}$ , and between  $I_{\sigma}$  and the Hirshfeld charge is 0.999, 0.999, 0.998, respectively. In agreement with our previous results,<sup>14,15</sup> these strong correlations among different quantities are merely another testament of the

fact that the electron density determines everything in the ground state in density functional theory. All information-theoretic quantities are functions of the electron density, so it is not unusual that these quantities are interrelated, as have been theoretically proved<sup>53</sup> and numerically demonstrated elsewhere.<sup>25,26</sup> Given their vastly different analytical forms, this kind of correlations from information-theoretic quantities is not trivial and has never been reported in the literature. They serve as one of the main results in this work.

Do we have to make use of the quantities from the transition state structure to observe above correlations? Do they still work well if results from reactants are employed? What if the solvent effect is taken into consideration? It is well known that the solvent effect might have a huge impact on the reactivity indices.<sup>69</sup> To answer these questions, shown in Table 3 are the correlation coefficients of the above strong correlations with different structures in different conditions. In the Table, we employed both the reactant and transition state structures, and considered the condition in both gas phase and in aqueous solvent. In all the cases, the strong correlations discussed above are still in existence. We do not observe noticeable differences in their correlation coefficients. These results demonstrate that one does not have to use the transition state structure, nor does one have to do it in the gas or solvent phase. Indeed, the only exception in Table is information gain of the transition state structure in the aqueous solvent, whose  $R^2$  is much smaller than others. In agreement with our previous work,<sup>15</sup> Table 3 shows that using the reactant alone, one is able to reliably predict the barrier height with the Hirshfeld charge and information gain. Moreover, this Table also suggests that other quantities from the reactant can do the similar job with the similar accuracy.

Shown in Figure 1 are the numerical results of three information-theoretic quantities,  $I_\sigma$ ,  $S_{\text{GBP}}$ , and  $I_G$ , for a total of 346 carbon atoms from 21 reactant and transition state structures as a function of their total electron population in both gas (Fig. 1a - 1c) and solvent (Fig. 1d - 1f) phases. Again, we employed Hirshfeld's shareholder approach to perform the atom-in-molecule partition. These scaling

results of atomic values with respect to the total electron population for Shannon entropy, Fisher information, and GBP entropy are similar to those shown in our previous study for other systems,<sup>15,26</sup> including transition state structures. What looks to be an exception is the result for the GBP entropy. In both the gas and solvent cases,  $R^2$  values of  $S_{\text{GBP}}$  vs.  $N_{\text{C}}$  are both larger than 0.95, enough to claim that their correlations are strong. However, looking at Figs. 1b and 1e, it appears that there exist two parallel lines between  $S_{\text{GBP}}$  vs.  $N_{\text{C}}$  with the same slope but different intercepts. The origin of this bi-linearity might be coming from the fact that there exist two kinds of carbon atoms, one on the reaction center and the other from the substituting groups.<sup>12</sup>

Following our previous efforts,<sup>30,31</sup> we decomposed the reaction barrier height into three components using the two distinct energy partition schemes, Eqs. (1) and (3). The purpose of this analysis is to examine the origin of the activation energy by finding out (i) whether or not there is a dominant energy contributor and (ii) if there is a strong correlation between the barrier height and any of the energy components. Table 4 shows the numerical results for the 21  $S_{\text{N}}2$  substitution reactions studied in this work in aqueous solvent. Similar results were also obtained for these species in gas phase. As can be seen from the Table, we have the barrier height  $\Delta E$  is always positive,  $\Delta E > 0$ , so are  $\Delta E_{\text{e}}$  and  $\Delta E_{\text{xc}}$  and  $\Delta E_{\text{s}}$ , indicating that these three components are positively contributing to the existence of the barrier height. Correlations of these three components with the barrier height in both gas (a-c) and solvent (e-f) phases are shown in Fig. 2. For the gas phase, the correlation coefficient between  $\Delta E$  and  $\Delta E_{\text{s}}$ , between  $\Delta E$  and  $\Delta E_{\text{xc}}$ , and between  $\Delta E$  and  $\Delta E_{\text{e}}$  is 0.711, 0.965, and 0.990, respectively, whereas in aqueous solvent, it becomes 0.783, 0.824, and 0.944, respectively. There are two points to make in regard to the results in Fig. 2. At first, in both gas and solvent phases, it is the electrostatic effect that possesses the strongest linear correlation with the barrier height. This result is consistent with our results reported elsewhere for other systems.<sup>31</sup> Secondly, the reasonably strong correlation between  $\Delta E$  and  $\Delta E_{\text{s}}$  in both cases suggest that the steric effect is a good indicator of the chemical reactivity for this

category of systems. The bigger the groups ( $R_1$ ,  $R_2$ , and  $R_3$ ) attached to the reactive center, the larger the barrier height and thus the slower the  $S_N2$  reaction. This result agrees well with the experimental findings available in textbooks as well as our theoretical work in the literature.<sup>29</sup>

With the two energy components from Eqs. (1) and (3) to fit with the  $S_N2$  reaction barrier height, we obtained slightly better results in both gas and solvent cases. Figure 3 exhibits the results, with  $R^2$  equal to or better than 0.99 in gas phase and  $R^2 > 0.95$  in aqueous solvent. These results are consistent with our previous studies for other systems such as conformational changes<sup>31</sup> and electrophilic aromatic substitution reactions,<sup>16</sup> confirming for one more time that the electrostatic interaction is the dominant contributor in determining the barrier height, but contributions from other effects such as exchange-correlation and steric interactions, though less significant, are indispensable.

Notice that, using virial theorem, the energy difference  $\Delta E$  between two systems should satisfy<sup>1</sup>

$$\Delta E = -\Delta T = \frac{1}{2}\Delta V, \quad (18)$$

where  $T$  and  $V$  denote the total kinetic and potential energies, respectively, of the systems in concern. Eq. (18) suggests that  $\Delta E$  is equal to either the entire kinetic energy difference or half of the total potential energy difference. These energy components are, however, not chemically meaningful. We often wish to obtain insights from such effects as steric, electrostatic, or quantum, which are missing in Eq. (18). More importantly, Eq. (18) does not work for density functional theory,<sup>1-4</sup> because a portion of the kinetic energy,  $T_c[\rho]$ , has already been incorporated in the exchange-correlation energy  $E_{xc}$ , making the DFT version of the virial theorem much more complicated.<sup>71-74</sup>

Before closing, we should caution the readers about the general applicability of these strong linear relationships obtained from this work to other systems. Based on the results we obtained previously, correlations from scaling properties<sup>15,25,26</sup> and energy partition analysis<sup>29-31,41-51</sup> shown in Figs. 1-3 should be generally applicable. However, for the strong correlations for the barrier height with various information-theoretic quantities, based on what we know for carbon and nitrogen systems,<sup>11,12</sup>

it is highly possible that they are not generally applicable, same as what we have observed for the acidity case.<sup>5-10</sup> Many competing terms showing good linear correlations in one system might suggest that a complete picture is still lacking. The search for a more general quantification approach using these and other DFRT descriptors is still in the works.

## 5. Concluding Remarks

Can we make use of the quantities from the information-theoretic approach to reliably predict the reactivity for bimolecular nucleophilicity substitution ( $S_N2$ ) reactions? Are the scaling properties observed in other systems for the quantities such as Shannon entropy, Fisher information, Ghosh-Berkowitz-Parr entropy and information gain still valid for  $S_N2$  reactions? What is the origin of the barrier height of these reactions? To answer these questions, in this work, we examined a total of 21 self-exchange  $S_N2$  reactions for compounds with the general formula of  $R_1R_2R_3C-F$  reacting with the fluoride anion, where  $R_1$ ,  $R_2$ , and  $R_3$  represent substituting alkyl groups such as  $-H$ ,  $-CH_3$ ,  $-C_2H_5$ ,  $-C_3H_7$ , and  $-C_4H_9$  groups. We confirmed that there existed similar scaling properties for information-theoretic quantities in these reactive systems. The barrier height of the  $S_N2$  reactions can reliably be predicted by not only the Hirshfeld charge and information gain at the regioselective carbon atom, but also other reactivity descriptors such as Shannon entropy, Fisher information, and Ghosh-Berkowitz-Parr entropy at the same carbon atom. The energy decomposition analysis result is in excellent agreement with our previous work reported elsewhere for other systems, suggesting that it is the electrostatic interaction plays the dominant role, while the roles played by exchange-correlation and steric effects are less significant but indispensable. These novel findings provide additional insights for the better understanding of the chemical reactivity as well as the factors governing the chemical reactivity for this extremely important category of chemical transformations, which finds widespread applications in

organic synthesis, enzyme chemistry, catalyst design and many other studies in different scientific disciplines.

### Acknowledgements

CYR acknowledges support from the National Natural Science Foundation of China (21503076), the Chinese Scholarship Council, the Construct Program of the Key Discipline in Hunan Province, and the Aid Program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province, PR China. PWA thanks NSERC for support from a Discovery Grant and an E. W. R. Steacie Fellowship.

### References

1. R. G. Parr and W. T. Yang, *Density Functional Theory for Atoms and Molecules*, Oxford University Press, New York, 1989.
2. P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793.
3. P. K. Chattaraj, U. Sarkar and D. R. Roy, *Chem. Rev.*, 2006, **106**, 2065.
4. S. B. Liu, *Acta Phys.-Chim. Sin.*, 2009, **25**, 590.
5. S.B. Liu, and L.G. Pedersen, *J. Phys. Chem. A*, 2009, **113**, 3648.
6. S.B. Liu, C.K. Schauer, and L.G. Pedersen, *J. Chem. Phys.*, 2009, **131**, 164107.
7. S.K. Burger, S.B. Liu, and P.W. Ayers, *J. Phys. Chem. A*, 2011, **115**, 1293.
8. Y. Huang, L.H. Liu, W.H. Liu, S.G. Liu, and S.B. Liu, *J. Phys. Chem. A*, 2011, **115**, 14697
9. D.B. Zhao, C.Y. Rong, D.L. Yin, and S.B. Liu, *J. Theor. Comp. Chem.*, 2013, **12**, 1350034.
10. Y. Huang, L.H. Liu, and S.B. Liu, *Chem. Phys. Lett.* 2012, **527**, 73-78.
11. S. B. Liu, C. Y. Rong and T. Lu, *J. Phys. Chem. A*, 2014, **118**, 3698.
12. X. Y. Zhou, C. Y. Rong, T. Lu and S. B. Liu, *Acta Phys.-Chim. Sin.*, **2014**, 30, 2055.
13. S. B. Liu, *J. Chem. Phys.*, 2014, **141**, 194109.
14. S. B. Liu, *J. Phys. Chem. A*, 2015, **119**, 3107.
15. W.J. Wu, Z.M. Wu, C.Y. Rong, Y. Huang, and S.B. Liu, *J. Phys. Chem. A*, DOI: 10.1021/acs.jpca.5b04309.

16. R. G. Parr, L. von Szentpaly, and S.B. Liu, *J. Am. Chem. Soc.*, 1999, **121**, 1922.
17. S. Wu, S.B. Liu, S. Sim, and L.G. Pedersen, *J. Phys. Chem. Lett.*, 2012, **3**, 2293.
18. P.P. Zhou, P.W. Ayers, S.B. Liu, and T.L. Li, *Phys. Chem. Chem. Phys.*, 2012, **14**, 9890.
19. N. Kumar, S.B. Liu, and P.M. Kozlowski, *J. Phys. Chem. Lett.* 2012, **3**, 1035.
20. S.B. Liu, D. H. Ess, and C. K. Schauer, *J. Phys. Chem. A*, 2011, **115**, 4738.
21. C. E. Shannon, *Bell Syst. Tech. J.*, 1948, **27**, 379.
22. R. A. Fisher, *Proc. Cambridge Philos. Soc.*, 1925, **22**, 700.
23. S. K. Ghosh, M. Berkowitz and R. G. Parr, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**, 8028.
24. S. Kullback, *Information Theory and Statistics*, Dover, Mineola, NY, 1997.
25. C. Rong, T. Lu, P. K. Chattaraj and S. Liu, *Indian J. Chem., Sect. A: Inorg., Bio-inorg., Phys., Theor. Anal. Chem.*, 2014, **53**, 970.
26. C.Y. Rong, T. Lu, P.W. Ayers, P.K. Chattaraj, and S.B. Liu, *Phys. Chem. Chem. Phys.*, 2015, **17**, 4977; *Phys. Chem. Chem. Phys.*, 2015, **17**, 11110-11111.
27. J.E. McMurry, *Organic Chemistry*, Wadsworth, Belmont, 1992.
28. M.B. Smith, J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, Wiley-Interscience, New York, 2007.
29. S. B. Liu, H. Hu and L. G. Pedersen, *J. Phys. Chem. A*, 2010, **114**, 5913.
30. S. B. Liu, *J. Chem. Phys.*, 2007, **126**, 244103.
31. S. B. Liu, *J. Phys. Chem. A*, 2013, **117**, 962.
32. C. F. von Weizsäcker, *Z. Phys.*, 1935, **96**, 431.
33. V. F. Weisskopf, *Science*, 1975, **187**, 605.
34. M. Levy and J. P. Perdew, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **32**, 2010.
35. S. B. Liu, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1996, **54**, 1328.
36. S. B. Liu and R. G. Parr, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1996, **53**, 2211.
37. S. B. Liu, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1996, **54**, 1328.
38. S. B. Liu, R. C. Morrison and R. G. Parr, *J. Chem. Phys.*, 2006, **125**, 174109.
39. S. B. Liu, P. W. Ayers, and R. G. Parr, *J. Chem. Phys.*, 1999, **111**, 6197.
40. S. B. Liu and R. G. Parr, *J. Phys. Chem. A*, 2007, **111**, 10422.
41. S. B. Liu and N. Govind, *J. Phys. Chem. A*, 2008, **112**, 6690.
42. S. B. Liu, N. Govind, and L. G. Pedersen, *J. Chem. Phys.*, 2008, **129**, 094104.
43. M. Torrent-Sucarrat, S. B. Liu, and F. De Proft, *J. Phys. Chem. A*, 2009, **113**, 3698.
44. D. H. Ess, S. B. Liu, and F. De Proft, *J. Phys. Chem. A*, 2010, **114**, 12952.

45. V. G. Tsirelson, A. I. Stash, and S. B. Liu, *J. Chem. Phys.*, 2010, **133**, 114110.
46. Y. Huang, A. G. Zhong, Q. S. Yang, and S. B. Liu, *J. Chem. Phys.* 2011, **134**, 084103.
47. R. O. Esquivel, S. B. Liu, J. C. Angulo, J. S. Dehesa, J. Antolin, and M. Molina-Espiritu, *J. Phys. Chem. A*, 2011, **115**, 4406.
48. D. B. Zhao, C. Y. Rong, S. Jenkins, S. R. Kirk, D. L. Yin, and S. B. Liu, *Acta Phys.-Chim. Sin.*, 2013, **29**, 43.
49. V. G. Tsirelson, A. I. Stash, and S. B. Liu, *Comput. Theor. Chem.*, 2013, **1006**, 92.
50. Y. J. Wang, D. B. Zhao, C. Y. Rong, and S. B. Liu, *Acta Phys.-Chim. Sin.* 2014, **29**, 2173.
51. D. Fang, J.-P. Piquemal, S. B. Liu, and G. A. Cisneros, *Theor. Chem. Acc.* 2014, **133**, 1484.
52. S.B. Liu, and C.K. Schauer, *J. Chem. Phys.*, 2015, **142**, 054107.
53. S. B. Liu, *J. Chem. Phys.*, 2007, **126**, 191107.
54. R. F. Nalewajski and R. G. Parr, *J. Phys. Chem. A*, 2001, **105**, 7391.
55. R. F. Nalewajski and R. G. Parr, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8879.
56. R. G. Parr, P. W. Ayers and R. F. Nalewajski, *J. Phys. Chem. A*, 2005, **109**, 3957.
57. P. W. Ayers, *Theor. Chem. Acc.*, 2006, **115**, 370.
58. F. L. Hirshfeld, *Theor. Chem. Acc.*, 1977, **44**, 129.
59. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
60. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.
61. R. Ditchfield, W. J. Hehre and J. Pople, *J. Chem. Phys.*, 1971, **54**, 724.
62. M. J. Frisch, G. W. Trucks and H. B. Schlegel, et al., Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford, CT, 2009.
63. C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, *J. Comp. Chem.*, 1996, **17**, 49.
64. K. Fukui, *Acc. Chem. Res.*, 1981, **14**, 363.
65. M. Cossi, N. Rega, G. Scalmani, and V. Barone, *J. Comp. Chem.*, 2003, **24**, 669-81.
66. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580.
67. A. D. Becke, *J. Chem. Phys.*, 1988, **88**, 2547.
68. R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, Oxford, 1990.
69. C. Y. Rong, S. X. Lian and S. B. Liu, *J. Nat. Sci. Hunan Norm. Univ.*, 2014, **37**, 31.
70. M. Levy and J.P. Perdew, *Phys. Rev. A* 1985, **32**, 2010.
71. S.B. Liu and R.G. Parr, *Phys. Rev. A* 1996, **53**, 2211.
72. S.B. Liu, *Phys. Rev. A* 1996, **54**, 1328.
73. S.B. Liu, A. Nagy, and R.G. Parr, *Phys. Rev. A* 1999, **59**, 1131.
74. S.B. Liu, R.C. Morrison, and R.G. Parr, *J. Chem. Phys.*, 2006, **125**, 174109.

Table 1. The molecular value of Shannon entropy, Fisher information, GBP entropy and information gain for the 21 reactants of  $S_N2$  reactions in aqueous solvent studied in this work. Also shown is the shape function counterpart of the Shannon entropy and Fisher information. The correlation coefficient  $R^2$  between these information-theoretic quantities and the total electron number  $N$  of the molecules is displayed in the last row. Atomic units.

| $R_1R_2R_3$          | $N$ | $S_S$   | $S_\sigma$ | $I_F$    | $I_\sigma$ | $S_{GBP}$ | $I_G$  |
|----------------------|-----|---------|------------|----------|------------|-----------|--------|
| $C_2H_5C_2H_5C_2H_5$ | 76  | 78.457  | 5.363      | 2928.490 | 38.533     | 494.871   | 2.5051 |
| $C_2H_5C_2H_5CH_3$   | 68  | 68.359  | 5.225      | 2676.706 | 39.363     | 442.883   | 2.3461 |
| $C_2H_5C_2H_5H$      | 60  | 58.301  | 5.066      | 2425.464 | 40.424     | 391.018   | 2.2217 |
| $C_2H_5CH_3CH_3$     | 60  | 58.255  | 5.065      | 2424.913 | 40.415     | 390.892   | 2.1888 |
| $C_2H_5CH_3H$        | 52  | 48.056  | 4.875      | 2173.225 | 41.793     | 338.927   | 2.0186 |
| $C_2H_5HH$           | 44  | 37.844  | 4.644      | 1921.361 | 43.667     | 286.927   | 1.8433 |
| $C_3H_7C_2H_5C_2H_5$ | 84  | 88.639  | 5.486      | 3180.458 | 37.863     | 546.898   | 2.6594 |
| $C_3H_7C_2H_5CH_3$   | 76  | 78.550  | 5.364      | 2928.701 | 38.536     | 494.918   | 2.5049 |
| $C_3H_7C_2H_5H$      | 68  | 68.407  | 5.225      | 2677.132 | 39.370     | 442.982   | 2.3269 |
| $C_3H_7C_3H_8C_2H_5$ | 92  | 98.829  | 5.596      | 3432.452 | 37.309     | 598.933   | 2.8204 |
| $C_3H_7C_3H_7C_3H_7$ | 100 | 109.008 | 5.695      | 3684.408 | 36.844     | 650.957   | 2.9720 |
| $C_3H_7C_3H_7CH_3$   | 84  | 88.735  | 5.487      | 3180.683 | 37.865     | 546.949   | 2.6609 |
| $C_3H_7C_3H_7H$      | 76  | 78.661  | 5.366      | 2929.373 | 38.544     | 495.070   | 2.5183 |
| $C_3H_7CH_3CH_3$     | 68  | 68.444  | 5.226      | 2676.913 | 39.366     | 442.927   | 2.3469 |
| $C_3H_7CH_3H$        | 60  | 58.242  | 5.065      | 2425.211 | 40.420     | 390.960   | 2.1744 |
| $C_3H_7HH$           | 52  | 48.039  | 4.875      | 2173.392 | 41.796     | 338.966   | 2.0016 |
| $C_4H_9HH$           | 60  | 58.289  | 5.066      | 2425.606 | 40.427     | 391.043   | 2.1816 |
| $CH_3CH_3CH_3$       | 52  | 48.138  | 4.877      | 2173.099 | 41.790     | 338.893   | 2.0283 |
| $CH_3CH_3H$          | 44  | 37.998  | 4.648      | 1921.620 | 43.673     | 286.974   | 1.8974 |
| $CH_3HH$             | 36  | 27.656  | 4.352      | 1669.378 | 46.372     | 234.892   | 1.6881 |
| $HHH$                | 28  | 17.586  | 3.960      | 1418.115 | 50.647     | 182.985   | 1.6020 |
| $R^2$                |     | 1.000   | 0.942      | 1.000    | 0.855      | 1.000     | 0.997  |

Table 2. The atomic value of Shannon entropy, Fisher information, GBP entropy and information gain on the reactive carbon center for the 21 transition state structures of  $S_N2$  reactions in gas phase studied in this work. Also shown is the shape function counterpart of the Shannon entropy and Fisher information. The Hirshfeld charge on the reaction carbon center and the barrier height of these reactions are tabulated as well. Atomic units except for the barrier height, which is in kcal/mol.

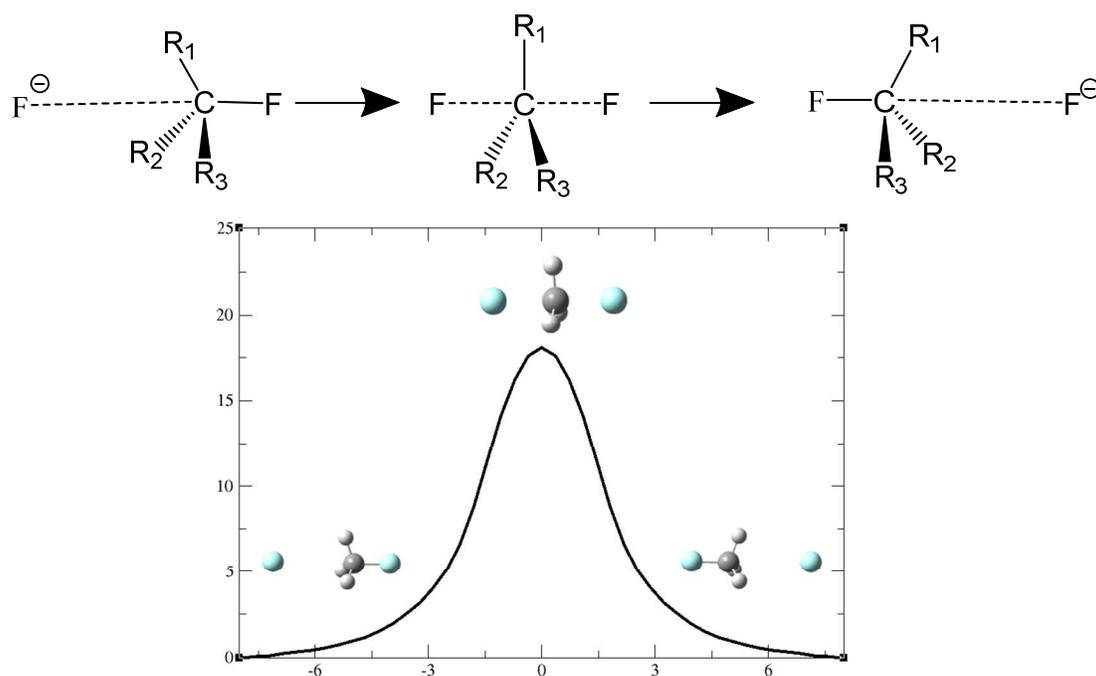
| $R_1R_2R_3$          | $S_S$  | $S_G$  | $I_F$    | $I_G$   | $S_{GBP}$ | $I_G$  | Hirshfeld | $N_C$  | Barrier |
|----------------------|--------|--------|----------|---------|-----------|--------|-----------|--------|---------|
| $C_2H_5C_2H_5C_2H_5$ | 4.1304 | 2.4720 | 246.0801 | 42.1417 | 37.9468   | 0.0876 | 0.1607    | 5.8394 | 24.87   |
| $C_2H_5C_2H_5CH_3$   | 4.1403 | 2.4737 | 246.0947 | 42.1415 | 37.9529   | 0.0880 | 0.1603    | 5.8397 | 25.48   |
| $C_2H_5C_2H_5H$      | 4.4002 | 2.5204 | 246.5392 | 41.8373 | 38.3825   | 0.0952 | 0.1072    | 5.8928 | 19.04   |
| $C_2H_5CH_3CH_3$     | 4.1558 | 2.4764 | 246.1133 | 42.1370 | 37.9640   | 0.0870 | 0.1592    | 5.8408 | 26.21   |
| $C_2H_5CH_3H$        | 4.4445 | 2.5281 | 246.6084 | 41.8325 | 38.4137   | 0.0950 | 0.1049    | 5.8951 | 19.93   |
| $C_2H_5HH$           | 4.7668 | 2.5844 | 247.1427 | 41.5653 | 38.8458   | 0.1021 | 0.0541    | 5.9459 | 14.04   |
| $C_3H_7C_2H_5C_2H_5$ | 4.1300 | 2.4719 | 246.0733 | 42.1409 | 37.9457   | 0.0886 | 0.1607    | 5.8393 | 24.79   |
| $C_3H_7C_2H_5CH_3$   | 4.1397 | 2.4736 | 246.0879 | 42.1413 | 37.9515   | 0.0878 | 0.1604    | 5.8396 | 25.39   |
| $C_3H_7C_2H_5H$      | 4.3983 | 2.5201 | 246.5294 | 41.8395 | 38.3784   | 0.0962 | 0.1077    | 5.8923 | 19.13   |
| $C_3H_7C_3H_8C_2H_5$ | 4.1296 | 2.4718 | 246.0665 | 42.1404 | 37.9444   | 0.0884 | 0.1608    | 5.8392 | 24.72   |
| $C_3H_7C_3H_7C_3H_7$ | 4.1287 | 2.4717 | 246.0589 | 42.1407 | 37.9425   | 0.0883 | 0.1610    | 5.8390 | 24.65   |
| $C_3H_7C_3H_7CH_3$   | 4.1386 | 2.4734 | 246.0788 | 42.1409 | 37.9496   | 0.0876 | 0.1606    | 5.8394 | 25.31   |
| $C_3H_7C_3H_7H$      | 4.3956 | 2.5196 | 246.5137 | 41.8385 | 38.3753   | 0.0962 | 0.1080    | 5.8920 | 19.05   |
| $C_3H_7CH_3CH_3$     | 4.1545 | 2.4762 | 246.1041 | 42.1368 | 37.9620   | 0.0869 | 0.1594    | 5.8406 | 26.10   |
| $C_3H_7CH_3H$        | 4.4414 | 2.5275 | 246.5967 | 41.8366 | 38.4077   | 0.0962 | 0.1057    | 5.8943 | 20.03   |
| $C_3H_7HH$           | 4.7634 | 2.5838 | 247.1247 | 41.5624 | 38.8432   | 0.1020 | 0.0541    | 5.9459 | 14.49   |
| $C_4H_9HH$           | 4.7626 | 2.5837 | 247.1227 | 41.5620 | 38.8429   | 0.1019 | 0.0541    | 5.9459 | 14.87   |
| $CH_3CH_3CH_3$       | 4.1739 | 2.4795 | 246.1366 | 42.1383 | 37.9734   | 0.0862 | 0.1588    | 5.8412 | 27.13   |
| $CH_3CH_3H$          | 4.4647 | 2.5316 | 246.6172 | 41.8105 | 38.4417   | 0.0943 | 0.1015    | 5.8985 | 19.66   |
| $CH_3HH$             | 4.8138 | 2.5924 | 247.2093 | 41.5587 | 38.8801   | 0.1016 | 0.0516    | 5.9484 | 14.67   |
| $HHH$                | 5.2125 | 2.6606 | 247.8293 | 41.2701 | 39.3688   | 0.1056 | -0.0051   | 6.0051 | 9.22    |
| $R^2$                | 0.952  | 0.955  | 0.958    | 0.984   | 0.975     | 0.984  | 0.980     | 0.980  | ---     |

Table 3. Correlation coefficients between the reaction barrier height and information-theoretic quantities for the  $S_N2$  reactions studied in this work in both gas phase and aqueous solvent. The correlation is considered with respect to both the reactant complex and the transition state.

| $R^2$              | $S_S$ | $S_\sigma$ | $I_F$ | $I_\sigma$ | $S_{GBP}$ | $I_G$ | Hirshfeld | $N_C$ |
|--------------------|-------|------------|-------|------------|-----------|-------|-----------|-------|
| Reactant (gas)     | 0.945 | 0.946      | 0.955 | 0.946      | 0.952     | 0.974 | 0.951     | 0.951 |
| TS (gas)           | 0.952 | 0.955      | 0.958 | 0.984      | 0.975     | 0.984 | 0.980     | 0.980 |
| Reactant (solvent) | 0.962 | 0.964      | 0.968 | 0.980      | 0.976     | 0.986 | 0.979     | 0.979 |
| TS (solvent)       | 0.967 | 0.971      | 0.963 | 0.994      | 0.987     | 0.761 | 0.990     | 0.990 |

Table 4. Energy component results from the two decomposition approaches, Eqs. (1) and (3), for a total of 21  $S_N2$  reaction systems in aqueous solvent studied in this work. Units in kcal/mol.

| $R_1R_2R_3$          | $\Delta T_s$ | $\Delta E_{xc}$ | $\Delta E_e$ | $\Delta E_s$ | $\Delta E_q$ | $\Delta E$ |
|----------------------|--------------|-----------------|--------------|--------------|--------------|------------|
| $C_2H_5C_2H_5C_2H_5$ | -46.58       | 18.53           | 43.53        | 77.91        | -105.96      | 32.24      |
| $C_2H_5C_2H_5CH_3$   | -43.61       | 18.02           | 41.33        | 77.49        | -103.07      | 32.50      |
| $C_2H_5C_2H_5H$      | -35.11       | 15.75           | 27.97        | 77.12        | -96.49       | 27.07      |
| $C_2H_5CH_3CH_3$     | -44.33       | 18.06           | 42.62        | 79.07        | -105.35      | 32.81      |
| $C_2H_5CH_3H$        | -35.42       | 16.50           | 31.93        | 76.66        | -95.58       | 26.83      |
| $C_2H_5HH$           | -30.86       | 15.67           | 25.69        | 75.94        | -91.14       | 22.36      |
| $C_3H_7C_2H_5C_2H_5$ | -47.67       | 18.77           | 44.42        | 78.07        | -106.97      | 32.29      |
| $C_3H_7C_2H_5CH_3$   | -43.87       | 18.08           | 40.77        | 78.77        | -104.56      | 32.49      |
| $C_3H_7C_2H_5H$      | -34.16       | 16.36           | 29.64        | 77.26        | -95.06       | 26.80      |
| $C_3H_7C_3H_8C_2H_5$ | -47.49       | 18.62           | 43.93        | 78.23        | -107.10      | 32.33      |
| $C_3H_7C_3H_7C_3H_7$ | -46.11       | 18.21           | 43.03        | 78.10        | -106.00      | 32.35      |
| $C_3H_7C_3H_7CH_3$   | -44.81       | 18.21           | 41.77        | 78.91        | -105.51      | 32.51      |
| $C_3H_7C_3H_7H$      | -34.79       | 15.82           | 28.34        | 77.13        | -96.11       | 27.10      |
| $C_3H_7CH_3CH_3$     | -42.71       | 17.74           | 40.92        | 79.03        | -104.00      | 32.82      |
| $C_3H_7CH_3H$        | -34.91       | 16.49           | 31.27        | 76.19        | -94.60       | 26.86      |
| $C_3H_7HH$           | -29.95       | 15.43           | 24.59        | 75.72        | -90.23       | 22.26      |
| $C_4H_9HH$           | -27.42       | 14.27           | 21.54        | 75.80        | -88.95       | 22.27      |
| $CH_3CH_3CH_3$       | -41.79       | 17.16           | 41.80        | 78.70        | -103.33      | 33.14      |
| $CH_3CH_3H$          | -34.46       | 15.16           | 29.32        | 74.20        | -93.50       | 25.87      |
| $CH_3HH$             | -30.54       | 15.44           | 26.34        | 74.69        | -89.79       | 22.56      |
| $HHH$                | -18.87       | 10.67           | 10.99        | 75.02        | -83.22       | 18.56      |



Scheme 1. The identity  $S_N2$  reaction studied in this work.  $R_1$ ,  $R_2$ , and  $R_3$  are alky functional groups, which could be H,  $CH_3$ ,  $C_2H_5$ ,  $C_3H_8$ , or  $C_4H_{10}$ . The IRC calculation shown above was done at the B3LYP/6-311+G(d) level of theory for the system of  $R_1=R_2=R_3=H$  in aqueous solvent with the CPCM used as the implicit solvent model.

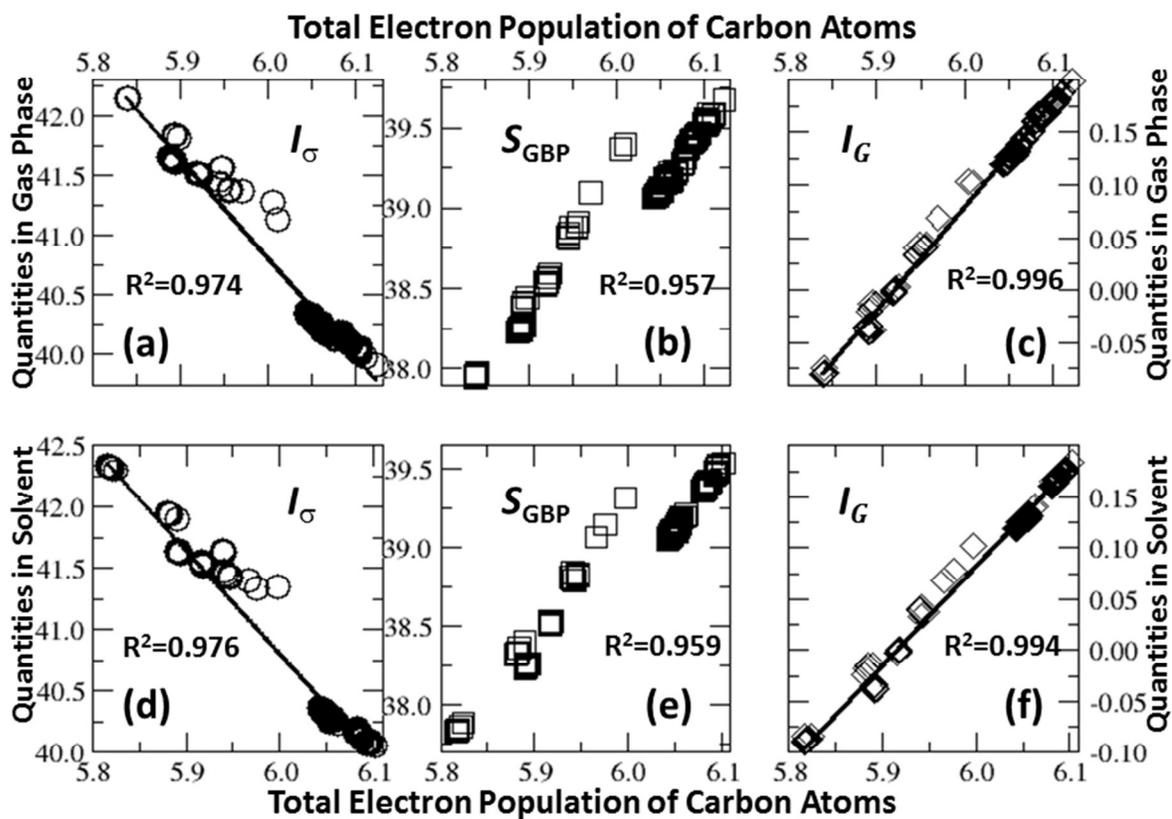


Figure 1. Strong linear scaling properties of informatics-theoretical quantities,  $I_\sigma$ ,  $S_{GBP}$ , and  $I_G$ , obtained for a total of 346 carbon atoms from the 21 reactant and transition state structures with respect to the total electron populations on the carbon atom using Hirshfeld's shareholder partition scheme in both gas (a-c) and solvent (d-f) phases.

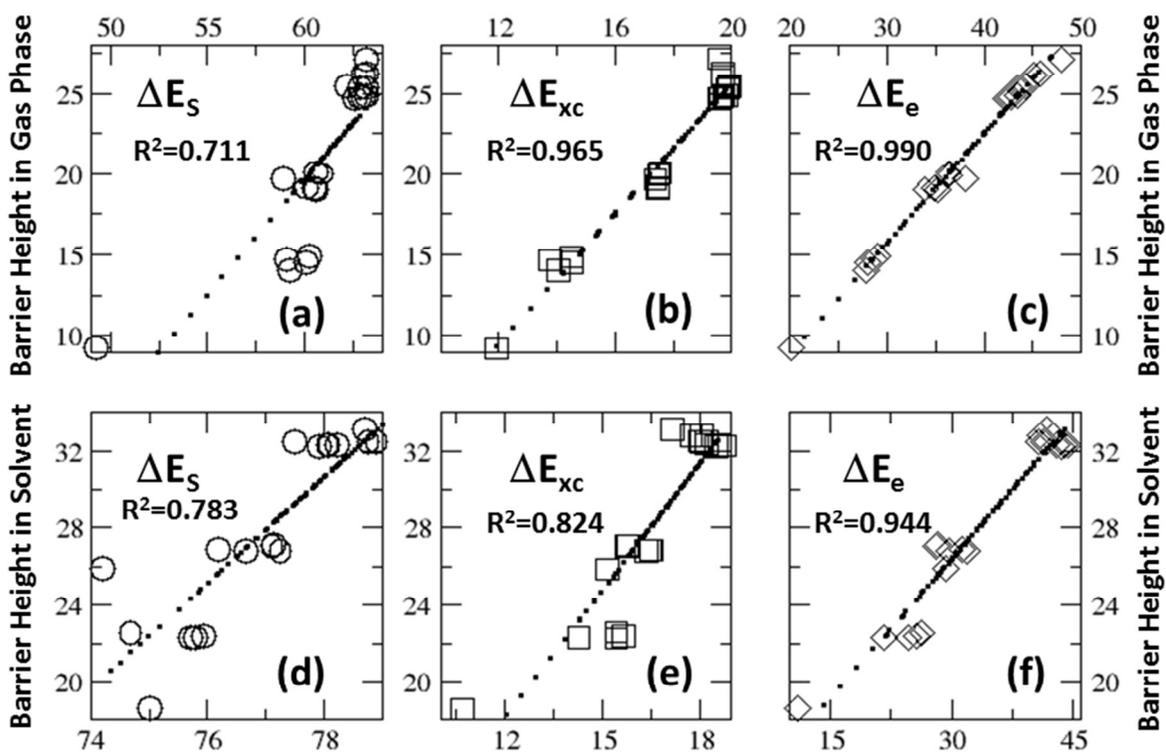


Figure 2. Strong linear correlations for 21 bimolecular nucleophilic substitution reactions between the computed barrier height and three individual energy components: (a,d) fitting with the steric energy component,  $\Delta E_s$ ; (b,e) fitting with exchange-correlation components  $\Delta E_{xc}$ ; and (c,f) fitting with the electrostatic term  $\Delta E_e$  and steric energy components  $\Delta E_s$  in both (a-c) gas and (d-f) solvent phases.

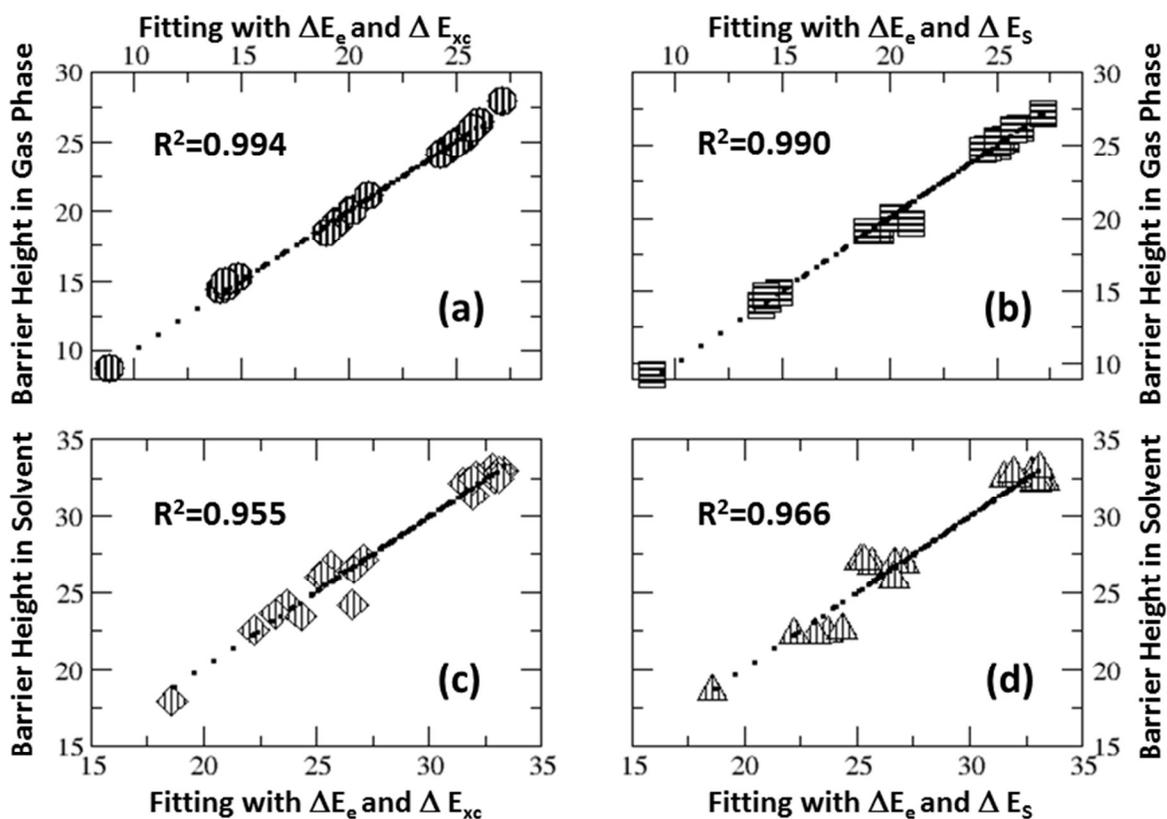
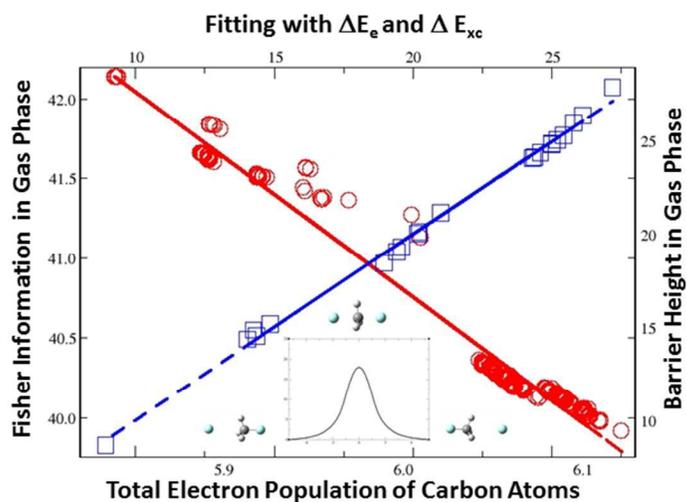


Figure 3. Strong linear correlations for 21 bimolecular nucleophilic substitution reactions between the computed barrier height and two-variable fitting: (a,c) with both electrostatic  $\Delta E_e$  and exchange-correlation components  $\Delta E_{xc}$  and (b,d): two-term fitting with both electrostatic  $\Delta E_e$  and steric energy components  $\Delta E_s$  in both (a-b) gas and (c-d) solvent phases.

## TOC Graphic



Strong linear correlations were unveiled between barrier heights of bimolecular nucleophilic substitution ( $S_N2$ ) reactions and quantities from the information-theoretic approach.