Energy Decomposition Scheme Based on the Generalized Kohn–Sham Scheme

Peifeng Su,* Zhen Jiang, Zuochang Chen, and Wei Wu

The State Key Laboratory of Physical Chemistry of Solid Surfaces, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, Fujian 361005, China

ABSTRACT: In this paper, a new energy decomposition analysis scheme based on the generalized Kohn–Sham (GKS) and the localized molecular orbital energy decomposition analysis (LMO-EDA) scheme, named GKS-EDA, is proposed. The GKS-EDA scheme has a wide range of DFT functional adaptability compared to LMO-EDA. In the GKS-EDA scheme, the exchange, repulsion, and polarization terms are determined by DFT orbitals; the correlation term is defined as the difference of the GKS correlation energy from monomers to supermolecule. Using the new definition, the GKS-EDA scheme avoids the error of LMO-EDA which comes from the separated treatment of $E_X$ and $E_C$ functionals. The scheme can perform analysis both in the gas and in the condensed phases with most of the popular DFT functionals, including LDA, GGA, meta-GGA, hybrid GGA/meta-GGA, double hybrid, range-separated (long-range correction), and dispersion correction. By the GKS-EDA scheme, the DFT functionals assessment for hydrogen bonding, vdW interaction, symmetric radical cation, charge-transfer, and metal–ligand interaction is performed.

A. INTRODUCTION

As one of the useful tools for intermolecular interactions in quantum chemistry, the energy decomposition analysis (EDA) scheme divides the total interaction energy into several individual terms with ab initio molecular orbital (MO) theory or density functional theory (DFT) methods.1–23 As one of the recently developed EDA schemes, the LMO-EDA (localized molecular orbital energy decomposition analysis) method provides analysis for open- or closed-shell interacting systems in the gas phase using Hartree–Fock (HF) orbitals or Kohn–Sham (KS) orbitals.24 It divides the total HF or KS interaction energy $\Delta E_{\text{TOT}}$ into $\Delta E_{\text{ele}}$ (electrostatic), $\Delta E_{\text{ex}}$ (exchange), $\Delta E_{\text{disp}}$ (dispersion), $\Delta E_{\text{pol}}$ (polarization), and $\Delta E_{\text{corr}}$ (correlation) terms. Due to its simplicity and robustness, the LMO-EDA scheme has been widely used for various weak and strong interactions.24–40

For the LMO-EDA analysis at the KS-DFT level, the exchange and repulsion terms are defined as the differences of the exchange ($E_X$) and correlation ($E_C$) functionals ongoing from monomers to supermolecule respectively. The repulsion and polarization terms are determined by KS orbitals and exchange functionals in certain intermediate steps. For most of DFT functionals, because the aim of the parametrization is to obtain accurate estimation of the total $E_{\text{SOC}}$ the separated treatment of $E_X$ and $E_C$ in the LMO-EDA scheme would introduce the error into the individual interaction terms, leading to confused insight for some cases. For example, it is known that BLYP is not suitable for $\text{vdW}$ interaction due to the lack of dispersion correction. However, for the T-shaped benzene dimer by BLYP, which will be discussed in the next section of this paper, the LMO-EDA analysis shows that the dispersion interaction plays the most important role for the $\text{vdW}$ interaction. It does not make sense. Therefore, the analysis accuracy of LMO-EDA with DFT functional needs to be further improved. Moreover, the current LMO-EDA scheme cannot work with dispersion-corrected and range-separated functionals, resulting in the limited application of LMO-EDA. The functional adaptability of LMO-EDA need to be expanded.

One of the possible ways for improving the LMO-EDA scheme is to introduce the generalized Kohn–Sham (GKS) scheme proposed by Levy et al.41 The GKS scheme maps a real system to an interacting model system represented by a single Slater determinant. In 2010, Baer et al. applied the GKS scheme to present systematic interpretations for range-separated and hybrid functionals, showing the accuracy of range-separated functionals for complicated electronic structure.42

The motivation of this work is to present a generalized Kohn–Sham based energy decomposition analysis scheme, named GKS-EDA, for interaction analysis in the gas and the condensed phase with various DFT functionals. Sequentially, GKS-EDA is employed in the assessment of various DFT functionals, which can be divided into five categories in this work.

The first category is local functional, involving local density approximation (LDA), generalized gradient approximation (GGA), and meta generalized gradient approximation (meta-GGA). The second category is a hybrid functional involving the exact HF exchange for functional description, including hybrid GGA and hybrid meta-GGA. The most popular hybrid GGA functional is B3LYP.33,43 The hybrid meta-GGA functionals, for example, TPSSh,45 M06-2X,46 M08-HX,47 etc., whose total energy depends on the occupied orbitals not only through the
HF exchange terms but also through the noninteracting spin kinetic energy densities, are shown to have better performance than local and hybrid GGA functionals.\textsuperscript{37} The third category is double hybrid functional, which depends not only on occupied orbitals but also on unoccupied orbitals, for example, B2PLYP.\textsuperscript{40} The fourth category is the dispersion-corrected functional, including vdW-DF,\textsuperscript{49,50} DFT-D,\textsuperscript{51–53} DCACP,\textsuperscript{54} etc. DFT-D takes the dispersion effect into account with a damped dispersion correction, vdW-DF includes dispersion via nonlocal correlation in $E_{XC}$ functional, whereas DCACP takes the form of pseudopotential for dispersion. The final category is range-separated functional, also called long-range correction operators, regarded as external potential operators.

The nonlocal functionals, including hybrid, double hybrid, and range-separated types, are usually beyond the original Kohn–Sham framework because their exchange parts depend both on KS orbitals and density. They can be treated by the generalized Kohn–Sham (GKS) scheme.\textsuperscript{31} In the GKS scheme, not only the noninteracting kinetic energy but also the part of electron–electron interaction is treated exactly. As denoted by Levy and Baer et al., the GKS form is flexible.\textsuperscript{41,42} When the HK functional is defined as

\[
F_\text{HK}(\rho) = \min_{\Phi} \langle \Phi | \hat{T} + \hat{V}_{\text{ex}} | \Phi \rangle
\]

the GKS scheme is known as the Hartree–Fock–Kohn–Sham (HF-KS) scheme. By HF-KS, the correlation energy can be expressed as

\[
E_{\text{C}}^{\text{GKS}} = F(\rho) - F_\text{HK}(\rho)
\]

$\rho$ is achieved from the orbitals $\phi_i$ computed by the following equations:

\[
\delta E_{\text{C}}^{\text{GKS}} = \frac{\delta E_{\text{C}}}{\delta \rho} = \langle \Phi | \hat{V}_{\text{xc}} | \Phi \rangle - j(\rho)
\]

\[
E_{\text{C}}^{\text{GKS}} = E_{\text{C}}(\rho) + (1 - \gamma)(E_{\text{X}}(\rho) - E_{\text{X}}^{\text{GKS}})
\]

In eq 8, $j(\rho)$ is the electron–electron repulsion energy. In eq 10, $E_{\text{X}}(\rho)$ and $E_{\text{C}}(\rho)$ are the exchange and correlation functionals, respectively. $\gamma$ is the portion of the exact HF energy, ranging from 0 to 1.0. If double hybrid functional is applied, $E_{\text{C}}(\rho)$ is partially replaced by portion of MP2 correlation energy. In the KS scheme, the unknown energy is a density functional instead of an orbital functional, but in the GKS scheme, the exchange energy is exact, determined only by DFT orbitals, the electronic correlation energy is accounted for by $E_{\text{C}}^{\text{GKS}}$, determined by density and orbitals simultaneously. As claimed by Baer et al., $E_{\text{C}}^{\text{GKS}}$ encapsulates the entire immensity of the electronic structure problem, and it is this functional for which approximation must be crafted under DFT.\textsuperscript{42}

The parentheses in the left-hand side of eq 6 can be written as

\[
F_{\text{GKS}} = F^D + v_{\text{C}}^{\text{GKS}}
\]

\[
F^D = -\frac{1}{2} \hat{V}^2 + v_{\text{X}}^{\text{GKS}} + v_{\text{e}} + v_{\text{ne}}
\]

For a supermolecule consisting of monomers, the total interaction energy $\Delta E_{\text{TOT}}$ is expressed as

\[
\Delta E_{\text{TOT}} = E_{\text{c}}^{\text{GKS}} + E_{\text{X}}^{\text{GKS}} + E_{\text{V}}^{\text{GKS}}
\]
\[ \Delta E_{\text{TOT}} = \langle \Phi_S | H_{\text{GKS}} | \Phi_S \rangle - \sum_M \langle \Phi_M | H_{\text{GKS}} | \Phi_M \rangle \]

\[ \quad = \left( \langle \Phi_S | H_{\text{D}} | \Phi_S \rangle - \sum_M \langle \Phi_M | H_{\text{D}} | \Phi_M \rangle \right) \]

\[ \quad + \left( \langle \Phi_S | H_{\text{GKS}} | \Phi_S \rangle - \sum_M \langle \Phi_M | H_{\text{GKS}} | \Phi_M \rangle \right) \]

\[ \quad = \Delta E_{\text{ex}} + \Delta E_{\text{ex}} + \Delta E_{\text{ex}} + \Delta E_{\text{pol}} \]  

(14)

These interaction terms can be obtained gradually by several intermediate states. In step 0, there is not any interaction among monomers. The wave function of step 0, \( \Phi^{\text{DP}} \), is the product of monomers’ wave function:

\[ \Phi^{\text{DP}} = \prod_M \Phi_M \]  

(15)

When the counterpoise (CP) method is applied for the basis set superposition error (BSSE), the Kohn–Sham orbitals in \( \Phi_M \) are expanded with supermolecule’s basis set.\(^{68}\)

The energy of the step 0 can be expressed as

\[ E_S^{(0)} = \langle \Phi^{\text{DP}} | H_{\text{D}} | \Phi^{\text{DP}} \rangle \]

\[ \quad = \sum_M \langle \Phi_M | H_{\text{D}} | \Phi_M \rangle \]  

(16)

At step 1, there is not any orbital overlap among monomers. The wave function \( \Phi^{\text{DP}} \) is unchanged. If the sum of monomers’ GKS operators \( F_{\text{D}} \) is replaced by the supermolecule’s GKS operator \( F_{\text{DP}} \), the corresponding energy \( E_S^{(1)} \) can be written as

\[ E_S^{(1)} = \langle \Phi^{\text{DP}} | F_{\text{D}} | \Phi^{\text{DP}} \rangle \]  

(17)

From the step 0 to the step 1, the electrostatic interaction arising from the variation of the GKS operators with the wave function \( \Phi^{\text{DP}} \), can be defined as

\[ \Delta E_{\text{ex}} = \langle \Phi^{\text{DP}} | F_{\text{D}}^{(1)} | \Phi^{\text{DP}} \rangle - \sum_M \langle \Phi_M | F_{M}^{\text{DP}} | \Phi_M \rangle \]  

(18)

For the step 2, the orbital overlap among monomers is taken into considered. The wave function must be in the antisymmetric and normalized form:

\[ \Phi^{\text{ASN}} = \hat{A} \Phi^{\text{DP}} \]  

(19)

In eq 19, \( \hat{A} \) is the antisymmetrization operator. With the requirement for antisymmetry and orthonormality of DFT orbitals, the corresponding approximate energy \( E_S^{(2)} \) can be expressed as

\[ E_S^{(2)} = \langle \Phi^{\text{ASN}} | F_{\text{D}}^{\text{DP}} | \Phi^{\text{ASN}} \rangle \]  

(20)

The electronic density from the antisymmetrized and normalized wave function \( \Phi^{\text{ASN}} \), compared to that from \( \Phi^{\text{DP}} \), decreases between monomers and increases at the side of each monomer remote from the other, leading to repulsive force among monomers.\(^{67}\) The energy difference between steps 1 and 2 is defined as exchange repulsion interaction:

\[ \Delta E_{\text{ex}} = \langle \Phi^{\text{ASN}} | F_{\text{D}}^{\text{DP}} | \Phi^{\text{ASN}} \rangle - \langle \Phi^{\text{DP}} | F_{\text{D}}^{\text{DP}} | \Phi^{\text{DP}} \rangle \]  

(21)

As proposed by Stone and Li et al.,\(^{20,66,67}\) the \( \Delta E_{\text{ex}} \) can be further divided into the exchange term \( \Delta E_{\text{ex}} \) and the repulsion term \( \Delta E_{\text{ex}} \). The exchange term \( \Delta E_{\text{ex}} \) is explicitly expressed as the variation of the GKS exchange energy \( E_{X,GKS}^{(1)} \) from step 1 to step 2.

\[ \Delta E_{\text{ex}} = E_{X,GKS}^{(2)} - E_{X,GKS}^{(1)} \]  

(22)

The repulsion energy is then defined as

\[ \Delta E_{\text{rep}} = \Delta E_{\text{ex}} - \Delta E_{\text{ex}} \]  

(23)

At step 3, with the SCF procedure, \( \Phi^{\text{ASN}} \) is optimized as \( \Phi_S \). The polarization energy can be defined as

\[ \Delta E_{\text{pol}} = \langle \Phi_S | F_{\text{D}}^{\text{DP}} | \Phi_S \rangle - \langle \Phi^{\text{ASN}} | F_{\text{D}}^{\text{DP}} | \Phi^{\text{ASN}} \rangle \]  

(24)

Equation 24 means that the polarization term is equal to orbital relaxation energy by the SCF procedure. In some EDA schemes, orbital relaxation energy includes polarization and charge transfer. However, as said by Davidson et al.: “This partitioning is not clean because it depends on grouping the basis set into functions associated with each monomer ... In the limit of a complete basis set on both A and B, the charge transfer and polarization in these definitions become equivalent.”\(^{69}\) In GKS-EDA, it is preferred that only polarization energy is used since monomers’ orbitals are expanded using supermolecule’s basis set for BSSE correction.

The second bracket in the right-hand side of eq 13, illustrating the difference of the GKS correlation energy from monomers to supermolecule, can be defined as the correlation term:

\[ \Delta E_{\text{corr}} = \langle \Phi_S | H_{\text{GKS}}^{(1)} | \Phi_S \rangle - \sum_M \langle \Phi_M | H_{\text{GKS}}^{(1)} | \Phi_M \rangle \]  

(25)

Using eq 10, the correlation term can be further written as

\[ \Delta E_{\text{corr}} = \Delta E_{C}(\rho) + (1 - \gamma)(\Delta E_X(\rho) - \Delta E_{X,GKS}(\rho)) \]  

(26)

where \( \Delta E_{C}(\rho) \) and \( \Delta E_X(\rho) \) are the difference of the exchange and correlation functionals between monomers and supermolecule.

If double hybrid functional is applied, \( \Delta E_{\text{corr}} \) can be expressed as

\[ \Delta E_{\text{corr}} = (1 - \gamma)(\Delta E_X - \Delta E_{X,GKS}) + (1 - C_{MP2})(\Delta E_{C} + C_{MP2}\Delta E_{MP2}^{(2)}) \]  

(27)

where \( \Delta E_{MP2}^{(2)} \) is the MP2 correlation energy from DFT orbitals with the amount accounted by \( C_{MP2} \).

Specifically, when dispersion corrected DFT is applied, an extra term \( \Delta E_{\text{disp}} \) accounting the dispersion interaction is introduced:

\[ \Delta E_{\text{disp}} = E_{\text{disp}}^{(2]} - \sum M \]  

(28)

If Grimme’s DFT-D scheme is specified, the \( E_{\text{disp}}^{(2]} \) can be expressed as\(^{31–33} \)
In eq 29 the summation is over all atom pairs in the 
supermolecule, \( s_i \) are global (functional dependent) scaling 
factors, \( C_{AB}^{n} \) is the averaged \( n \)th dispersion coefficient for atom 
pair \( AB \), \( R_{AB} \) is the distance between atoms A and B, and \( f_{\text{damp}} \) are 
damping functions determining the range of the dispersion 
correction.

In summary, for the gas phase, the GKS-EDA scheme divides 
the total interaction energy into the following terms with the 
final term optional for DFT-D:

\[
\Delta E_{\text{TOT}} = \Delta E_{\text{ele}} + \Delta E_{\text{ex}} + \Delta E_{\text{rep}} + \Delta E_{\text{pol}} + \Delta E_{\text{corr}} + \Delta E_{\text{damp}}
\]

It is worthwhile to point out the difference between LMO-
EDA and GKS-EDA here. In LMO-EDA, the contribution of 
\( E_{\text{XC}} \) functionals is divided into \( E_X \) and \( E_C \) which uniquely 
determine the \( \Delta E_{\text{ele}} \) and \( \Delta E_{\text{damp}} \) terms, respectively. The \( \Delta E_{\text{rep}} \) and \( \Delta E_{\text{pol}} \) terms are determined by the KS orbitals and \( E_X \) functional together. In GKS-EDA, the contribution of the \( E_{\text{XC}} \) functional is not separated. The \( \Delta E_{\text{corr}} \) is defined as the GKS 
correlation energy from monomers to supermolecule, while the 
other terms are governed by DFT orbitals.

2. GKS-EDA Scheme in Solution. Intermolecular 
interactions in various solvated environments are significant, 
which can be well understood by condensed phase energy 
decomposition analysis schemes by Tomasi,70 de1 Valle,71 
Robert,72 Bickelhaupt,73 Fedorov,74 and Su75 etc. EDA-PCM is 
proposed by our group on the basis of LMO-EDA and implicit 
solvation models. In this paper, only 

\[
G^{\text{TOT}} = \langle \Phi^{\text{SOL}} | f_{\text{GKS}} | \Phi^{\text{SOL}} \rangle + \Delta G^{\text{SOL}}
\]

The first parentheses in the right-hand side of eq 34, describing 
the variation of the solvation free energy by KS determinant 
from monomers to supermolecule, can be further decomposed 
into electrostatic, exchange, repulsion, polarization, and 
correlation terms:

\[
\langle \Phi^{\text{SOL}} | f_{\text{GKS}} | \Phi^{\text{SOL}} \rangle = \sum_M \langle \Phi^{\text{M}} | f_{\text{GKS}} | \Phi^{\text{M}} \rangle
\]

\[
= \Delta G_{\text{ele}}^{\text{SOL}} + \Delta G_{\text{ex}}^{\text{SOL}} + \Delta G_{\text{rep}}^{\text{SOL}} + \Delta G_{\text{pol}}^{\text{SOL}} + \Delta G_{\text{corr}}^{\text{SOL}}
\]

The expressions of \( \Delta G_{\text{ele}}^{\text{SOL}} \), \( \Delta G_{\text{ex}}^{\text{SOL}} \), \( \Delta G_{\text{rep}}^{\text{SOL}} \), \( \Delta G_{\text{pol}}^{\text{SOL}} \), and \( \Delta G_{\text{corr}}^{\text{SOL}} \) are the same with 
\( \Delta E_{\text{ele}} \), \( \Delta E_{\text{ex}} \), \( \Delta E_{\text{rep}} \), \( \Delta E_{\text{pol}} \), and \( \Delta E_{\text{corr}} \), respectively. The notation “G” is applied instead of “E” because 
these interaction terms are determined by \( \Phi^{\text{SOL}} \) and \( \rho^{\text{SOL}} \) in the 
condensed phase instead of \( \Phi \) and \( \rho \) in the gas phase.

The second parentheses in the right-hand side of eq 34, which 
accounts the free energy penalty by environment due to 
monomers’ interaction, defined as the desolvation term \( \Delta G_{\text{desol}}^{\text{SOL}} \):

\[
\Delta G_{\text{desol}}^{\text{SOL}} = \sum_M \Delta G_{\text{SOL}}^{\text{M}}
\]

Finally, if dispersion correction DFT is applied, the dispersion 
term \( \Delta E_{\text{damp}} \) is introduced. When the Grimme’s correction DFT 
is specified, the \( \Delta E_{\text{damp}} \) definition is same with the \( \Delta E_{\text{damp}} \) one.

In summary, the total interaction energy in the condensed 
phase is decomposed into electrostatic, exchange, repulsion, 
polarization, correlation, and dispersion terms:

\[
\Delta G^{\text{TOT}} = \Delta G_{\text{ele}} + \Delta G_{\text{ex}} + \Delta G_{\text{rep}} + \Delta G_{\text{pol}} + \Delta G_{\text{corr}} + \Delta G_{\text{damp}}
\]

C. COMPUTATIONAL DETAILS

All calculations were performed with the quantum chemistry 
program package GAMESS.86 Based on LMO-EDA, the GKS-
EDA scheme is implemented, which may be available in the 
released version of GAMESS recently. The counterpoise (CP) 
method is applied for basis set superposition error (BSSE) 
without dispersion correction.88
The Journal of Physical Chemistry A

As for the test examples in the gas phase, the four complexes shown in Figure 1 are selected, including the water dimer, benzene dimer, C\textsubscript{6}H\textsubscript{4}⋯F\textsubscript{2} and Ar\textsubscript{2}+ complexes.

Figure 1. Geometries of the water dimer, benzene dimer, C\textsubscript{6}H\textsubscript{4}⋯F\textsubscript{2} and Ar\textsubscript{2}+ complexes.

benzene dimer, C\textsubscript{6}H\textsubscript{4}⋯F\textsubscript{2} and Ar\textsubscript{2}+. The aug-cc-pVnZ (n = D, T) and cc-pVTZ basis sets were employed, denoted as ACCD, ACCT, and CCT, respectively, in the following discussions. The geometries of the water dimer, C\textsubscript{6}H\textsubscript{4}⋯F\textsubscript{2}, and Ar\textsubscript{2}+ are optimized at the MP2/ACCD level, and the benzene dimer is optimized at the MP2/ACCD level.

As for the test examples in the condensed phase, the copper ligand model molecule for the active site of the type 1 copper protein environment.

Table 1. DFT Functionals Used in This Work

<table>
<thead>
<tr>
<th>type</th>
<th>functional name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGA</td>
<td>BLYP</td>
</tr>
<tr>
<td>hybrid GGA</td>
<td>B3LYP, B97</td>
</tr>
<tr>
<td>meta-GGA</td>
<td>M06-L</td>
</tr>
<tr>
<td>hybrid meta-GGA</td>
<td>M06-2X</td>
</tr>
<tr>
<td>range-separated</td>
<td>LC-BLYP, CAM-B3LYP, M11, wB97, wB97X</td>
</tr>
<tr>
<td>dispersion correction, GGA</td>
<td>BLYP-D, B97-D</td>
</tr>
<tr>
<td>range-separated + dispersion corrections</td>
<td>wB97X-D</td>
</tr>
</tbody>
</table>

D. RESULTS AND DISCUSSIONS

1. Water Dimer. Table 2 shows the GKS-EDA analysis results with the ACCT basis set for the water dimer interaction. The best theoretical estimated interaction energy is ∼5.02 kcal/mol at the CCSD(T)/complete basis set (CBS) level, whereas the corresponding experimental value is −5.40 kcal/mol.

By the GKS-EDA analysis, as a typical hydrogen bond, besides the electrostatic interaction, the water dimer interaction also includes the obvious correlation and dispersion contributions. The Δ\textsubscript{E\textsubscript{dip}} term only exists in the analysis by B97-D, BLYP-D, and wB97X-D functionals. First, the BLYP, B3LYP, and B97 functionals underestimate the total interaction energies because of the small correlation terms. As a double hybrid functional, B2PLYP slightly enhances the description of the total binding energy compared to BLYP and B3LYP due to the small repulsive term. The meta-GGA functional M06-L provides the smallest Δ\textsubscript{E\textsubscript{corr}} and Δ\textsubscript{E\textsubscript{disp}} among all the test functionals, leading to the moderate total binding energy of −4.61 kcal/mol. Owing to the larger Δ\textsubscript{E\textsubscript{disp}} and Δ\textsubscript{E\textsubscript{corr}} terms, M06-2X's total interaction energy is better than M06-L.

Second, the range-separated DFT functionals, except M11, tend to mimic the dispersion interaction by relatively large correlation energy, in agreement with the comment by Ziegler and Grimme et al. As a result of the appropriate description for the correlation interaction, CAM-B3LYP is able to predict the hydrogen bonding quite well, whereas, because of the quite large correlation terms, LC-BLYP, wB97, and wB97X somewhat overestimate the total interaction energies. The total interaction energy by M11 is the smallest among all the tested range-separated functionals due to the small correlation term.

Third, in the GKS-EDA analysis by B97-D, BLYP-D, and wB97X-D, there are two significant terms for the hydrogen bonding, including Δ\textsubscript{E\textsubscript{corr}} and Δ\textsubscript{E\textsubscript{disp}}. B97-D and BLYP-D provide the similar Δ\textsubscript{E\textsubscript{disp}} values. Their performance is attributed to the Δ\textsubscript{E\textsubscript{corr}} terms, leading to the total energies of −4.92 and −4.30 kcal/mol by BLYP-D and B97-D, respectively. By wB97X-D, given the additional dispersion term, the Δ\textsubscript{E\textsubscript{corr}} value is relatively small to avoid the overestimation, leading to a satisfied total interaction energy.

2. Benzene Dimer. Table 3 shows the GKS-EDA analysis results for the T-shaped benzene dimer using the CCD, CCT, and ACCT basis sets by B97-D, BLYP-D, wB97X-D, LC-BLYP, M06-2X, and B2PLYP. The GKS-EDA analysis shows that the performance is controlled by the dispersion and correlation terms. The other interaction terms by the different functionals are similar.

First, the dispersion correction DFT functionals predict the binding energies quite well, close to the estimated CCSD(T)/CBS value of −2.84 kcal/mol by Kim et al. With the variation of the basis sets, their total interaction energies along with the range-separated methods provide the similar performance is controlled by the dispersion and correlation terms. The other interaction terms by the different functionals are similar.

First, the dispersion correction DFT functionals predict the binding energies quite well, close to the estimated CCSD(T)/CBS value of −2.84 kcal/mol by Kim et al. With the variation of the basis sets, their total interaction energies along with the range-separated methods provide the similar performance is controlled by the dispersion and correlation terms. The other interaction terms by the different functionals are similar.
individual interaction terms remain stable. The total interaction energies are governed by the dispersion terms, which are $-6.23$, $-5.63$, and $-3.36$ kcal/mol by B97-D, BLYP-D, and wB97X-D, respectively, with the ACCT basis set. For wB97X-D, the small dispersion energy is compensated by the large correlation term, leading to the proper total interaction energy.

Second, similar to the analysis for the water dimer, LC-BLYP and M06-2X provide extra correlation terms to mimic the dispersion interaction, leading to the attractive binding energies. In detail, with the ACCT basis set, the correlation values of LC-BLYP and M06-2X, $-5.17$ and $-5.95$ kcal/mol respectively, are close to the sum of $\Delta E_{\text{corr}}$ and $\Delta E_{\text{disp}}$ in the analysis results by B97-D, BLYP-D, and wB97X-D. With the larger correlation value, the performance of M06-2X is better than that of LC-BLYP.

Third, as for B2PLYP, the performance is more sensitive to the basis set variations. Although the MP2 correlation energy is involved in the correlation term, the $\Delta E_{\text{corr}}$ values, $-2.75$, $-3.34$, and $-3.57$ kcal/mol using CCD, CCT, and ACCT, respectively, are smaller than those by M06-2X and LC-BLYP, leading to the repulsive total interaction energy.

Finally, to explore the error of LMO-EDA mentioned in the Introduction, the LMO-EDA and GKS-EDA analysis results for the T-shaped benzene dimer by the BLYP functional are listed in Table 4. By the two EDA schemes, it is shown that BLYP provides the repulsive description for the benzene interaction.
One of the theoretical viewpoints considers the charge transfer 
nism of the addition reaction of halogen molecule to ethylene.

Truhlar et al. First, the local functionals, involving BLYP, 
whereas the viewpoint is that the dispersion interaction is more important 
energy arises from the large correlation value, LC-BLYP overpredicts 

corr \ and the positive 

and B97, respectively, because of the additional dispersion 
give even larger total interaction energies compared to BLYP 


Table 5 shows the GKS-EDA 

Table 6. Electron Density Difference Maps (EDDMs) of the 

the red and blue parts denote the increase and decrease of the electron density arising from the interaction, respectively.

As for GKS-EDA, the positive total interaction by BLYP arises 
from the lack of the dispersion interaction and the relatively 
small correlation term. However, as illustrated by LMO-EDA, 
the dispersion interaction is the largest contribution ranging 
from –4.18 to –4.28 kcal/mol with the various basis sets, even 
larger than the ΔEdisp value. The LMO-EDA result contradicts 
our knowledge of the vdW interaction, whereas the GKS-EDA 
analysis is physical meaningful.

3. C2H4···F2 Complex. Table 5 shows the GKS-EDA analysis for the C2H4···F2 complex with the ACCT basis set. 

The ethylene–halogen complex is important for the mechanism of the addition reaction of halogen molecule to ethylene. 

One of the theoretical viewpoints considers the charge transfer from the π orbital of C2H4 toward the halogen molecule as the main contribution of the stabilization energy, whereas another viewpoint is that the dispersion interaction is more important than the charge transfer effect. As illustrated by Chai et al., most of DFT functionals tend to overestimate the binding energy.

The GKS-EDA analysis illustrates that the DFT performance is mainly attributed to the polarization and correlation terms. The total binding energies vary from –0.54 to –3.46 kcal/mol, compared to the best estimated value of –1.07 kcal/mol by Truhlar et al. First, the local functionals, involving BLYP, BLYP-D, M06-L, and B97-D, overestimate the total interaction energy along with the large correlation and the positive polarization terms. By BLYP, the ΔEcorr is –12.19 kcal/mol whereas the ΔEpol is 6.32 kcal/mol, showing that the polarization energy is a repulsive effect. BLYP-D and B97-D give even larger total interaction energies compared to BLYP and B97, respectively, because of the additional dispersion term. As for the hybrid functionals B97 and B3LYP, the analysis results show the proper total interaction energies with the positive ΔEpol and large ΔEcorr values.

Second, for all the range-separated DFT functionals along with M06-2X and B2PLYP, their analysis results are similar, showing the small negative ΔEpol and the relatively small ΔEcorr values. The polarization interaction is not important, the total interaction is governed by the correlation term, regarded as the typical vdW interaction. For example, the wB97X interaction energy is –1.18 kcal/mol, with ΔEdisp = –2.64 kcal/mol, ΔEpol = –6.92 kcal/mol, ΔEcorr = 12.82 kcal/mol, ΔEpol = –0.63 kcal/mol, and ΔEcorr = –3.55 kcal/mol. All these functionals except LC-BLYP avoid the overestimation of the binding energy. Arising from the large correlation value, LC-BLYP overpredicts the interaction energy. As concluded by Corrinboeuf et al., the large portion of the exact HF exchange energy is helpful to improve the interaction description for the ethylene-halogen complex.

Finally, the behavior of the GKS-EDA polarization interaction can be examined by the electron density difference map (EDDM), which is defined as the difference between the electron density maps of monomers and that of supermolecule. The EDDMs of BLYP, B3LYP, LC-BLYP, and B2PLYP for the C2H4···F2 and the water dimer are shown in Table 6, in which the red and blue parts in the EDDMs denote the increase and decrease of the electron density arising from the interaction, respectively.

Table 6. Electron Density Difference Maps (EDDMs) of the Water Dimer and C2H4···F2 with the ACCT Basis Set

<table>
<thead>
<tr>
<th></th>
<th>DFT</th>
<th>water dimer</th>
<th>C2H4···F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2PLYP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M06-2X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAM-B3LYP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3LYP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLYP</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The red and blue parts in the EDDMs denote the increase and decrease of the electron density arising from the interaction, respectively.

As for the water dimer, it is shown that the EDDMs from the various DFT functionals are similar, according to the fact that their polarization values are close. As for the C2H4···F2 complex, the EDDMs are quite different. The EDDM by BLYP shows the extinct electron transfer from the π orbital of C2H4 to the σu* orbital of F2. It leads to the instability of the energy by the SCF procedure and the positive polarization interaction, which can be compensated by the large correlation energy. The EDDM of B3LYP is analogous to the BLYP one, compared to the similarity of their positive polarization terms, whereas for the functionals providing small negative polarization values, their EDDMs show that the electron density of the F atom that is
close to \(\text{C}_2\text{H}_4\) decreases while that of the other F atom increases. The small variation of electron density illustrate that there is no obvious electron transfer between \(\text{C}_2\text{H}_4\) and \(\text{F}_2\). In general, the polarization term, which is in agreement with the EDDMs quite well, provides a physically meaningful description for the orbital interaction.

In general, with the help of the EDDM analysis, it is found that for the local and some hybrid functionals, the electron transfer is involved in the total interaction, leading to the positive polarization and large correlation values, whereas for nonlocal functionals with a relatively large portion of exact HF exchange energy, the total interaction is regarded as a vDW interaction.

4. Symmetric Radical Cation \(\text{Ar}_2^+\). The GKS-EDA analysis results for the symmetric radical cation \(\text{Ar}_2^+\) with the ACCT basis set are shown in Table 7. The experimental exchange energy, the total interaction is regarded as a vDW that for the local and some hybrid functionals, the electron EDDMs quite well, provides a physically meaningful analysis results for the symmetric radical cation \(\text{Ar}_2^+\).

Finally, for the range-separated functionals except LC-BLYP, attributed by the smaller correlation terms, the total interaction energies are improved. For example, the CAM-B3LYP total interaction energy is \(-39.66\) kcal/mol, smaller than the B3LYP value because of the decreased correlation term. The LC-BLYP correlation value is the largest among these range-separated functionals, resulting in the overestimated binding energy of \(-43.39\) kcal/mol.

5. Active Site of Aurin (PDB 1e5y) in a Solvated Environment. In the active site of the type-1 copper proteins, the Cu cation is coordinated by a cysteine thiolate ligand and two histidine N\(^\text{\delta}\) ligands. An axial coordination is provided by a methionine thioether ligand in plastocyanin, amicyanin, pseudoazurin, rusticyanin, and azurin. In azurin, a second axial coordination is provided by a glycine backbone carbonyl oxygen ligand.\(^{103}\) According to the previous LMO-EDA analysis,\(^{30}\) the \(\text{Cu}^+\text{−S}^\text{−}\) (Cys) is the strongest among these Cu\(^+\)-ligand bonds in the active site, showing the largest degree of covalence. The \(\text{Cu}^+\text{−S}^\text{−}\) (Cys) and the two \(\text{Cu}^+\text{−N}^\text{\delta}\) (His) ligand bonds compose the basic framework of the copper active site. The fourth ligand, methionine, is somewhat far away from the copper cation compared with the cystine and histidine residues. The bond strength of Cu\(^+\) with the cystine and histidine is significantly elongated when the B3LYP is employed for optimization. However, the long axial Cu\(^+\) bond in protein environment has not been analyzed quantitatively.

Finally, for the range-separated functionals except LC-BLYP, attributed by the smaller correlation terms, the total interaction energies are improved. For example, the CAM-B3LYP total interaction energy is \(-39.66\) kcal/mol, smaller than the B3LYP value because of the decreased correlation term. The LC-BLYP correlation value is the largest among these range-separated functionals, resulting in the overestimated binding energy of \(-43.39\) kcal/mol.

### Table 7. GKS-EDA Analysis for \(\text{Ar}_2^+\) with the ACCT Basis Set (kcal/mol)

<table>
<thead>
<tr>
<th>Method</th>
<th>(\Delta E^{\text{ex}})</th>
<th>(\Delta E^{\text{cor}})</th>
<th>(\Delta E^{\text{pol}})</th>
<th>(\Delta E^{\text{disp}})</th>
<th>(\Delta E^{\text{corr}})</th>
<th>(\Delta E^{\text{tot}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLYP</td>
<td>-9.68</td>
<td>-31.03</td>
<td>67.95</td>
<td>-38.71</td>
<td>-35.38</td>
<td>-46.85</td>
</tr>
<tr>
<td>B3LYP</td>
<td>-9.26</td>
<td>-30.07</td>
<td>66.10</td>
<td>-39.40</td>
<td>-29.43</td>
<td>-42.06</td>
</tr>
<tr>
<td>B2PLYP</td>
<td>-8.94</td>
<td>-29.31</td>
<td>64.54</td>
<td>-40.30</td>
<td>-22.89</td>
<td>-36.90</td>
</tr>
<tr>
<td>LC-BLYP</td>
<td>-9.88</td>
<td>-31.73</td>
<td>68.78</td>
<td>-40.75</td>
<td>-29.80</td>
<td>-43.39</td>
</tr>
<tr>
<td>CAM-B3LYP</td>
<td>-9.35</td>
<td>-30.85</td>
<td>64.48</td>
<td>-40.20</td>
<td>-26.23</td>
<td>-39.66</td>
</tr>
<tr>
<td>BLYP-D</td>
<td>-9.68</td>
<td>-31.03</td>
<td>67.95</td>
<td>-38.71</td>
<td>-35.38</td>
<td>-0.12</td>
</tr>
<tr>
<td>M06-L</td>
<td>-7.25</td>
<td>-25.58</td>
<td>57.78</td>
<td>-35.65</td>
<td>-33.18</td>
<td>-43.88</td>
</tr>
<tr>
<td>M06-2X</td>
<td>-8.84</td>
<td>-29.34</td>
<td>64.43</td>
<td>-39.98</td>
<td>-24.11</td>
<td>-37.84</td>
</tr>
<tr>
<td>M11</td>
<td>-8.50</td>
<td>-29.20</td>
<td>64.12</td>
<td>-40.20</td>
<td>-23.24</td>
<td>-37.01</td>
</tr>
<tr>
<td>B97</td>
<td>-8.92</td>
<td>-29.18</td>
<td>64.52</td>
<td>-39.07</td>
<td>-30.05</td>
<td>-42.70</td>
</tr>
<tr>
<td>B97-D</td>
<td>-8.61</td>
<td>-28.44</td>
<td>63.64</td>
<td>-37.90</td>
<td>-33.58</td>
<td>-0.05</td>
</tr>
<tr>
<td>wB97X</td>
<td>-8.57</td>
<td>-28.61</td>
<td>63.30</td>
<td>-39.71</td>
<td>-24.61</td>
<td>-38.19</td>
</tr>
</tbody>
</table>

Finally, for the range-separated functionals except LC-BLYP, attributed by the smaller correlation terms, the total interaction energies are improved. For example, the CAM-B3LYP total interaction energy is \(-39.66\) kcal/mol, smaller than the B3LYP value because of the decreased correlation term. The LC-BLYP correlation value is the largest among these range-separated functionals, resulting in the overestimated binding energy of \(-43.39\) kcal/mol.
The B3LYP interaction energy, 5.87 kcal/mol, is repulsive due to the small $\Delta G^{\text{corr}}$ value. Despite the fact that CAM-B3LYP provides the larger correlation energy than the B3LYP one, it is still unable to predict the attractive binding energy. As for B97-D, M11, wB97-D, and M06-2X functionals, the Met121 residue can serve as the weak ligand binding to the copper center. As for B97-D and wB97X-D, the attractive interaction energies are attributed by the dispersion terms, $-14.73$ and $-14.27$ kcal/mol, respectively. For M11 and M06-2X, their correlation terms are larger than those of B97-D and wB97X-D, compensating the lack of the dispersion interaction. In conclusion, the correlation and dispersion terms are the main factors for this ligand interaction.

E. CONCLUSION
In this paper, a new energy decomposition analysis scheme based on LMO-EDA and EDA-PCM, called GKS-EDA, is presented. Compared to the LMO-EDA scheme, the GKS-EDA scheme shows improvement on the DFT functional adaptability. The GKS-EDA scheme can perform interaction analysis in the gas phase and solvated environments by various DFT functionals, including local, hybrid, double hybrid, range-separated and dispersion correction functionals, whereas the current version of LMO-EDA cannot work with the range-separated and dispersion correction functionals. The GKS-EDA interaction terms are defined according to the generalized Kohn–Sham scheme. The exchange, repulsion, and polarization terms are determined by DFT determinant; the correlation term is attributed by the GKS correlation energy; the dispersion term, which is optional for dispersion correction DFT, defined as the difference of the dispersion corrections from monomers to supermolecule without BSSE correction. With the new definition, the GKS-EDA scheme avoids the error of LMO-EDA arising from the separated treatment of $E_x$ and $E_c$ functionals. With the GKS-EDA scheme, the assessment of a series DFT functionals is performed. Some points are highlighted as follows:

1. In most of the examples, the DFT performance is mainly attributed by the correlation interaction. The range-separated DFT functionals, along with M06-2X, providing relatively large correlation energy, tend to mimic the dispersion energy for the hydrogen bonding and vdW interaction. The B2PLYP performs well for the hydrogen bonding and the charge transfer complex, but not suitable for the vdW interactions due to the inadequate description of the correlation interaction.

2. For the $\text{C}_2\text{H}_4\cdots\text{F}_2$ complex, the tested functionals can be regarded as two groups. By the first group, the interaction belongs to electron transfer; whereas by the second group, the interaction can be regarded as vdW interaction. The positive polarization energy by the first group functionals can be well understood by the EDDM, showing the electron transfer from the $\pi$ bonding orbital of $\text{C}_2\text{H}_4$ to the $\sigma^*$ orbital of $\text{F}_2$.

3. As for the interaction of the active site in the type 1 copper protein, the long weak ligand of the copper cation provided by Met121 belongs to the vdW interaction, which is weakened by the protein environment.

Table 8. GKS-EDA Analysis of the Model Molecule for the Active Site of the Type 1 Copper Protein with the 6-31G(d) Basis Set (Azurin, PDB ID 1e5y, kcal/mol)

<table>
<thead>
<tr>
<th>Functional</th>
<th>$\Delta G^{\text{xc}}$</th>
<th>$\Delta G^{\text{ex}}$</th>
<th>$\Delta G^{\text{disp}}$</th>
<th>$\Delta G^{\text{TOT}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>B97-D</td>
<td>-12.22</td>
<td>-37.18</td>
<td>-1.21</td>
<td>-14.73</td>
</tr>
<tr>
<td>wB97X-D</td>
<td>-11.08</td>
<td>-34.08</td>
<td>-2.88</td>
<td>-32.68</td>
</tr>
<tr>
<td>CAM-B3LYP</td>
<td>-11.76</td>
<td>-36.45</td>
<td>-3.16</td>
<td>-36.45</td>
</tr>
<tr>
<td>M11</td>
<td>-10.69</td>
<td>-31.63</td>
<td>-2.57</td>
<td>-33.13</td>
</tr>
<tr>
<td>B3LYP</td>
<td>-11.07</td>
<td>-36.80</td>
<td>-1.90</td>
<td>-38.77</td>
</tr>
<tr>
<td>M06-2X</td>
<td>-10.84</td>
<td>-33.82</td>
<td>-2.15</td>
<td>-36.05</td>
</tr>
<tr>
<td>B2PLYP</td>
<td>-11.13</td>
<td>-32.68</td>
<td>-3.47</td>
<td>-36.05</td>
</tr>
</tbody>
</table>

Figure 2. Model molecule for the type 1 copper protein (PDB 1e5y). The red and blue parts are the Met121 residue and the remaining part of the active site, respectively.

Note: This project is supported by the Ministry of Science and Technology of China (No. 2011CB808504), the Natural Science Foundation of China (Nos. 21373165, 21273176, 2112010203S), and the Natural Science Foundation of Fujian Province, China (No. 2013J01058).


(87) Nar, H.; Messerschmidt. A. To be Published. DOI: 10.22201/pdb1elsey/pdb. http://www.rcsb.org/pdb/explore.explore.do?structureId=1E5Y.