JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

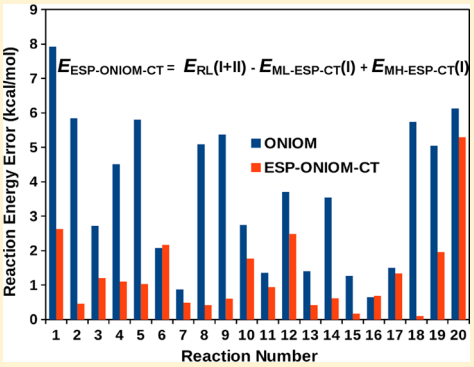
Electrostatic Potential-Based Method of Balancing Charge Transfer Across ONIOM QM:QM Boundaries

K. V. Jovan Jose and Krishnan Raghavachari*

Department of Chemistry, Indiana University, Boomington, Indiana 47405, United States

Supporting Information

ABSTRACT: The inability to describe charge redistribution effects between different regions in a large molecule can be a source of error in an ONIOM hybrid calculation. We propose a new and an inexpensive method for describing such charge-transfer effects and for improving reaction energies obtained with the ONIOM method. Our method is based on matching the electrostatic potential (ESP) between the model system and the real system. The ESP difference arising due to charge redistribution is overcome by placing an optimum one electron potential at a defined buffer region. In our current implementation, the link atom nuclear charge is optimized iteratively to produce a model low ESP distribution equal to that in the real low calculation. These optimum charges are relatively small in magnitude and corroborate physical intuition. This new ESP-ONIOM-CT method is independent of any arbitrary definition of charges, is defined on the basis of a physical observable, and is less basis set dependent than previous approaches. The method is easily extended for studying reactions involving multiple link atoms. We present a thorough benchmark of this method on test sets consisting of one- and two-link atom reactions. Using reaction energies of four different test sets each with four different combinations of high:low levels of theory, the accuracy of ESP-ONIOM-CT improved by 40−60% over the ONIOM method.



1. INTRODUCTION

Rapid scaling of accurate *ab initio* computational methods with system size is the barrier for their applications to large molecules.[1−6] Approximate methods often fail to give quantitatively useful results. A popular approach to balance the accuracy and scaling of computational methods is to use hybrid energy schemes,[7−16] which treat the chemically important region of the large molecule (region I) at a high level of theory and the rest of the molecule (region II, environment) at a low level of theory. Perhaps the most widely used QM/QM embedding scheme is the ONIOM approach of Morokuma and co-workers,[17−23] where the extrapolated total energy expression is

$$E^{\text{total}} \cong E^{\text{low}}(\text{I} + \text{II}) - E^{\text{low}}(\text{I}) + E^{\text{high}}(\text{I}) \quad (1)$$

where high and low refer to the levels of theory. Note that a separate calculation of region II is not performed. Only calculations on the entire system (I+II, low) and region I (high, low) are needed. In the standard notation used in ONIOM, this is equivalently written as

$$E^{\text{total}} \cong E_{\text{RL}} - E_{\text{ML}} + E_{\text{MH}} \quad (2)$$

where R stands for the "real system" (regions I + II), M stands for the "model system" (region I), and H and L stand for the high and low levels of theory. Thus, it requires three separate well-defined calculations. The nature of the hybrid-energy expression allows one to couple different computational chemistry methods without modification. ONIOM studies are thus possible with any combination of methods for which energies and gradients are available.

In covalently bonded systems, the model system is constructed from region I by saturating the "dangling bonds" with the addition of link-atoms (typically H atoms). This truncation of region I with a capping H atom is a possible source of error in the ONIOM method. In addition, the electronic coupling between the regions via Coulomb, exchange, charge-transfer, etc. is accounted only at the low level of theory through the E_{RL} subcalculation. This description can be deficient when region II has highly electron withdrawing/donating functional groups. Hence, the electronic nature of region II may shift the electron density associated with region I either in positive or negative directions. The resulting charge imbalance can impose limitations on the performance of the ONIOM method.

There have been many previous approaches to treat this charge imbalance between regions in different hybrid schemes. Merz and co-workers[24] have developed an approach to match chemical potentials of different regions[25,26] by transferring charge from one region to another. Lin and co-workers have also addressed this problem in the context of QM/MM methods.[27−29] In their approach, the fractional number of electrons is obtained by taking an ensemble average of integer charge states, whose weights are determined by the equalization of chemical potentials. The Generalized Hybrid Orbital (GHO) method of Gao and co-workers[30−32] is another hybrid method which also can describe charge transfer. Other workers have also treated the problem of cross-system charge redistribution, though in other contexts.[33−35]

Received: June 20, 2014
Published: August 27, 2014

   dx.doi.org/10.1021/ct5005356 | J. Chem. Theory Comput. 2014, 10, 4351−4359

In our previous work, we developed the ONIOM-CT method[36]—ONIOM including charge transfer—that has the ability to describe such charge redistribution between the regions with only a slight increase in computational effort. This is accomplished via a novel "Model System Preparation" step in which a one-electron potential is optimized to shift density into or out of a defined buffer region. In our initial implementation, we treated the link atoms on the model system as an electron buffer region and altered the link-atom nuclear charges to shift electron density into or out of the region to achieve charge balance.

The model system preparation step in ONIOM-CT was based on atomic charges from different population analyses such as Löwdin, Mulliken, and Hirshfeld schemes. Though our method improved the ONIOM reaction energies, it is based on atomic charges and not on a physical observable. Atomic charges depend strongly on the population analysis used and on the basis set employed: usually the reaction energies do not improve with large basis functions at the low level. The tuned link-atom nuclear charges were almost always greater than one, irrespective of the electronic nature of region II functional groups. Moreover, extension of this method to multiple link atoms is not straightforward. If the same charges are used for all link atoms, these charges do not agree with physical intuition. We have attempted to define a regional fitness function for each link atom and optimized the corresponding link atom nuclear charges. Though the magnitudes of the optimum link atom nuclear charges are different, some of these charges tend to be large and overpolarize the model system. The development of better performing methods is desirable.

The aim of this work is to develop an alternative method that is independent of the arbitrary definition of charges and to improve the ONIOM method with physically intuitive link atom charges. Hence, we have come up with a new method based on the electrostatic potential (ESP) to balance the charge transfer at the model system boundaries. The ESP is a physical observable and hence independent of any arbitrary definition of the charges. The following section furnishes a brief description of the ESP-based ONIOM charge transfer method.

2. METHODS

This section describes the theory behind the electrostatic potential-based ONIOM charge transfer (ESP-ONIOM-CT) method. As shown in Figure 1, the real system consists of regions I and II, and the model system is constructed out of the chemically active region I by saturating the dangling valency with a hydrogen atom. It is assumed that the hydrogen link atom saturates the unsatisfied valency of the supporting atom in region I and replaces region II. However, the electrostatic potential (ESP) experienced by region I, that occurs in both real low (RL) and model low (ML) calculations, may not be the same. Schematically, Figure 1 illustrates this ESP difference in the third column with color coding for the positive and negative ESP.

The electrostatic potential, $V(\mathbf{r})$, is a scalar property, and classically it is defined at a given point as the energy required to bring a unit positive charge from infinity to this point.[37−39] Hence, the ESP at a point $\mathbf{r}$, due to the discrete set of N nuclei with charges Z_A and continuous electron density, $\rho(\mathbf{r})$ is given as

$$V(\mathbf{r}) = \sum_{A=1}^{N} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r}')\, \mathrm{d}^3 r'}{|\mathbf{r} - \mathbf{r}'|} \tag{3}$$

The two terms in this expression are the bare nuclear potential and electronic contributions, respectively. The ESP is sensitive to the push−pull effects of electron density within the bonding region.[40] Moreover, Politzer and Parr[41] have demonstrated a relation connecting the total energy of the molecule and the nuclear point electrostatic potential, $V(\mathbf{r}_{0,m})$, of the molecule.[37] While constructing the model system out of a real system in the ONIOM method, the electronic rearrangements in the former system are noticeable through the ESP changes at the corresponding nuclear points and in the bonding region between the supporting and the host atoms.

In the ONIOM method, employing the correct energies for the model low and model high are very important for attaining a correct extrapolated local property of interest. Hence in the ESP-ONIOM-CT model, we minimize the model system electronic rearrangements due to truncation effects at the boundaries. In other words, we try to improve the model system energy representation by performing a link atom nuclear charge optimization by fitting the ESP distribution of the real and model systems, at a low level of theory. Moreover, ESP is a physical observable and less sensitive to basis set effects,[37,42] thus we feel that the ESP observable is a good descriptor.

Figure 2 describes the algorithm involved in the ESP-ONIOM-CT method. Since this method is an improved ONIOM method, it follows a similar architecture. The algorithm begins with the evaluation of the RL single point energy and the corresponding ESP. For defining the regional ESP, we have taken into account all the nuclear positions in the model system, $V(\mathbf{r}_{0,m})$ (except the link atom), along with a set of points along the vector connecting the supporting atom and the link atom, $V(\mathbf{r}_l)$, up to half the distance to the link atom.

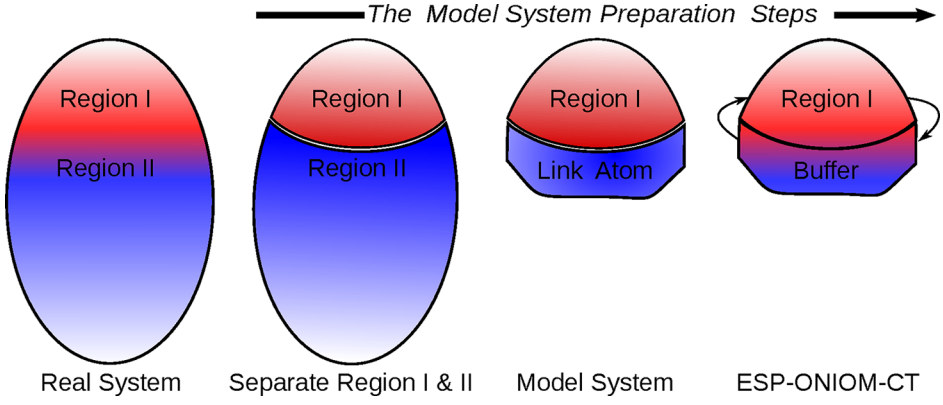


Figure 1. Schematic illustration of the steps involved in the ESP-ONIOM-CT model system preparation at a low level of theory. The red and blue colors correspond to the positive and negative electrostatic potentials, respectively. Refer to the text for further details.

The second step is the model system preparation step. In this step, we improve the ONIOM model by constraining the regional ESP in the model system to be the same as that in the real system. The simplest way to balance the ESP difference between ML and RL is by placing an extra point charge at the buffer region in the model calculation. In the current implementation, the buffer region is considered to be only the link atom, and this extra point charge is placed on top of this link atom (effectively changing the nuclear charge of the link atom). For performing the link-atom nuclear charge optimization, a corresponding fitness function is defined for each link-atom as

$$f_I(V(\mathbf{r}_{0,1}), ..., V(\mathbf{r}_{0,m}), V(\mathbf{r}_i)) = \sum_{i \in I} [V^{RL}(\mathbf{r}_i) - V^{ML}(\mathbf{r}_i)]^2 \quad (4)$$

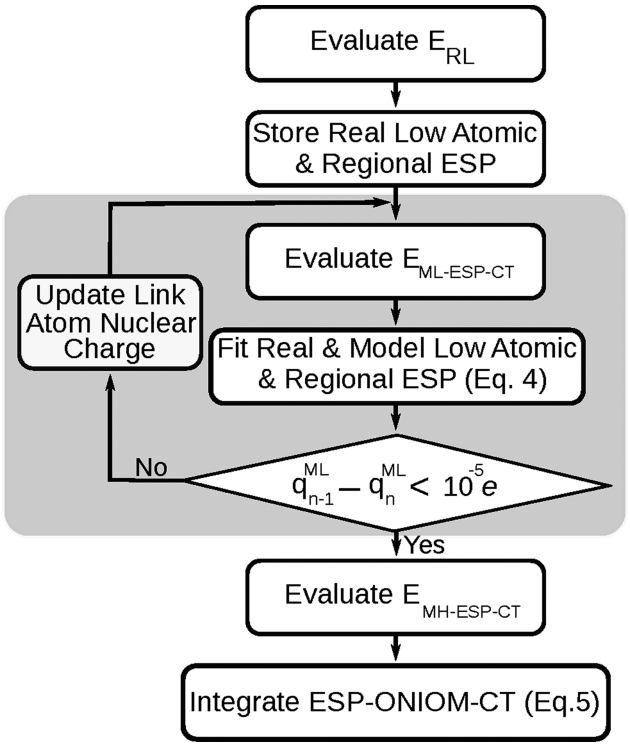


Figure 2. Algorithm describing the steps involved in the ESP-ONIOM-CT energy evaluation. The model system preparation steps within the ESP-ONIOM-CT method are depicted within the gray box. These later steps are precisely illustrated in Figures 1 and 3 for one- and two-link atom test cases, respectively. Refer to the text for further details.

Then the link atom nuclear charge is optimized through the Gauss−Newton nonlinear least-squares algorithm. This step is illustrated in the fourth column of Figure 1. This iterative optimization of the link-atom nuclear charges in practice requires only around 3−5 ML subcalculations to obtain convergence to $10^{-5}e$. Due to the relatively low cost of the ML calculation, this can be done without a significant increase in the overall computational cost. Once the optimal link-atom nuclear charges are obtained for the ML-ESP-CT subcalculation, the MH-ESP-CT subcalculation is then performed using the same link-atom nuclear charges with no further buffer region optimization. The ESP-ONIOM-CT energy expression is thus given as

$$E_{\text{ESP-ONIOM-CT}} = E_{\text{RL}}(\text{I} + \text{II}) - E_{\text{ML-ESP-CT}}(\text{I}) + E_{\text{MH-ESP-CT}}(\text{I}) \quad (5)$$

where $E_{\text{ML-ESP-CT}}$ and $E_{\text{MH-ESP-CT}}$ denote the Model-ESP-CT energies with the low and high levels of theory, respectively. Note that the only difference between eq 2 and eq 5 is that the model system has link atoms with modified nuclear charges. The individual steps of the ESP-ONIOM-CT calculation are shown in Figure 2. Note that it is only the difference in the electronic polarization between MH and ML subcalculations that contributes to the ESP-ONIOM-CT correction. The nuclear term is not needed since it cancels exactly.

Another source of error in the ONIOM approach is the mismatch of the basis functions in the high and low levels of theory. In the standard ONIOM method, the capping hydrogen link atom is not a part of the real system. Since the model high and model low calculations are performed at different levels of theory and basis sets, the link atom effects from these two calculations do not cancel exactly. This difference may become more prominent once the link atom nuclear charge is tuned. We have taken this into account by employing the same set of basis functions for the link atom at model high and model low calculations. Thus, in the ESP-ONIOM-CT model, we treated the link atom at the same low level basis set.

Table 1. List of the Combinations of High and Low Levels of Theory Employed Throughout This Paper

no.	high level	low level
1a	MP2/6-311+G(d,p)	HF/3-21G
1b	MP2/6-31+G(d)	HF/3-21G
2a	B3LYP/6-311+G(d,p)	HF/3-21G
2b	B3LYP/6-31+G(d)	HF/3-21G

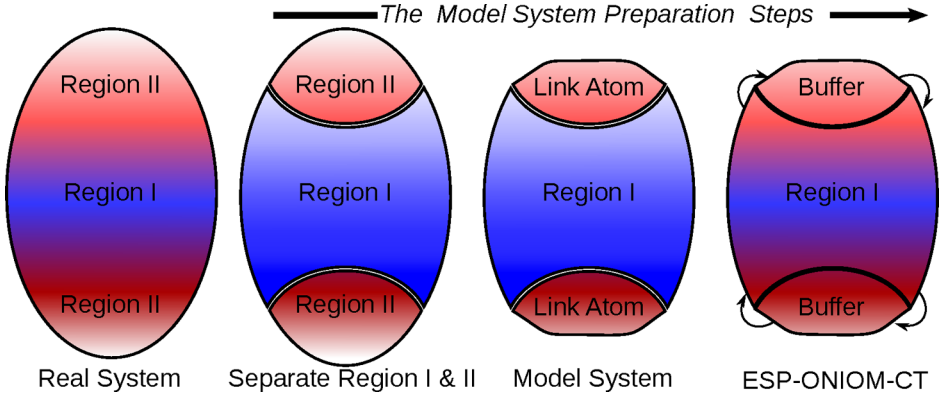


Figure 3. Schematic illustration of the steps involved in the ESP-ONIOM-CT model system preparation involving multiple-link atoms at a low level of theory. The red and blue colors correspond to the positive and negative electrostatic potentials, respectively. Refer to the text for further details.

dx.doi.org/10.1021/ct5005356 | J. Chem. Theory Comput. 2014, 10, 4351−4359

Finally, our method in the current form can be employed to explore reactions involving multiple link atoms. Figure 3 schematically illustrates the general way to extrapolate the ESP-based model for multiple link atoms. The middle region in each column corresponds to region I; the top and bottom parts correspond to region II. The dangling bonds in region I are saturated with two link hydrogen atoms. For optimizing the link atom nuclear charges, a separate fitness function is defined, as in section 2, for each link atom. The Gauss−Newton nonlinear least-squares optimization is carried out to constrain the model system ESP to be the same as that in the real system. Here the optimum charges are physically intuitive, and the resulting reaction energies are closer to the actual high values. This algorithm is currently implemented as a set of *perl* scripts usable with Gaussian 09,[43] to facilitate the computation of the ESP-ONIOM-CT energy. The following section furnishes the benchmark studies of the performance of our method on four test sets of reactions.

Table 2. One-Link Atom Reaction Set Employed for Benchmarking the ONIOM, ONIOM−CT, and ESP−ONIOM−CT Method[a]

	protonation		
1)	X_3C-//-$CH_2(OH)$ + H^+	→	X_3C-//-$CH_2(OH_2^+)$
2)	X_3C-//-$CH_2(O^-)$ + H^+	→	X_3C-//-$CH_2(OH)$
3)	X_3C-//-$CH_2(NH_2)$ + H^+	→	X_3C-//-$CH_2(NH_3^+)$
4)	X_3C-//-$CH_2(NH^-)$ + H^+	→	X_3C-//-$CH_2(NH_2)$
5)	X_3C-//-(COO^-) + H^+	→	X_3C-//-(COOH)
	H-abstraction		
6)	X_3C-//-$CH_2(OH)$	→	X_3C-//-$CH_2(O^\bullet)$ + $H^\bullet$
7)	X_3C-//-$CH_2(NH_2)$	→	X_3C-//-$CH_2(NH^\bullet)$ + $H^\bullet$
	reduction		
8)	X_3C-//-$CH_2(O^\bullet)$ + e^-	→	X_3C-//-$CH_2(OH^-)$
9)	X_3C-//-$CH_2(NH^\bullet)$ + e^-	→	X_3C-//-$CH_2(NH^-)$
	S_N2 reactions		
10)	X_3C-//-CH_2F + Cl^-	→	X_3C-//-CH_2Cl + F^-

[a]In each system, the corresponding model system is depicted on the right side of the double slash. All the benchmark results are reported in Table 4 and refer to the text for further details.

Table 3. Two-Link Atom Reaction Set Employed for Benchmarking ONIOM, ONIOM-CT, and ESP-ONIOM-CT Methods[a]

	protonation		
1)	X_3C-//-CH(OH)-//-CY_3 + H^+	→	X_3C-//-$CH(OH_2^+)$-//-CY_3
2)	X_3C-//-$CH(O^-)$-//-CY_3 + H^+	→	X_3C-//-CH(OH)-//-CY_3
3)	X_3C-//-$CH(NH_2)$-//-CY_3 + H^+	→	X_3C-//-$CH(NH_3^+)$-//-CY_3
4)	X_3C-//-$CH(NH^-)$-//-CY_3 + H^+	→	X_3C-//-$CH(NH_2)$-//-CY_3
5)	X_3C-//-$CH(COO^-)$-//-CY_3 + H^+	→	X_3C-//-CH(COOH)-//-CY_3
	H-abstraction		
6)	X_3C-//-CH(OH)-//-CY_3	→	X_3C-//-$CH(O^\bullet)$-//-CY_3 + $H^\bullet$
7)	X_3C-//-$CH(NH_2)$-//-CY_3	→	X_3C-//-$CH(NH^\bullet)$-//-CY_3 + $H^\bullet$
	reduction		
8)	X_3C-//-$CH(O^\bullet)$-//-CY_3 + e^-	→	X_3C-//-$CH(O^-)$-//-CY_3
9)	X_3C-//-$CH(NH^\bullet)$-//-CY_3 + e^-	→	X_3C-//-$CH(NH^-)$-//-CY_3
	S_N2 reactions		
10)	X_3C-//-CHF-//-CY_3 +Cl^-	→	X_3C-//-CHCl-//-CY_3 +F^-

[a]In each system, the corresponding model system is depicted in between the double slashes. All the benchmark results are reported in Table 6 and refer to the text for further details.

3. RESULTS AND DISCUSSION

For evaluating the performance of the ESP-ONIOM-CT method relative to the ONIOM and ONIOM-CT methods, we have taken representative test sets of chemical reactions involving one- and two-link atoms. The reaction sets listed in Tables 2 and 3 are chosen such that region I is the chemically active region, and region II has strong electron withdrawing/donating groups (X = F and CH_3, i.e., with CF_3 and *t*-butyl groups). The different high:low combinations, listed in Table 1, are chosen to highlight some of the effects of choosing methods with varying disparities both in electron correlation (MP2:HF and B3LYP:HF) and in basis set size (6-311+G(*d*,*p*):3-21G and 6-31+G(*d*):3-21G). To perform a thorough benchmark comparison with the earlier defined ONIOM-CT model and to avoid the known problems with performing Mulliken and Löwdin population analysis with large basis sets,[44] we have employed only the 3-21G basis set for the low level of theory in this study. The following sections furnish evaluation of the reaction energies on benchmark systems with one and two link atoms.

ESP distributions for neutral, anionic, and cationic species are quite different from each other. Further, ESP is very sensitive to the changes in the electronic structure of a molecule and is not transferable from the reactant to the product (particularly for reactions involving changes in the charge state as in many of the reactions we have considered). Thus, we cannot place the same link atom charge in both reactant and product to account for the regional charge transfer in each of these systems. Hence, in all cases, we have optimized the link atom charges separately for each reactant and product.

3.1. Reaction Energies of One-Link Atom Test Sets.

Table 2 reports the one-link atom reaction test set employed for benchmarking the ESP-ONIOM-CT method. Each quantity reported (mean absolute deviation (MAD), maximum deviation) is listed for each high:low combination with the averaged value given at the bottom. All the ONIOM-CT results listed for comparison were obtained via a Mulliken population analysis. Throughout this paper, a deviation is defined as the difference

Table 4. Comparison of the Performance of the ONIOM, ONIOM-CT, and ESP-ONIOM-CT Methods on One-Link Atom Reaction Set, Reported in Table 2[a]

	MAD (max deviation)			MAD (X = F)			MAD (X = CH_3)		
high:low	ONIOM	ONIOM-CT	EONIOM-CT	ONIOM	ONIOM-CT	EONIOM-CT	ONIOM	ONIOM-CT	EONIOM-CT
1a	3.59 (7.99)	1.82 (5.82)	1.59 (3.46)	4.16	1.67	1.40	3.03	1.97	1.80
1b	3.47 (7.90)	1.92 (4.78)	1.57 (3.39)	3.30	1.85	1.16	3.64	2.06	1.98
2a	4.13 (9.11)	2.26 (4.45)	1.97 (3.35)	3.82	2.26	1.67	4.43	2.80	2.27
2b	4.20 (9.29)	1.98 (7.38)	1.87 (4.86)	3.92	2.10	1.88	4.48	1.92	1.86
average	**3.85 (8.57)**	**2.00 (5.61)**	**1.75 (3.77)**	**3.80**	**1.97**	**1.53**	**3.90**	**2.19**	**1.98**

[a]The three sets of columns correspond to the MAD (mean average deviation) and maximum deviation for the full set of reactions and the MAD for the subset of reactions with X = F and CH_3 respectively. All the reported values are in kcal/mol and refer to the text for further details.

between reaction energies obtained from actual high calculations and those from the ONIOM/ONIOM-CT/ESP-ONIOM-CT calculations. Therefore, a zero deviation would indicate that the ONIOM-like method produced a reaction energy identical to that of the actual high level method.

The overall mean absolute deviation (MAD) for each method is listed in columns 2−4 in Table 4. On comparing ONIOM and ESP-ONIOM-CT models, it is notable that the mean absolute deviation (MAD) for all 80 data points is reduced from 3.85 to 1.75 kcal/mol, respectively, showing a significant improvement by 55% on the reaction energies. It is especially encouraging to see that for every high:low combination the ESP-ONIOM-CT method provides significant improvement. Moreover, on comparison of ONIOM-CT with ESP-ONIOM-CT, the MAD in the latter (1.75 kcal/mol) corresponds to a 12.5% improvement over the former (2.00 kcal/mol). By separating the test set into two subsets, one for X = F and one for X = CH_3, it is observed that the accuracy of the standard ONIOM method for the two groups is quite similar (3.80 and 3.90 kcal/mol, respectively). With the ONIOM-CT method, the MADs of 1.97 and 2.19 kcal/mol for the X = F and X = CH_3 reactions, correspond to an improvement by 48% and 44%, respectively, over the ONIOM method. Use of the ESP-ONIOM-CT method reduces the MAD further to 1.53 and 1.98 kcal/mol for the X = F and X = CH_3 reactions, respectively. Overall, the ESP-ONIOM-CT method shows significant improvement by 60% and 49%, respectively, over the ONIOM method.

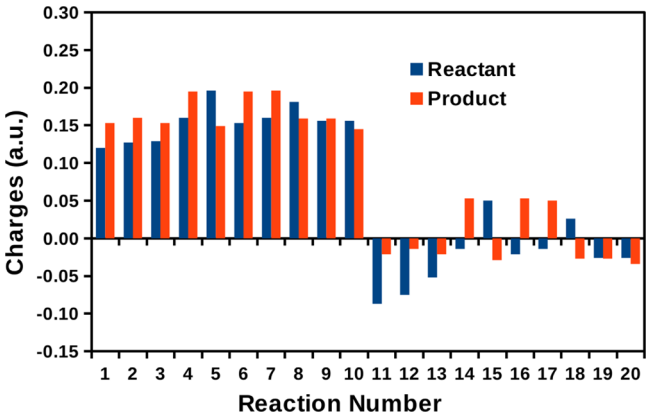


Figure 4. Distribution of the link atom charges in the reaction set 1 to 20 as reported in Table 2. In each reaction, link atoms are arranged in the order {Reactant, Product}. Reactions 1 to 10 are with link atoms replacing $-CF_3$ functional groups and reactions 11 to 20 with link atoms replacing $-C(CH_3)_3$ functional groups, respectively.

Figure 4 shows the optimized link atom charges obtained during the iterative optimization for the test set of molecules reported in Table 2. The optimized link atom charges are arranged in the order {Reactant, Product} for each reaction, from 1 to 20. Reactions 1 to 10 correspond to those in which hydrogen link atom is replacing the $-CF_3$ functional group, while reactions 11 to 20 correspond to those in which the link atom is replacing the $-C(CH_3)_3$ functional group.

The optimum charges obtained in the ESP-ONIOM-CT method are fairly small in magnitude. For example, for the 10 reactants and 10 products from Table 2, we obtain an average value of +0.160*e* for X = F and an average value of −0.013*e* for

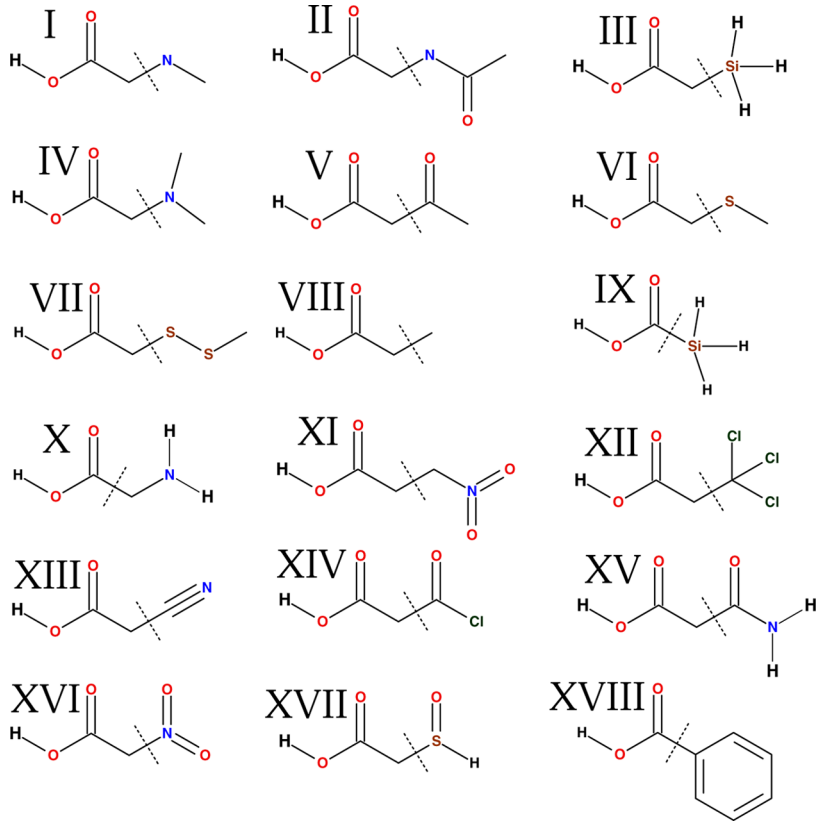


Figure 5. Structures **I** to **XVIII** are the one-link atom test systems chosen to benchmark the ESP-ONIOM-CT method. The hydrogen atoms are explicitly depicted only at the reaction centers and at the functional groups, and the rest of the dangling valencies are assumed to be saturated with hydrogen atoms. The model system is shown to the left side of the dotted line partition in each structure. See Table 5 and the text for further details.

dx.doi.org/10.1021/ct5005356 | J. Chem. Theory Comput. 2014, 10, 4351−4359

X = CH_3, using the model chemistry MP2/6-311+G(d,p):HF/3-21G. The sign and magnitude of these charges are consistent with expectations from the strong electron withdrawing nature of the $-CF_3$ group and the weak electron donating nature of the *t*-butyl group.

In the next test set, the ESP-ONIOM-CT method is benchmarked on deprotonation reactions of 18 test molecules including several where a heteroatom (i.e., other than sp^3 hybridized carbon atom) is replaced by a link H atom. The corresponding reactant molecules are depicted in Figure 5, and Table 5 reports the deprotonation energies of these one-link atom reaction set. The ONIOM-CT method (not listed) has convergence trouble and yields unphysically high link-atom charges (>0.5*e*) for some of the systems. The ONIOM and ESP-ONIOM-CT results for four different high:low combinations are listed in Table 5. The MAD for the full set is lowered from 4.01 kcal/mol for ONIOM to 2.41 kcal/mol for ESP-ONIOM-CT, an improvement of over 40%. However, the maximum deviation still remains large for some high:low combinations, indicating the challenging nature of some of the molecules in the test set.

Table 5. Comparison of the Performance of the ONIOM and ESP-ONIOM-CT Methods on the Deprotonation Reaction on One-Link Atom Test Systems, Depicted in Figure 5, with the Link Atom Substituting a Hetero-Host Atom Other than an sp^3 Hybridized Carbon Host Atom[a]

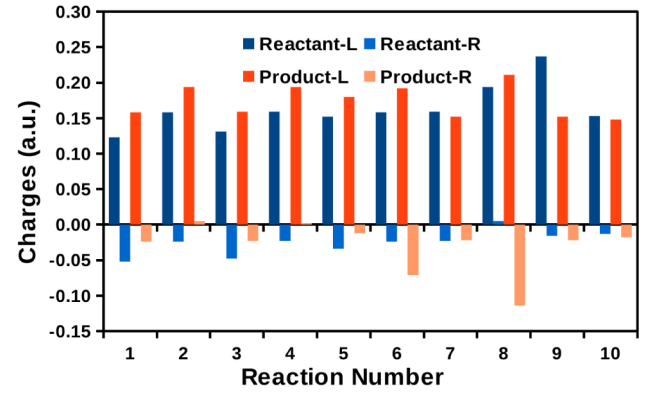
	MAD		max deviation	
high:low	ONIOM	EONIOM-CT	ONIOM	EONIOM-CT
1a	4.58	2.64	12.22	9.69
1b	4.12	2.17	7.17	4.81
2a	3.98	2.41	7.73	5.87
2b	3.37	2.42	7.64	5.81
average	**4.01**	**2.41**	**8.69**	**6.55**

[a]The three columns correspond to the MAD (mean average deviation) for the full set of reactions and maximum deviation of errors in the reaction energies, respectively. All the reported values are in kcal/mol. Refer to the text for further details.

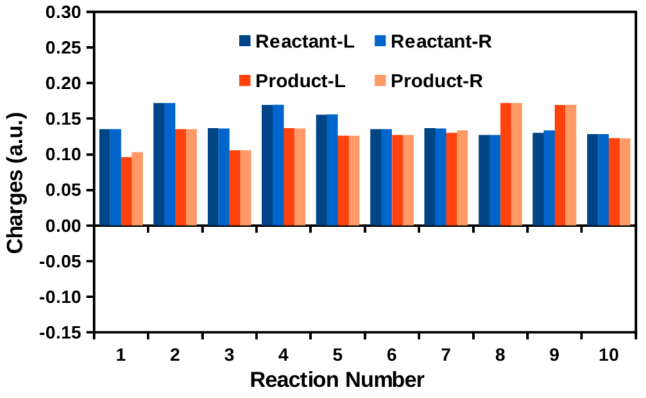
3.2. Reaction Energies of Two-link Atom Test Sets. As discussed in the Methods section, the ESP-ONIOM-CT method can be extrapolated to optimize the link atom nuclear charge and to explore the reaction involving multiple link atoms. In this section, we have benchmarked the performance of this method on two-link atom reactions similar to that for the one-link atom reactions described previously. Table 3 reports the two-link atom reaction test set employed for benchmarking the ESP-ONIOM-CT method. The model system is shown at the middle of the wavy lines, and the substituents, $-CF_3$ and $-C(CH_3)_3$, correspond to the highly electron withdrawing and electron donating functional groups, respectively.

In Table 6, we provide the results for the full test set of reactions using each high:low combination listed in Table 1. Table 6 reports the mean absolute deviation (MAD), and the maximum deviation for each high:low combination involving one electron withdrawing and one electron donating group. Here, the MAD in the reaction energy improves by 60% over the ONIOM method. Also providing an indication of reliability, the maximum deviation for each high:low combination is also decreased by up to 77% over the ONIOM method.

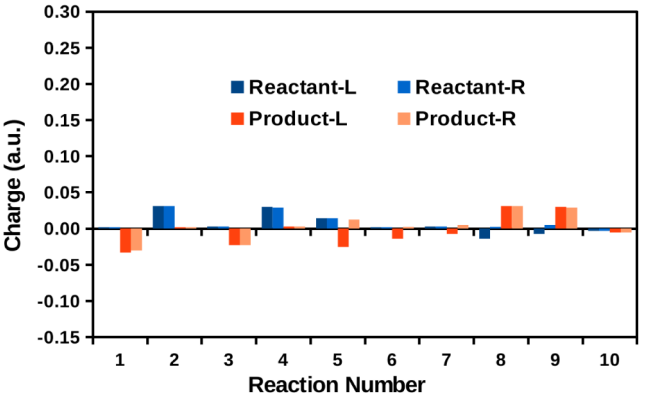
On comparison, the ESP-ONIOM-CT method outperforms the ONIOM-CT method. The ESP-ONIOM-CT brings down the MAD from 2.47 to 1.61 kcal/mol. This is an improvement of 35% over the ONIOM-CT method. The maximum deviation is also significantly smaller for the ESP-ONIOM-CT method.



(a)



(b)



(c)

Figure 6. Distribution of the extra link atom charges in the reaction set 1 to 10 reported in Table 3. For each reaction, link atom nuclear charges are arranged in the order {Reactant-Left, Reactant-Right, Product-Left, and Product-Right}. Three sets of calculations are performed on these test systems. In a, the left-hand side link atom replaces a $-CF_3$ functional group, and the right-hand side link atom replaces the $-C(CH_3)_3$ functional group, respectively. In b, both left- and right-hand side link atoms replace $-CF_3$ functional groups. In c, both left- and right-hand side link atoms replace $-C(CH_3)_3$ functional groups.

Figure 6 shows both optimized link atom charges for test sets, reported in Table 3. It includes three subsets: containing (a) one

dx.doi.org/10.1021/ct5005356 | J. Chem. Theory Comput. 2014, 10, 4351−4359

Table 6. Comparison of the Performance of the ONIOM, ONIOM-CT, and ESP-ONIOM-CT Methods on Two-Link Atom Reaction Set, Reported in Table 3[a]

	MAD			max deviation		
high:low	ONIOM	ONIOM-CT	EONIOM-CT	ONIOM	ONIOM-CT	EONIOM-CT
1a	5.06	2.99	1.14	11.59	12.58	2.63
1b	3.58	2.35	1.35	6.53	4.46	3.22
2a	3.84	2.29	1.93	6.84	6.14	5.74
2b	3.87	2.26	2.03	6.80	6.08	5.99
average	**4.09**	**2.47**	**1.61**	**7.94**	**7.32**	**4.40**

[a]The model system in each of the test sets is bonded to the host groups, with X = F and Y = CH_3. The columns correspond to the MAD (mean average deviation) and maximum deviation of errors in the reaction energies. All of the reported values are in kcal/mol. Refer to the text for further details.

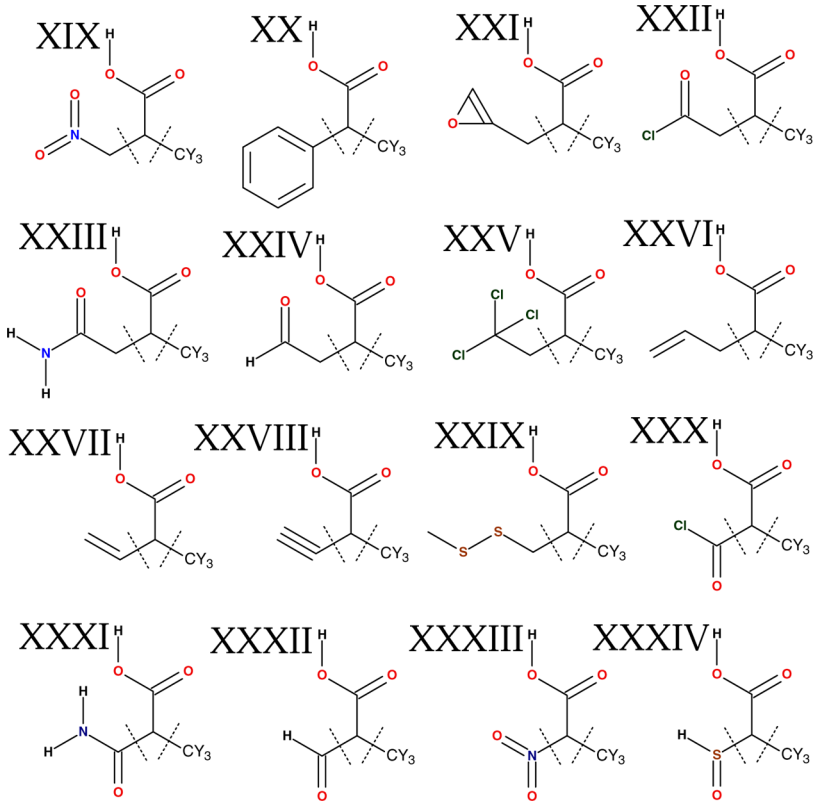


Figure 7. Structures [**XIX** to **XXXIV**], the two-link atom test systems chosen to benchmark the ESP-ONIOM-CT method. In each structure, the corresponding model system is shown in between the two dotted line partitions. The hydrogen atoms are explicitly depicted only at the reaction centers and at the functional groups, and the rest of the dangling valencies are assumed to be saturated with hydrogen atoms. See Table 7 and the text for further details.

electron withdrawing and one electron donating group, (b) two electron withdrawing groups, and (c) two electron donating groups. From Figure 6, it is clear that the link atom charges for the reactant and product are quite different in a few systems. In all cases, the derived charges are chemically meaningful. For a link atom that replaces a $-CF_3$ functional group, the range of the extra link atom charge is between 0.1 and 0.2 au, whereas in those reactions in which a link atom replaces a $-C(CH_3)_3$ functional group, the range is between −0.08 and +0.05 au, respectively. While the $-CF_3$ group is strongly electron withdrawing in all cases, the t-butyl group is slightly electron donating (on average) with a stronger influence from the model system.

For examples where both link atoms (1) replace the $-CF_3$ group and (2) replace the $-C(CH_3)_3$ group, Figure 6b and c show the magnitude of the optimized charges obtained. The link atom nuclear charges are similar when both the replacing functional groups are the same. We evaluated the performance in such cases using the ONIOM MP2/6-311+G(d,p):HF/3-21G model. For a sample set of 10 reactions involving two $-CF_3$ groups, the deviations improved from 7.53 kcal/mol for ONIOM to 3.13 kcal/mol for ESP-ONIOM-CT. For a sample set of 10 reactions involving two $-C(CH_3)_3$ groups, the deviations improved from 3.84 kcal/mol for ONIOM to 1.95 kcal/mol for ESP-ONIOM-CT. The performance improvement is similar to that with one electron withdrawing and one electron donating group.

Finally, we benchmarked the performance of the ESP-ONIOM-CT method by studying the deprotonation energies of molecules involving two-link atoms including involving a range of functional groups. The corresponding reactant molecules are depicted in Figure 7, and Table 7 reports the deprotonation energies of these two link atom reaction set. In all these test systems, the left substituent is an electron withdrawing/donating functional group, and the right side is either a Y = F or Y = CH_3 functional group.

dx.doi.org/10.1021/ct5005356 | J. Chem. Theory Comput. 2014, 10, 4351−4359

Table 7. Comparison of the Performance of the ONIOM, ONIOM-CT and ESP-ONIOM-CT Methods on the Deprotonation Reactions of Two-Link Atom Molecule Set, Depicted in Figure 7[a]

	MAD (max deviation)			MAD (X = F)			MAD (X = CH_3)		
high:low	ONIOM	ONIOM-CT	EONIOM-CT	ONIOM	ONIOM-CT	EONIOM-CT	ONIOM	ONIOM-CT	EONIOM-CT
1a	4.69 (9.48)	3.49 (9.33)	2.35 (5.35)	6.07	4.36	2.80	3.30	2.61	1.89
1b	4.27 (8.24)	2.57 (7.55)	2.39 (4.93)	5.34	2.84	2.75	3.19	2.30	2.03
2a	4.23 (9.35)	3.92 (9.80)	2.73 (6.89)	5.49	4.92	3.37	2.96	2.91	2.09
2b	4.30 (9.41)	3.09 (9.60)	2.85 (6.99)	5.59	3.89	3.46	3.01	2.29	2.24
average	**4.37 (9.12)**	**3.27 (9.07)**	**2.58 (6.04)**	**5.62**	**4.00**	**3.10**	**3.12**	**2.53**	**2.06**

[a]The corresponding model system has two-link atoms, which replaced the bonded host carbon and hetero-carbon atoms. All of the reported values are in kcal/mol. Refer to the text for further details.

Inspection of the overall mean absolute deviation (MAD) for each method reported in Table 7 shows that the ESP-ONIOM-CT method provides significant overall improvement over the ONIOM and ONIOM-CT methods. It is notable that the mean absolute deviation (MAD) for all the data points in this challenging test set is reduced from 4.37 to 2.58 kcal/mol, indicating a significant improvement by 41% on the reaction energies. It is especially encouraging to see that for every high:low combination the ESP-ONIOM-CT method provides significant improvements, at times reducing the MAD by 50%.

We have separated the test set into two subsets, one for X = F and one for X = CH_3. Using the ONIOM method, the X = F reactions have an overall MAD of 5.62 kcal/mol, while the analogous reactions with X = CH_3 have a MAD of 3.12 kcal/mol. The ESP-ONIOM-CT method improves the performance discrepancy between the two subsets of reaction with MADs of 3.10 and 2.06 kcal/mol for X = F and X = CH_3, respectively.

We also carried out a few additional calculations to test the stability of the ESP-ONIOM-CT energy with increasing the low-level basis set size. As an example, using the high:low combination MP2/6- 311+G(d,p):HF/6-311+G(d,p), we found an improved performance of the ESP-ONIOM-CT method compared to the ONIOM-CT method. The method is thus stable with respect to the level of theory and the basis set.

There are a variety of models for procuring the ESP fitted charges in the literature that differ in the way of choosing the points in space for performing the electrostatic potential least-square fit.[45] In all these ESP fitting models, ESP points used for fitting the charges must lie outside the van der Waals surface of the molecule. As Kollman and co-workers[45] have pointed out, buried charges tends to be poorly determined because even the closest surface points at which the ESP is evaluated are relatively far away and in general closer to a different atomic charge. However, the fitness function we have employed for optimizing the link atom nuclear charge is a function of the model system nuclear ESP, and the regional ESP between the supporting and link atoms. These selected sets of points for performing the ESP fit are not under the influence of multiple neighboring atoms. Hence the ESP-ONIOM-CT method is not bounded by the buried atom problem.

4. CONCLUSIONS

In this work, we have proposed an inexpensive approach, ESP-ONIOM-CT, to incorporate inter-regional charge redistribution effects within the ONIOM formalism. In this method, link-atom centers are treated as an electron buffer region, and their nuclear charges are shifted to balance the electrostatic potential of region I in the model system to be the same as that in the real system. This new ESP-ONIOM-CT method is independent of any arbitrary definition of charges, is defined on the basis of a physical observable, and is less basis set dependent than previous approaches. The link-atom optimization of the charges is done in an iterative procedure that is computationally inexpensive. Our method is easily extended for exploring reaction energies on systems with multiple link atoms. We have benchmarked the ESP-ONIOM-CT method on reaction sets having one- and two link-atoms, using four different combinations of high:low levels of theory. The optimum link-atom nuclear charges are small in magnitude and are physically intuitive. From these benchmark studies, ESP-ONIOM-CT reaction energies improved significantly by 40−60% over the standard ONIOM method.

■ ASSOCIATED CONTENT

Supporting Information

The performance for the individual systems discussed in the paper and the Cartesian coordinates of all the species are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kraghava@indiana.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge support from NSF grant CHE-1266154 at Indiana University.

■ REFERENCES

(1) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 084108−084119.
(2) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *124*, 114104.
(3) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kallay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129−4141.
(4) Ochterski, J. W.; Petersson, G. A.; M, J. A., Jr *J. Chem. Phys.* **1996**, *104*, 2598.
(5) Karton, A.; Rabinovich, E.; Martin, J. M.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.
(6) Tajti, A.; Szalay, P.; Csaszar, A.; Kallay, M.; Gauss, J.; Valeev, E.; Flowers, B.; Vazquez, J.; Stanton, J. *J. Chem. Phys.* **2004**, *121*, 11599.
(7) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(8) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.
(9) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(10) Aqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523−2544.
(11) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170−1179.
(12) Mordasini, T.; Thiel, W. *Chimia* **1998**, *52*, 288−291.
(13) Monard, G.; Merz, K. M. *Acc. Chem. Res.* **1999**, *32*, 904−911.

(14) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(15) Field, M. J. *J. Comput. Chem.* **2002**, *23*, 48−58.
(16) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185−199.
(17) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959−1967.
(18) Svensson, M.; Humbel, S.; Froese, R.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.
(19) Karadakov, P. B.; Morokuma, K. *Chem. Phys. Lett.* **2000**, *317*, 589−596.
(20) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419−1432.
(21) Vreven, T.; Mennucci, B.; da Silva, C.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62−72.
(22) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125−132.
(23) Rega, N.; Iyengar, S.; Voth, G.; Schlegel, H.; Vreven, T.; Frisch, M. *J. Phys. Chem. B* **2004**, *108*, 4210−4220.
(24) Gogonea, V.; Westerhoff, L. M.; Merz, K. M. *J. Chem. Phys.* **2000**, *113*, 5604.
(25) Yang, W.; Lee, T. *J. Chem. Phys.* **1995**, *103*, 5674.
(26) Dixon, S. L.; Merz, K. M. *J. Chem. Phys.* **1996**, *104*, 6643.
(27) Zhang, Y.; Lin, H. *J. Chem. Theory Comput.* **2008**, *4*, 414−425.
(28) Zhang, Y.; Lin, H. *Theor. Chem. Acc.* **2010**, *126*, 315−322.
(29) Pezeshki, S.; Lin, H. *Mol. Simul.* **2014**, *1*, 1−22.
(30) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.
(31) Pu, J. Z.; Gao, G. L.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 632−650.
(32) Pu, J. Z.; Gao, J. L.; Truhlar, D. G. *ChemPhysChem* **2005**, *6*, 1853−1865.
(33) Aviram, A.; Ratner, M. A. *Chem. Phys. Lett.* **1974**, *29*, 277−283.
(34) Nitzan, A.; Ratner, M. A. *Science* **2003**, *300*, 1384−1389.
(35) Pacheco, A.; Iyengar, S. S. *J. Chem. Phys.* **2010**, *133*, 044106.
(36) Mayhall, N. J.; Raghavachari, K. *J.Chem. Theory Comput.* **2010**, *6*, 3131−3136.
(37) Politzer, P.; Truhlar, D. G. *Chemical Applications of Atomic and Molecular Electrostatic Potentials: Reactivity, Structure, Scattering, and Energetics of Organic, Inorganic, and Biological Systems.*; Springer: Boston, MA, 1981.
(38) Shrivastava, I. H.; Gadre, S. R. *Int. J. Quantum Chem.* **1994**, *49*, 397−407.
(39) Gadre, S. R.; Shirsat, R. N. *Electrostatics of Atom and Molecules*; Universities Press: Hyderabad, 2000.
(40) Lakshmi, B.; Samuelson, A. G.; Jovan Jose, K. V.; Gadre, S. R.; Arunan, E. *New J. Chem.* **2005**, *29*, 371−377.
(41) Politzer, P.; Parr, R. G. *J. Chem. Phys.* **1974**, *61*, 4258−4262.
(42) Gadre, S. R.; Kulkarni, S. A.; Suresh, C.; Shrivastava, I. H. *Chem. Phys. Lett.* **1995**, *239*, 273−281.
(43) Frisch, M. J. et al. *Gaussian 09*, Revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
(44) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons: Chichester, England, 2007; pp 293−296.
(45) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.