Computational Study of Chemical Reactivity Using Information-Theoretic Quantities from Density Functional Reactivity Theory for Electrophilic Aromatic Substitution Reactions

6 AUTHORS, INCLUDING:

Chunying Rong
Hu Nan Normal University
32 PUBLICATIONS **338** CITATIONS

SEE PROFILE

Tian Lu
Beijing Kein Research Center for Natural Scien…
31 PUBLICATIONS **852** CITATIONS

SEE PROFILE

Shubin Liu
University of North Carolina at Chapel Hill
155 PUBLICATIONS **5,453** CITATIONS

SEE PROFILE

Computational Study of Chemical Reactivity Using Information-Theoretic Quantities from Density Functional Reactivity Theory for Electrophilic Aromatic Substitution Reactions

Wenjie Wu,[†] Zemin Wu,[‡] Chunying Rong,*[,‡,§] Tian Lu,[∥] Ying Huang,*[,†] and Shubin Liu*[,‡,⊥]

[†]School of Pharmacy, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China

[‡]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

[§]Department of Chemistry, McMaster University, Hamilton, Ontario Canada, L8S 4M1

[∥]Beijing Kein Research Center for Natural Sciences, Beijing 100022, People's Republic of China

[⊥]Research Computing Center, University of North Carolina, Chapel Hill, North Carolina 27599-3420, United States

ABSTRACT: The electrophilic aromatic substitution for nitration, halogenation, sulfonation, and acylation is a vastly important category of chemical transformation. Its reactivity and regioselectivity is predominantly determined by nucleophilicity of carbon atoms on the aromatic ring, which in return is immensely influenced by the group that is attached to the aromatic ring a priori. In this work, taking advantage of recent developments in quantifying nucleophilicity (electrophilicity) with descriptors from the information-theoretic approach in density functional reactivity theory, we examine the reactivity properties of this reaction system from three perspectives. These include scaling patterns of information-theoretic quantities such as Shannon entropy, Fisher information, Ghosh–Berkowitz–Parr entropy and information gain at both molecular and atomic

levels, quantitative predictions of the barrier height with both Hirshfeld charge and information gain, and energetic decomposition analyses of the barrier height for the reactions. To that end, we focused in this work on the identity reaction of the monosubstituted-benzene molecule reacting with hydrogen fluoride using boron trifluoride as the catalyst in the gas phase. We also considered 19 substituting groups, 9 of which are ortho/para directing and the other 9 meta directing, besides the case of R = −H. Similar scaling patterns for these information-theoretic quantities found for stable species elsewhere were disclosed for these reactions systems. We also unveiled novel scaling patterns for information gain at the atomic level. The barrier height of the reactions can reliably be predicted by using both the Hirshfeld charge and information gain at the regioselective carbon atom. The energy decomposition analysis ensued yields an unambiguous picture about the origin of the barrier height, where we showed that it is the electrostatic interaction that plays the dominant role, while the roles played by exchange-correlation and steric effects are minor but indispensable. Results obtained in this work should shed new light for better understanding of the factors governing the reactivity for this class of reactions and assisting ongoing efforts for the design of new and more efficient catalysts for such kind of transformations.
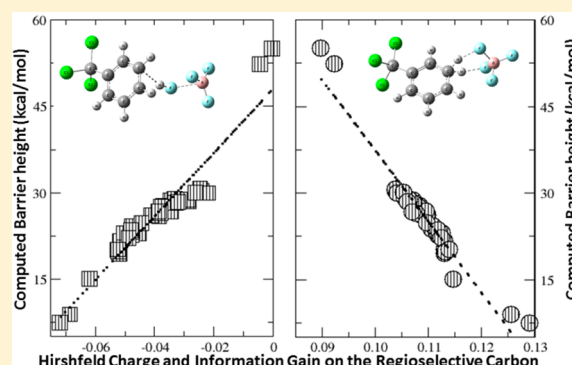
1. INTRODUCTION

To accurately predict the chemical reactivity, that is, regioselectivity, stereoselectivity, barrier heights, rate constants, and so forth, for such chemical reactions as electrophilic aromatic substitution[1−3] is still an unaccomplished task. Recently, we have developed a quantitative approach employing information-theoretic quantities under the framework of density functional reactivity theory[4−7] to predict the regioselectivity and barrier heights.[8−12] According to density functional theory,[4,5] any properties including molecular structure and chemical reactivity in the ground state of a molecular system are determined by its electron density. Information-theoretic quantities such as Shannon entropy,[13] Fisher information,[14] Ghosh–Berkowitz–Parr entropy[15,16] and information gain[17] are functionals of the electron density and its associated

quantities (e.g., density gradient and Laplacian, etc.), so they should be natural descriptors of chemical reactivity. On the other hand, we recently revealed that these information-theoretic quantities follow interesting scaling patterns within the molecular environment with respect to the total electron population at the scales of atoms, molecules, and atoms in molecules.[18] These amazing scaling properties should provide the changing patterns to understand chemical reactivity for atoms and molecules.

In this work, we examine the behavior of these quantities for a category of chemical transformations, electrophilic aromatic

substitution reactions, quantify their reactivity using these descriptors, and reveal their scaling properties with respect to the electron population. An electrophilic aromatic substitution is a process where one atom or group on the aromatic benzene ring is replaced by an incoming electrophile.[1−3] The regioselectivity and reactivity of this reaction is governed by the group, either ortho/para or meta directing, that is already attached to the ring. As we have recently pointed out, we can employ the quantities like information gain and Hirshfeld charge to determine the regioselectivity and quantify the reactivity.[8−12] Here, we try to address three issues related to this kind of reaction systems. At first, we examine the behavior of employing both Hirshfeld charge and information gain as descriptors to predict reaction barrier heights. Second, with the tools we recently developed in quantifying chemical effects, such as steric effect under the framework of density functional reactivity theory,[19−31] we decompose the barrier height and determine the energy components that play dominant and indispensable roles in contributing to the barrier heights of these reactions. Finally, we reveal the scaling properties of information-theoretic quantities in these systems with respect to the electronic population at both molecular and atoms-in-molecules levels. These results should provide useful information in further understanding the factors governing the reactivity for this vastly important category of reactions.

2. THEORETICAL FRAMEWORK

Key quantities in information theory is the Shannon entropy[13]

$$S_S = -\int \rho(\mathbf{r}) \ln \rho(\mathbf{r}) d\mathbf{r} \tag{1}$$

with $\rho(\mathbf{r})$ as the ground state electron density of an N-electron system, which satisfies the normalization condition, $\int \rho(\mathbf{r}) d\mathbf{r} = N$, and the Fisher information defined by[14]

$$I_F = \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} \tag{2}$$

with $\nabla \rho(\mathbf{r})$ being the density gradient. Earlier, we proved that for atoms and molecules there exists an equivalent formula for the Fisher information in terms of the Laplacian of the electron density.[32] The third quantity is called the Ghosh−Berkowitz−Parr (GBP) entropy[15,16]

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

where $t(\mathbf{r};\rho)$ is the kinetic energy density, which is related to the total kinetic energy T_S via $\int t(\mathbf{r};\rho) d\mathbf{r} = T_S$, and $t_{TF}(\mathbf{r};\rho)$ is the Thomas−Fermi kinetic energy density given by $t_{TF}(\mathbf{r};\rho) = c_k \rho^{5/3}(\mathbf{r})$, with k as the Boltzmann constant (set to be unity for convenience in this work), $c = (5/3) + \ln(4\pi c_k/3)$, and $c_k = (3/10)(3\pi^2)^{2/3}$. Following the original literature,[15,16] we use $t(\mathbf{r};\rho) = \sum_i (1/8)((\nabla \rho_i \cdot \nabla \rho_i)/(\rho_i)) - (1/8)\nabla^2 \rho$, where ρ_i are the orbital densities.

These three quantities, which were defined by the electron density $\rho(\mathbf{r})$, can also be redefined 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_σ, I_σ, and S_{GBP}^σ,[18] respectively. With the shape function, one can readily prove that[18]

$$S_\sigma = \frac{S_S}{N} + \ln N \text{ and } I_\sigma = \frac{I_F}{N} \tag{4}$$

For the GBP entropy, there exists no explicit relationship between S_{GBP} and S_{GBP}^σ. These quantities are all functions of the electron density and thus electron population N. Strong linear correlation between these quantities and molecular and atomic electron populations have recently been observed for a list of stable molecular species.[18]

Very recently, we have found out that the 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[8−12]

$$I_G = \sum_A \int \rho_A \ln \frac{\rho_A}{\rho_A^0} dr \tag{5}$$

is a quantitative measure of electrophilicity, nucleophilicity, and regioselectivity. In eq 5, ρ_A is the electron density on Atom (or Group) A in a molecule, whose total electron density is ρ and the total number of electrons is N, and ρ_A^0 is the counterpart of Atom (or Group) A in the reference state, which can be neutral atom, or ion, or group, and so forth. This result originates from the "minimum information gain principle", proved earlier by Nalewajski and Parr, who showed that minimizing eq 5 leads to the "'stockholder partition'" of the electron density proposed by Hirshfeld[33−36]

$$\rho_A = \frac{\rho_A^0}{\sum_A \rho_A^0} \rho \tag{6}$$

If we introduce a new variable, $x = (\rho_A - \rho_A^0)/\rho_A$, and employ $\ln(1/(1-x)) \approx x$ as the first-order approximation, eq 5 can be simplified as[9]

$$I_G \approx \sum_A \int (\rho_A - \rho_A^0) dr = -\sum_A q_A \equiv 0 \tag{7}$$

where q_A is the Hirshfeld charge on Atom (or Group) A. This result shows that under the first-order approximation, the information gain not only gives rise to the Hirshfeld charge distribution but it also simply vanishes. This result has been verified in a number of systems and we called it the "information conservation principle".[9]

Meanwhile, from the energetic perspective, in DFT the conventional energy decomposition scheme to partition the total energy difference ΔE of a chemical reaction is as follows[1,19]

$$\Delta E[\rho] = \Delta T_s[\rho] + \Delta E_e[\rho] + \Delta E_{xc}[\rho] \tag{8}$$

where ΔE is taken between the transition state and the reactant, T_s stands for the noninteracting kinetic energy, E_e denotes the electrostatic energy, and E_{xc} represents the exchange-correlation energy component. The electrostatic term consists of the nuclear-electron attraction, V_{ne}, classical electron−electron Coulombic repulsion, J, and nuclear−nuclear repulsion, V_{nn}, $E_e[\rho] = V_{ne}[\rho] + J[\rho] + V_{nn}$. Alternatively, a new partition scheme was recently proposed[19,31]

$$\Delta E[\rho] = \Delta E_s[\rho] + \Delta E_e[\rho] + \Delta E_q[\rho] \tag{9}$$

where the total energy difference comes from the contribution from three distinct but independent physiochemical effects, steric ΔE_s, electrostatic ΔE_e, and Fermionic quantum due to exchange and correlation effects ΔE_q. The common term in above two schemes is the electrostatic contribution ΔE_e. Notice that the steric energy is simply the Weizsäcker kinetic energy $T_W[\rho]$,[37]

which is closely related to the Fisher information, eq 2[19,32]

$$E_S[\rho] \equiv T_W[\rho] = \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} \equiv \frac{I_F}{8} \qquad (10)$$

Studies for a number of systems have shown that it is the electrostatic term ΔE_e that plays the dominant role in governing the total energy difference, but contributions from other terms, though less significant, are also indispensable.[19−31]

In this work, we will examine the role of these quantities and relationships for a number of electrophilic aromatic substitution reactions. In particular, we are interested to see (i) what the performance of applying Hirshfeld charge and information gain in predicting the reaction barrier height of these reactions is (2) whether the electrostatic component is still the dominant contributor to the barrier height; and (3) if there exist similar strong linear correlations from the information-theoretic quantities for these systems with transition state involved.

3. COMPUTATIONAL DETAILS

In this work, we considered a total of 19 monosubstituted benzene compounds (Scheme 1), 9 of which are derivatives with

Scheme 1. Ortho-, Meta-, and Para- Position Carbon Atoms of the Monosubstituted Benzene with the Functional Group R, Which Could Be Either Ortho/Para- or Meta- Directing
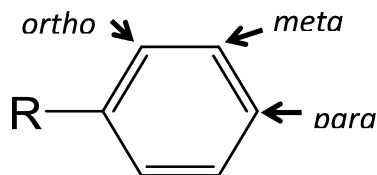
the ortho/para directing groups, R = −F, −Cl, −CH$_3$, −C$_2$H$_5$, −C$_3$H$_7$, −tBu, −NH$_2$, −-NMe$_2$, and −OH, and the other 9 with the meta-directing groups, −CCl$_3$, −CF$_3$, −CHO, −CN, −COF, −NH$_3^+$, −NO$_2$, −NO, and −SO$_3$H. Besides, we also included the case of R = −H (benzene). For each model system studied, no matter if it contains an ortho/para or a meta directing group, we examined its electrophilic substitution at both *meta* and

para positions, meaning that a total of 37 electrophilic aromatic substitution reactions are considered in this work. The reaction we investigated is the identity reaction of the monosubstituted-benzene molecule reacting with hydrogen fluoride using BF$_3$ as the catalyst in the gas phase (Scheme 2). Following the literature,[38,39] we represented the catalyst by one BF$_3$ molecule, and the reactant complex consisting of one monosubstituted benzene ring with one hydrogen fluoride (HF) and one BF$_3$. The QST2 technique was employed to obtain the optimized transition state structure of the reaction. The hybrid B3LYP density functional with the correlation consistent aug-cc-pVDZ basis set was employed for all the calculations.[40−42] After a transition-state search was accomplished, a single-point frequency calculation was performed to ensure that the final structure obtained (i) has only one imaginary frequency and (ii) the vibration mode of the negative frequency corresponds to the anticipated bond formation and breaking. In addition, intrinsic reaction coordinates (IRCs) were calculated to verify the relevance of transition-state structures. All calculations were performed with the Gaussian 09 D01 package with tight SCF convergence and ultrafine integration grids.[43] An illustrative example of the optimized reactant/product and transition state structures is shown in Scheme 3 for the reaction with the meta-directing R = −CCl$_3$. The Multiwfn 3.2 program[44] developed by one of the present coauthors was utilized to calculate the information-theoretic quantities in eqs 1−10 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 ROM062X/6-311+G(d) 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, Bader's zero-flux atoms-in-molecules criterion, and Hirshfeld's stockholder approach, are possible.[18] As have been demonstrated earlier, these three approaches yield similar results. In this work, we choose the Hirshfeld's stockholder approach[45] to partition the atoms in molecules.

Scheme 2. Electrophilic Aromatic Substitution Reaction Studied in This Work with −R as the Monosubstituted Group, Which Can Be Either Otho/Para-Directing Group or meta-Directing Group
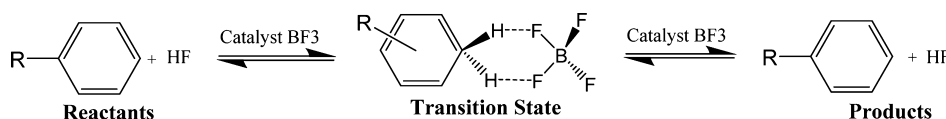
Scheme 3. An Illustrative Example of the Optimized Structure for the Reactant/Product and Transition State of the Electrophilic Aromatic Substitution Reaction with the Meta-Directing Group R = m-CCl$_3$

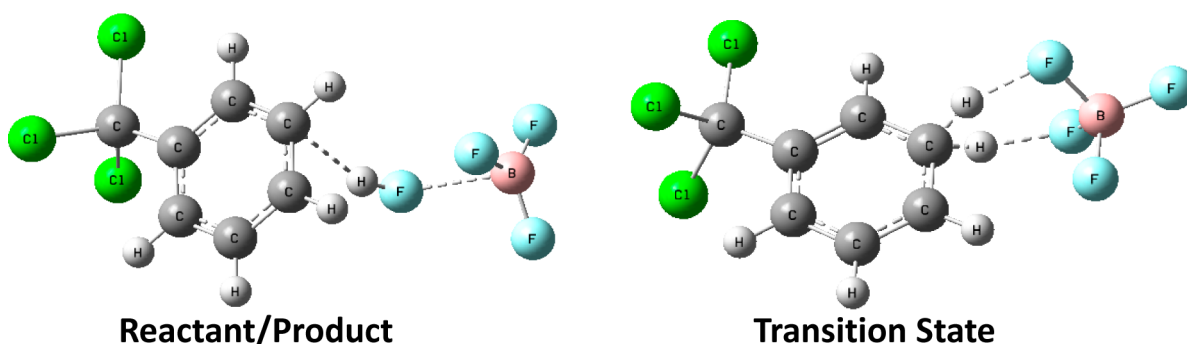**Reactant/Product**   **Transition State**

Table 1. Calculated Information-Theoretic Quantities Including Shannon Entropy, Fisher Information, Ghosh–Berkowitz–Parr Entropy and Information Gain for 37 Reaction Systems Studied in This Work[a]

|		transition state				reactant/product complex			
−R	N	S_S	I_F	I_{GBP}	I_G	S_S	I_F	I_{GBP}	I_G
m-CCl$_3$	140	18.48	10778.09	909.03	0.991	19.43	10781.99	909.51	0.968
m-CF$_3$	116	52.31	5952.28	753.98	0.998	53.25	5956.37	754.48	0.974
m-CHO	98	58.68	4675.59	637.62	1.053	59.63	4679.50	638.11	1.019
m-CN	96	58.21	4562.98	624.79	0.984	59.17	4567.30	625.31	0.969
m-COF	106	55.25	5249.15	689.36	0.997	56.22	5253.42	689.88	0.979
H	84	52.04	3977.50	546.45	0.953	52.98	3981.16	546.91	0.902
m-NH$_3^+$	92	57.02	4320.29	598.28	0.291	58.18	4326.26	599.06	0.315
m-NO$_2$	106	56.42	5216.70	689.50	0.985	57.38	5221.08	690.03	0.977
m-NO	98	55.09	4768.90	637.72	0.976	56.04	4772.94	638.24	0.959
p-CCl$_3$	140	18.46	10777.94	909.01	0.979	19.43	10781.99	909.51	0.968
p-CF$_3$	116	52.31	5952.05	753.96	0.983	53.25	5956.37	754.48	0.974
p-CHO	98	58.70	4675.33	637.61	1.034	59.62	4679.47	638.10	1.020
p-CN	96	58.23	4563.04	624.80	0.975	59.17	4567.30	625.31	0.969
p-COF	106	55.27	5248.81	689.34	0.973	56.21	5253.41	689.88	0.978
p-NH$_3^+$	92	57.08	4320.32	598.29	0.278	58.20	4326.27	599.07	0.318
p-NO$_2$	106	56.44	5216.48	689.48	0.963	57.38	5221.08	690.03	0.976
p-NO	98	55.13	4768.77	637.71	0.950	56.04	4772.96	638.24	0.957
p-SO$_3$H	124	46.56	7257.56	805.63	1.111	47.50	7262.03	806.16	1.112
p-SO$_3$H	124	46.54	7257.76	805.64	1.132	47.51	7262.07	806.17	1.115
p-Cl	100	37.40	6160.10	649.98	0.903	38.37	6163.75	650.46	0.872
p-Et	100	72.41	4482.10	650.52	1.307	73.24	4484.90	650.90	1.230
p-F	92	48.72	4551.82	598.32	0.921	49.64	4555.33	598.80	0.878
m-Cl	100	37.41	6159.72	649.97	0.897	38.37	6163.74	650.46	0.872
p-Me	92	62.21	4229.62	598.47	1.134	63.12	4232.90	598.91	1.063
m-Et	100	72.35	4481.59	650.50	1.290	73.19	4484.74	650.87	1.228
m-F	92	48.70	4551.29	598.32	0.903	49.62	4555.27	598.79	0.876
m-Me	92	62.21	4229.45	598.48	1.118	63.06	4232.68	598.86	1.060
m-NH$_2$	92	57.94	4318.53	598.41	1.036	58.80	4322.03	598.79	0.996
m-NMe$_2$	108	78.26	4821.98	702.48	1.364	79.06	4825.22	702.81	1.316
m−OH	92	53.50	4427.03	598.40	0.982	54.40	4430.70	598.83	0.945
m-Pr	108	82.56	4733.86	702.54	1.456	83.46	4737.24	702.94	1.393
m-tBu	116	92.61	4985.42	754.38	1.607	93.43	4988.48	754.74	1.541
p-NH$_2$	92	57.93	4318.28	598.36	1.021	58.82	4322.10	598.81	0.993
p-NMe$_2$	108	78.26	4822.09	702.42	1.350	79.09	4825.24	702.84	1.316
p−OH	92	53.51	4427.76	598.39	1.014	54.38	4430.63	598.81	0.946
p-Pr	108	82.62	4734.36	702.57	1.473	83.54	4737.50	703.00	1.394
p-tBu	116	92.60	4985.61	754.38	1.625	93.43	4988.48	754.74	1.541

[a]Atomic units.

4. RESULTS AND DISCUSSION

Table 1 presents the numerical results of four information-theoretic quantities, Shannon entropy, Fisher information, Ghosh–Berkowitz–Parr entropy and information gain defined through the electron density, for both the transition state and reactant/product of electrophilic aromatic substitution reactions studied in this work. Through eq 4, one should readily obtain the first two quantities in the representation of the shape function. The two strong linear correlations with respect to the total number of electrons in these systems obtained from the table are shown in Figure 1, where we find that Fisher information is reasonably proportional to the total N (Figure 1a). Shown in Figure 1b is an even stronger relationship, where the GBP entropy is almost perfectly linearly proportional to the total number of electrons. These relationships were seen in other systems, as have previously been reported.[18] But for transition state structures and reaction systems, Figure 1 demonstrates for the first time that these correlations are also applicable and valid.

Shown in Figure 2 are the numerical results of information-theoretic quantities for a total of 524 carbon (Figures 2a−c) and 612 hydrogen (Figures 2d−f) atoms in molecules. Again, we did the atom partition in this study by using Hirshfeld's shareholder scheme. These atomic scaling results 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.[18] The difference is that in this work, transition state structures and chemical reactions are involved. These results confirm that the strong correlations for the information theory quantities are generally applicable to any type of chemical systems. The origin and implication of these relationships, however, remain to be discovered and appreciated.

What is new in Figure 2 is the scaling property of information gain with respect to the total electron population for carbon (Figure 2c) and hydrogen (Figure 2f) atoms. They seemingly follow the same changing pattern. That is, as the total electron population increases, the information gain also increases. Because an increased Hirshfeld population represents an increase
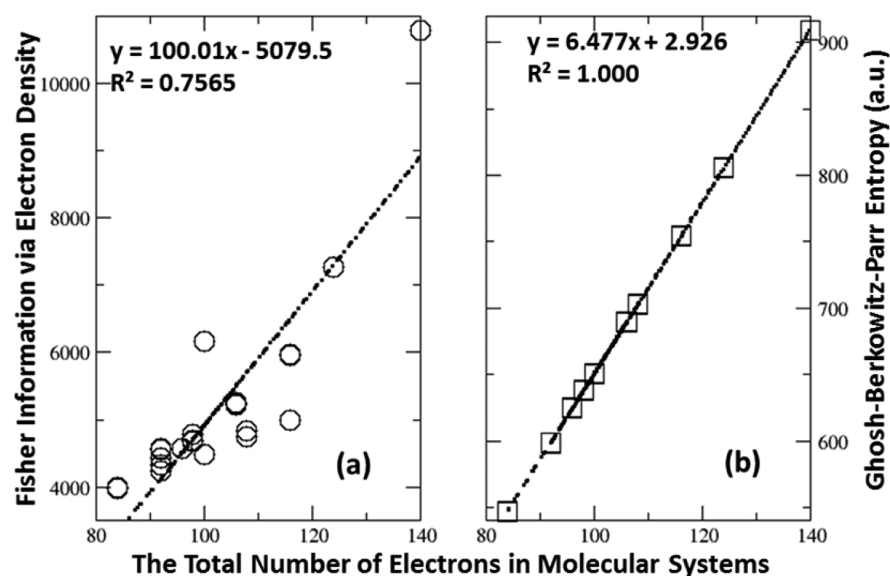
Figure 1. Strong linear correlations observed for the model systems studied in this work. (a) Fisher information as a function of the total number of electrons in the molecules and (b) Ghosh–Berkowitz–Parr entropy as a function of the total number of electrons in the molecules.
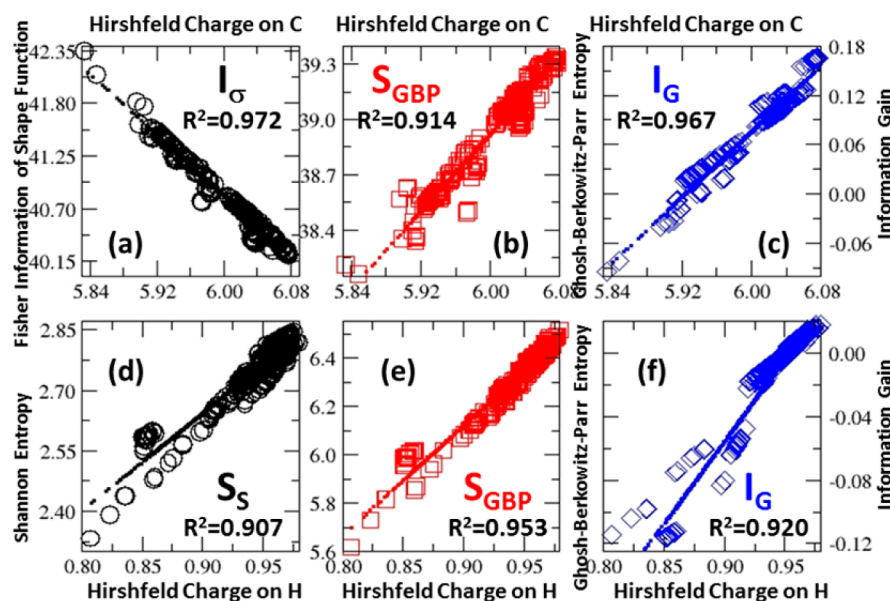

Figure 2. Scaling results for (a–c) carbon and (d–f) hydrogen atoms in molecules with Hirshfeld's stockholder partitioning approach for 524 carbon atoms and 612 hydrogen atoms in 37 molecular systems studied in this work.

capability to donate electrons, that is, an increase in nucleophilicity. From eq 7, as the first-order approximation, we see that the information gain is closely related to the Hirshfeld charge, but that approximation cannot tell how the information gain scales with respect to the total number of electron population.

Can these information-theoretic quantities be used to predict the reactivity for this category of reactions? That is, is it possible to apply these quantities to predict the barrier height of electrophilic substitution reactions? At the molecular level, we have tried to correlate molecular values of Shannon entropy, Fisher information, Ghosh–Berkowitz–Parr entropy and information gain with the calculated barrier height for the reactions studied, but none of them yield any significant correlation (results not shown). Much better results were obtained at the atomic level using the Hirshfeld partition. As previously reported by us,[9] the Hirshfeld charge at the regioselective site is a

quantitatively reliable descriptor of the chemical reactivity. This is again confirmed by Figure 3. Shown in the figure are two strong correlations that we obtained from the regioselective carbon atoms on the aromatic benzene ring. The first is from the Hirshfeld charge (Figure 3a) with the correlation coefficient for the 37 reactions $R^2 = 0.905$. In this case, we computed the Hirshfeld charge on the regioselective carbon atoms in Scheme 1, without considering the impact of the HF molecule and the BF_3 catalyst. The other, an even better one, is from the information gain in the transition state at the regioselective site, with $R^2 = 0.941$ (Figure 3b). In our previous study,[9] results for ortho/para and meta directing groups were separately presented. Shown in Figure 3a is the correlation, where both groups fitted into the same line. These strong linear relationships indicate that the more the negative Hirshfeld charge on the regioselective site, the more the information gain on the atom, the stronger its
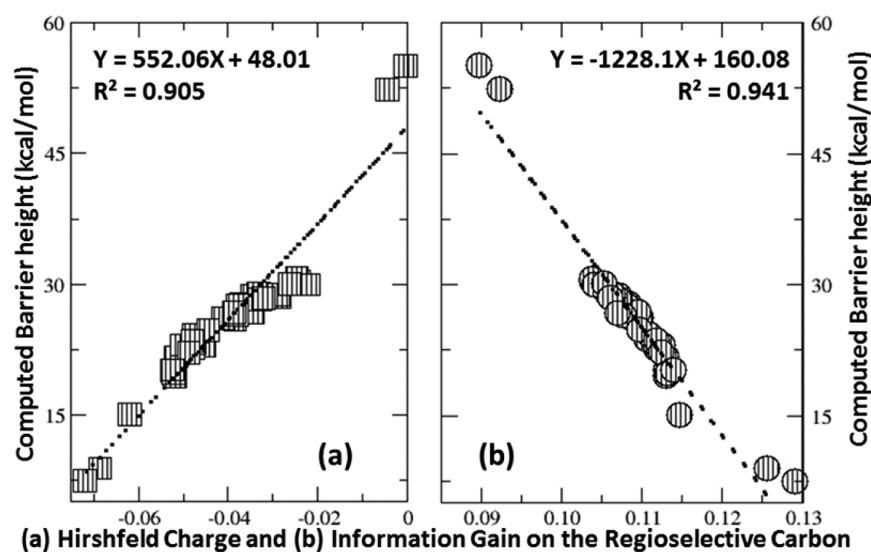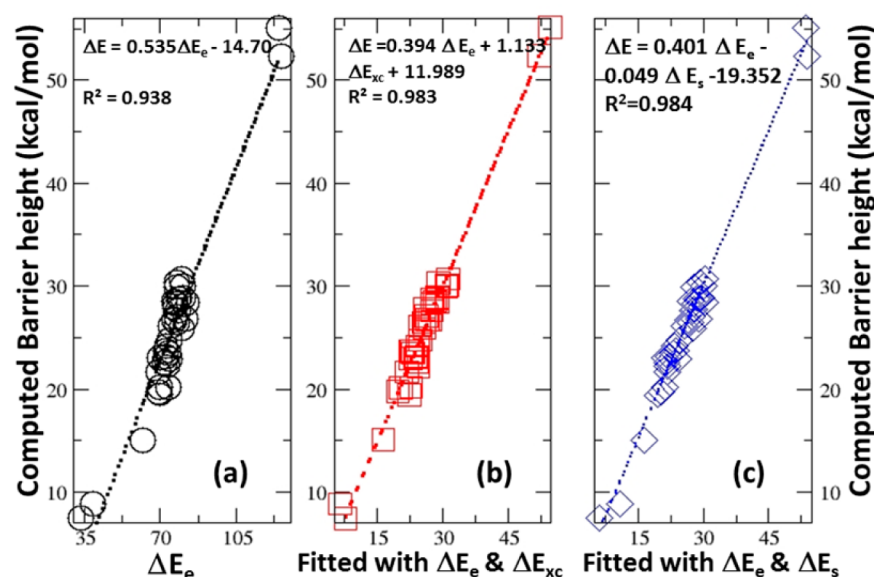
Figure 3. Strong linear correlations for 37 electrophilic aromatic substitution reactions between the computed barrier heights and (a) the Hirshfeld charge on the regioselective carbon atoms in Scheme 1 and (b) the information gain on the regioselective carbon atom in the transition state in Scheme 2.

Table 2. Results of the Energy Decomposition Analysis for the Barrier Height of 37 Reaction Systems Studied in This Work[a]

−R	ΔT_s	ΔE_{xc}	ΔE_e	ΔE_s	ΔE_q	ΔLUMO	ΔHOMO	$.\Delta E$
m-CCl$_3$	−38.38	−15.58	79.84	−305.65	251.69	−56.92	−28.35	25.88
m-CF$_3$	−35.95	−14.29	77.32	−320.37	270.13	−63.50	−31.17	27.09
m-CHO	−33.89	−14.90	74.91	−306.79	258.00	−45.96	−23.54	26.12
m-CN	−36.57	−13.69	78.99	−338.90	288.63	−52.94	−26.81	28.72
m-COF	−34.47	−14.54	76.79	−334.69	285.67	−44.88	−30.43	27.78
H	−33.03	−15.82	72.55	−286.80	237.94	−72.60	−33.60	23.69
m-NH$_3^+$	−64.96	−8.26	125.56	−467.99	394.77	−60.32	−38.08	52.34
m-NO$_2$	−38.09	−13.21	80.22	−343.64	292.34	−33.24	−28.08	28.92
m-NO	−37.16	−13.55	78.35	−316.85	266.14	−30.55	−24.38	27.63
p-CCl$_3$	−39.76	−15.20	81.74	−317.38	262.42	−58.87	−29.12	26.78
p-CF$_3$	−35.48	−13.12	76.90	−338.35	289.74	−63.71	−27.79	28.30
p-CHO	−35.56	−12.32	76.41	−324.57	276.69	−51.86	−27.47	28.53
p-CN	−37.21	−12.57	78.60	−334.20	284.42	−59.47	−33.20	28.81
p-COF	−35.39	−11.97	77.67	−360.24	312.89	−50.76	−24.55	30.31
p-NH$_3^+$	−63.96	−5.76	124.81	−466.98	397.27	−55.08	−33.17	55.10
p-NO$_2$	−38.51	−10.77	79.98	−360.57	311.29	−41.05	−24.80	30.70
p-NO	−37.20	−10.36	77.50	−329.26	281.70	−41.13	−28.76	29.95
p-SO$_3$H	−39.16	−11.36	80.58	−351.28	300.77	−56.81	−24.52	30.06
p-SO$_3$H	−40.41	−14.10	82.93	−338.08	283.57	−56.11	−27.37	28.42
p-Cl	−34.66	−16.50	74.11	−286.40	235.24	−70.70	−36.53	22.95
p-Et	−35.15	−15.19	69.77	−219.69	169.35	−72.75	−35.42	19.43
p-F	−31.97	−15.48	69.10	−274.89	227.45	−69.84	−33.70	21.65
m-Cl	−36.08	−14.85	77.42	−315.72	264.79	−69.70	−28.21	26.48
p-Me	−32.76	−16.81	69.71	−256.92	207.35	−72.72	−34.84	20.14
m-Et	−34.91	−14.49	72.03	−246.70	197.31	−72.40	−34.14	22.63
m-F	−35.74	−13.80	76.27	−311.81	262.28	−71.35	−30.62	26.73
m-Me	−34.69	−14.63	72.49	−253.50	204.18	−72.31	−34.11	23.17
m-NH$_2$	−34.17	−14.51	72.87	−274.72	226.04	−73.97	−29.46	24.19
m-NMe$_2$	−32.82	−14.03	69.87	−253.51	206.66	−74.23	−26.66	23.02
m−OH	−34.96	−14.39	74.11	−287.85	238.50	−73.26	−31.30	24.75
m-Pr	−34.45	−15.79	73.75	−264.60	214.36	−73.58	−33.66	23.51
m-tBu	−36.53	−14.56	73.27	−240.19	189.10	−72.95	−33.13	22.18
p-NH$_2$	−11.52	−18.42	38.81	−299.26	269.32	−38.06	−28.53	8.87
p-NMe$_2$	−10.37	−15.46	33.28	−246.91	221.09	−35.03	−26.63	7.45
p−OH	−29.26	−17.76	62.09	−224.51	177.49	−69.35	−40.28	15.07
p-Pr	−32.13	−17.38	69.27	−245.96	196.45	−72.56	−34.17	19.77
p-tBu	−37.29	−16.09	73.55	−225.04	171.66	−72.19	−35.57	20.17

[a] Also shown are HOMO and LUMO energies difference. Units in kcal/mol.

Figure 4. Strong linear correlations for 19 electrophilic aromatic substitution reactions between the computed barrier height and various energy components: (a) single-component fitting with the electrostatic term, ΔE_e; (b) two-term fitting with both electrostatic ΔE_e and exchange-correlation components ΔE_{xc}; and (c) two-term fitting with both electrostatic ΔE_e and steric energy components ΔE_s.

capability to donate electrons, and the lower the barrier height of the substitution reaction.

To better understand the origin of the barrier height of the reactions, we resort to the energy partition approaches in density functional theory, eqs 8 and 9. Table 2 shows the numerical results for the 37 electrophilic substitution reactions studied in this work. Also shown in the Table are their HOMO and LUMO energy differences among other places.[46] From the table, we can see that the barrier height ΔE is always positive, $\Delta E > 0$, as well as ΔE_e and ΔE_q quantities, suggesting that in the conventional DFT partition, eq 8, the only term that numerically contributes to ΔE positively is the electrostatic component, whereas in the alternative scheme, eq 9, both electrostatic and Fermionic quantum components work together to do the job. Figure 4 exhibits the strong relationships obtained from the data in the table. With one variable fitting, the only energy component strongly correlating with the barrier height is ΔE_e with $R^2 = 0.938$. This is shown in Figure 4a, suggesting that the dominant energy contributor to the barrier height is of the electrostatic nature. For the ΔE_q quantity, even though it is positively contributing, we do not find any meaningful correlation between ΔE and ΔE_q. For the two-variable fitting, we obtained two strong correlations, one with ΔE_e and ΔE_{xc} (Figure 4b), and the other with ΔE_e and ΔE_s (Figure 4c), with R^2 equal to 0.983 and 0.984, respectively. These results are consistent with our previous studies for other systems,[27] confirming that the electrostatic interaction is dominant in determining the barrier height, but contributions from other effects such as exchange-correlation and steric interactions, though minor, are indispensable. Notice that no significant correlation is seen from frontier orbitals, so one cannot employ frontier molecular orbitals to predict the barrier height.[46]

Put together, this work presents a computational study for a total of 19 electrophilic aromatic substitution reactions with both ortho/para and meta-directing groups demonstrating the usefulness and applicability in predicting chemical reactivity with information-theoretic quantities such as Shannon entropy, Fisher information, Ghosh–Berkowitz–Parr entropy, and information

gain. Scaling properties have been confirmed, and new linear correlations have been observed. The barrier height can be reliably predicted by either Hirshfeld charge or information gain on the regioselective atom. In addition, our energy decomposition analysis results showed that the electrostatic interaction is the dominant contributor to the reaction barrier height, whereas other effects such as exchange-correlation and steric interactions play minor but indispensable roles.

5. CONCLUSIONS

As one of the most fundamental transformation processes, the reaction of electrophilic aromatic substitutions involved in nitration, halogenation, sulfonation, Friedel–Crafts, and other reactions replaces an atom or a group on the aromatic system by an incoming electrophile. Its reactivity and regioselectivity are predominantly determined by nucleophilicity of carbon atoms on the aromatic ring, which in return is immensely influenced by the group that is already attached to the aromatic ring before the reaction takes place. In this work, we apply quantities from information theory, such as Shannon entropy, Fisher information, Ghosh–Berkowitz–Parr entropy, and information gain, to better understand scaling patterns and reactivity properties of this vastly important category of reactions. To that end, we chose 18 groups with both ortho/para- and meta- directing capabilities and examined the identity reaction of the monosubstituted-benzene molecule reacting with hydrogen fluoride using BF_3 as the catalyst in the gas phase. We found similar scaling properties for information-theoretic quantities in these reactive systems that were observed elsewhere in stable systems. We also unveiled novel scaling patterns for information gain at the atomic level. The barrier height of the reactions can reliably be predicted by using both the Hirshfeld charge and information gain at the regioselective carbon atom. The energy decomposition analysis subsequently ensued yields an unambiguous picture about the origin of the barrier height, for which we found that it is the electrostatic interaction that plays the dominant role, while the roles played by exchange-correlation and steric effects are minor but indispensable. These new insights obtained from this study

about the reactivity properties of different electrophilic aromatic substitution reactions not only showcase the usefulness and applicability of these information-theoretic quantities in understanding chemical reactivity of this important category of chemical transformations but also provide a practical pathway in designing novel and better catalysts whose reaction mechanisms involve electrophilic aromatic substitutions.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: chunyingrong@gmail.com (C.Y.R.).
*E-mail: 1125218638@qq.com (Y.H.).
*E-mail: shubin@email.unc.edu (S.B.L.). Telephone: (919) 962-4032.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Solomons, T. W. G.; Fryhle, C. B.; Snydeer, S. A. *Organic Chemistry*; Wiley: New York, 2013.
(2) Carey, F. A.; Sundberg, R. J. *Advanced Organic Chemistry Part A: Structure and Mechanism*; Kluwer/ Plenum: New York, 2000.
(3) Klein, D. R. *Organic Chemistry*; Wiley: New York, 2012.
(4) Parr, R. G.; Yang, W. T. *Density Functional Theory for Atoms and Molecules*; Oxford University: London, 1989.
(5) Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual density functional theory. *Chem. Rev.* **2003**, *103*, 1793−1873.
(6) Chattaraj, P. K.; Sarkar, U.; Roy, D. R. Electrophilicity index. *Chem. Rev.* **2006**, *106*, 2065−2091.
(7) Liu, S. B. Conceptual density functional theory and some recent developments. *Acta Phys.-Chim. Sin.* **2009**, *25*, 590−600.
(8) Liu, S. B. Where does the electron go? The nature of *ortho/para* and *meta* group directing in electrophilic aromatic substitution. *J. Chem. Phys.* **2014**, *140*, 194109.
(9) Liu, S. B.; Rong, C. Y.; Lu, T. Information conservation principle determines electrophilicity, nucleophilicity, and regioselectivity. *J. Phys. Chem. A* **2014**, *118*, 3698−3704.
(10) Zhou, X. Y.; Rong, C. Y.; Lu, T.; Liu, S. B. Hirshfeld charge as a quantitative measure of electrophilicity and nucleophilicity: Nitrogen-containing systems. *Acta Phys.-Chim. Sin.* **2014**, *30*, 2055−2062.
(11) Rong, C. Y.; Lu, T.; Liu, S. B. Dissecting molecular descriptors into atomic contributions in density functional reactivity theory. *J. Chem. Phys.* **2014**, *140*, 024109.
(12) Liu, S. B. Quantifying reactivity for electrophilic aromatic substitution reactions with Hirshfeld charge. *J. Phys. Chem. A* **2015**, *119*, 3107−3111.
(13) Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 623−656.
(14) Fisher, R. A. Theory of statistical estimation. *Math. Proc. Cambridge Philos. Soc.* **1925**, *22*, 700−725.
(15) Ghosh, S. K.; Berkowitz, M.; Parr, R. G. Transcription of ground-state density-functional theory into a local thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 8028−8031.
(16) Rong, C. Y.; Lu, T.; Chattaraj, P. K.; Liu, S. B. On the relationship among Ghosh−Berkowitz−Parr entropy, Shannon entropy and Fisher information. *Indian J. Chem. A* **2014**, *53*, 970−977.
(17) Kullback, S. *Information Theory and Statistics*; Dover: Mineola, NY, 1997.

(18) Rong, C. Y.; Lu, T.; Ayers, P. W.; Chattaraj, P. K.; Liu, S. B. Scaling properties of information-theoretic quantities in density functional reactivity theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 4977−4988; Corrections. *Phys. Chem. Chem. Phys.* **2015**, *17*, 11110−11111.
(19) Liu, S. B. Steric effect: A quantitative description from density functional theory. *J. Chem. Phys.* **2007**, *126*, 244103.
(20) Liu, S. B.; Govind, N. Toward understanding the nature of internal rotation barriers with a new energy partition scheme: ethane and *n*-butane. *J. Phys. Chem. A* **2008**, *112*, 6690−6699.
(21) Liu, S. B.; Govind, N.; Pedersen, L. G. Exploring the origin of the internal rotational barrier for molecules with one rotatable dihedral angle. *J. Chem. Phys.* **2008**, *129*, 094104.
(22) Torrent-Sucarrat, M.; Liu, S. B.; De Proft, F. Steric effect: Partitioning in atomic and functional group contributions. *J. Phys. Chem. A* **2009**, *113*, 3698−3702.
(23) Liu, S. B.; Hu, H.; Pedersen, L. G. Steric, quantum, and electrostatic effects on SN_2 reaction barriers in gas phase. *J. Phys. Chem. A* **2010**, *114*, 5913−5918.
(24) Ess, D. H.; Liu, S. B.; De Proft, F. Density functional steric analysis of linear and branched alkanes. *J. Phys. Chem. A* **2010**, *114*, 12952−12957.
(25) Tsirelson, V. G.; Stash, A. I.; Liu, S. B. Quantifying steric effect with experimental electron density. *J. Chem. Phys.* **2010**, *133*, 114110.
(26) Huang, Y.; Zhong, A. G.; Yang, Q.; Liu, S. B. Origin of anomeric effect: A density functional steric analysis. *J. Chem. Phys.* **2011**, *134*, 084103.
(27) Liu, S. B. Origin and nature of bond rotation barriers: A unified view. *J. Phys. Chem. A* **2013**, *117*, 962−965.
(28) Zhao, D. B.; Rong, C. Y.; Jenkins, S.; Kirk, S. R.; Yin, D. L.; Liu, S. B. Origin of the *cis*-effect: A density functional theory study of doubly substituted ethylenes. *Acta Phys.-Chim. Sin.* **2013**, *29*, 43−54.
(29) Tsirelson, V. G.; Stash, A. I.; Karasiev, V. V.; Liu, S. B. Pauli potential and Pauli charge from experimental electron density. *Comput. Theor. Chem.* **2013**, *1006*, 92−99.
(30) Wang, Y. J.; Zhao, D. B.; Rong, C. Y.; Liu, S. B. Towards understanding the origin and nature of the conformational stability of water clusters: A density functional theory and quantum molecular dynamics study. *Acta Phys.-Chim. Sin.* **2014**, *29*, 2173−2179.
(31) Fang, D.; Piquemal, J.-P.; Liu, S. B.; Cisneros, G. A. DFT-steric-based energy decomposition analysis of intermolecular interactions. *Theor. Chem. Acc.* **2014**, *133*, 1484.
(32) Liu, S. B. On the relationship between densities of Shannon entropy and Fisher information for atoms and molecules. *J. Chem. Phys.* **2007**, *126*, 191107.
(33) Nalewajski, R. F.; Parr, R. G. Information theory, atoms in molecules, and molecular similarity. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8879−8882.
(34) Nalewajski, R. F.; Parr, R. G. Information theory thermodynamics of molecules and their Hirshfeld fragments. *J. Phys. Chem. A* **2001**, *105*, 7391−7400.
(35) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. What is an atom in a molecule? *J. Phys. Chem. A* **2005**, *109*, 3957−3959.
(36) Ayers, P. W. Information theory, the shape function, and the Hirshfeld atom. *Theor. Chem. Acc.* **2006**, *115*, 370−378.
(37) von Weizsäcker, C. F. Zur theorie der kernmassen. *Eur. Phys. J. A* **1935**, *96*, 431−458.
(38) Alagona, G.; Scrocco, E.; Silla, E.; Tomasi, J. The catalytic effect of BF_3 on the electrophilic hydrogen exchange reaction in benzene. *Theoret. Chim. Acta* **1977**, *45*, 127−136.
(39) Heidrich, D. The transition state of electrophilic aromatic substitution in the gas phase. *Phys. Chem. Chem. Phys.* **1999**, *1*, 2209−2211.
(40) Becke, A. D. A new mixing of Hartree−Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372−1378.
(41) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.

(42) Dunning, T. H., Jr. Rates of convergence and error estimation formulas for the Rayleigh–Ritz variational method. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, P. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. et al. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.

(44) Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592.

(45) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acc.* **1977**, *44*, 129–138.

(46) Liu, S. B.; Schauer, C. K. Origin of molecular conformational stability: Perspectives from molecular orbital interactions and density functional reactivity theory. *J. Chem. Phys.* **2015**, *142*, 054107.