

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/236983190>

Reaching a Uniform Accuracy for Complex Molecular Systems: Long-Range-Corrected XYG₃ Doubly Hybrid Density Functional

ARTICLE in JOURNAL OF PHYSICAL CHEMISTRY LETTERS · MAY 2013

Impact Factor: 7.46 · DOI: 10.1021/jz400695u

CITATIONS

18

READS

61

2 AUTHORS, INCLUDING:



Igor Ying Zhang

Fritz Haber Institute of the Max Planck Society

34 PUBLICATIONS 747 CITATIONS

SEE PROFILE

Reaching a Uniform Accuracy for Complex Molecular Systems: Long-Range-Corrected XYG3 Doubly Hybrid Density Functional

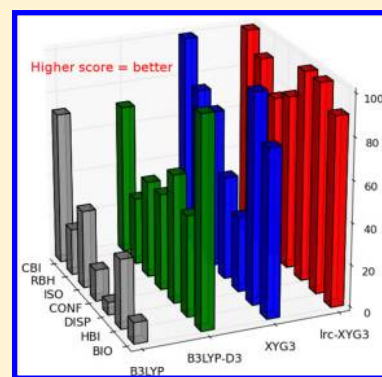
Igor Ying Zhang and Xin Xu*

Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, MOE Laboratory for Computational Physical Science, Collaborative Innovation Center of Chemistry for Energy Materials, Department of Chemistry, Fudan University, Shanghai 200433, China

Supporting Information

ABSTRACT: An unbiased understanding of complex molecular systems from first-principles critically demands theoretical methods with uniform accuracy for diverse interactions with different natures covering short-, medium-, and long-range correlations. Among the state-of-the-art density functional approximations (DFAs), doubly hybrid (DH) DFAs (e.g., XYG3 in this Letter) provide a remarkable improvement over the conventional DFAs (e.g., B3LYP in this Letter). Even though XYG3 works quite well in many cases of noncovalent bonding interactions (NCIs), it is incomplete in describing the pure long-range dispersive interactions. Here, we address such concerns by adding a scaled long-range contribution from the second-order perturbation theory (PT2). The long-range-corrected XYG3 (lrc-XYG3) is proposed without reparameterizing the three parameters in the original XYG3. Due to its overall excellent performance for all testing sets constructed for various purposes, lrc-XYG3 is the recommended method, which is expected to provide a balanced description of diverse interactions in complex molecular systems.

SECTION: Molecular Structure, Quantum Chemistry, and General Theory



The so-called exchange–correlation effect represents a nonclassical many-body effect that plays a fundamental role in forming a variety of chemical and physical interactions including, for example, covalent bond interactions (CBIs), stereoisomerizations (ISOs), hydrogen bond interactions (HBIs), and dispersive interactions (DISPs). These interactions are ubiquitous, and teamwork among them is always essential to exhibit unique features of a given complex system, for which comprehensive understanding critically demands a theoretical method with a uniform accuracy to all kinds of interactions.^{1,2} Development of such a theoretical method is the goal of the present work.

Density functional theory (DFT) provides a rigorous framework to consider the exchange–correlation effect,^{3–5} although, in practice, some density functional approximations (DFAs) have to be employed. There is increasing evidence showing that the popular conventional DFAs, for example, B3LYP,^{6–11} degrade as systems become large,^{12–16} underestimate reaction barrier heights (RBHs),^{15–21} fail in predicting stereoelectronic interactions^{15,21–24} and noncovalent bonding interactions (NCIs),^{15,16,20,21,25–30} and so forth. These unpleasant situations limit their proper applications to the complex systems.

In 2009, we proposed a doubly hybrid (DH) DFA, coined as XYG3,³¹ which contains three parameters in the form as

$$E_{xc}^{XYG3} = c_1 E_x^{HF} + (1 - c_1) E_x^S + c_2 \Delta E_x^{B88} + (1 - c_3) E_c^{LYP} + c_3 E_c^{PT2} \quad (1)$$

DH here suggests that XYG3 not only hybridizes the (semi)local exchange (e.g., E_x^S and ΔE_x^{B88})^{6,8} terms with the Hartree–Fock (HF) exchange E_x^{HF} but also hybridizes the terms from both the semilocal correlation (e.g., E_c^{LYP})⁹ and that based on the second-order perturbation theory (E_c^{PT2})

$$E_c^{PT2} = \sum_{i < j} \sum_{a < b} \frac{|\langle \varphi_i \varphi_j | \hat{v}_{ee} | \varphi_a \varphi_b \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \quad (2)$$

Here, \hat{v}_{ee} is the electron–electron repulsion operator. One salient characteristic, which differs from other DH functionals (see, e.g. refs 21, 26, and 32), is that all energy terms associated with XYG3 are evaluated by using the B3LYP self-consistent (generalized) Kohn–Sham (KS) orbitals $\{\varphi_{i,a}\}$ with eigenvalues $\{\epsilon_{i,a}\}$, where the subscripts (i,j) and (a,b) denote the occupied and unoccupied KS orbitals, respectively. The well-established B3LYP functional provides a good approximation to the real (yet unknown) KS functional to construct the zeroth-order Hamiltonian upon which the XYG3 functional relies.^{15,16,31,33} Parameters of $\{c_1 = 0.8033, c_2 = 0.2107, c_3 = 0.3211\}$ in eq 1 were determined empirically by fitting only to the data of the heats of formation (HOF) in the G3/99 set³⁴ but were tested independently with different systems including RBHs and NCIs.^{15,16,21,30,31}

Received: March 30, 2013

Accepted: April 30, 2013

Published: April 30, 2013



XYG3 cures, to a large extent, several aforementioned problems, being remarkably accurate for a wide range of systems and important chemical properties.^{15,16,21,30,31,35–39} As it contains about 80% HF exchange with its correlation made of LYP and PT2, all of which are free from the one-electron self-interaction error (SIE),^{40–44} XYG3 shows excellent performance on the calculations of reaction barrier heights,^{15,16,21,30,31,39} as well as other SIE-related problems (e.g., the SIE11 set²¹). Errors associated with conventional DFAs in the medium-range correlation are found to be largely responsible for the significant molecular size dependence problem,^{21,22,25–27} as well as their ineptitude in predicting the relative isomerization energies.^{21–24} By introducing a compatible portion of delocalized PT2 correlation, XYG3 presents a notable improvement for this issue^{15,16} (see also below).

Common DFAs miss the R^{-6} decay behavior in the long-range correlation^{1,2} and hence fail badly for dispersion-dominant NCIs.^{16,20,21,25–30} The delocalized E_c^{PT2} captures the correct long-range behavior by construction. It is, therefore, unsurprising to see an overall improvement of XYG3 for the description of many NCI systems.^{15,16,30,31} On the other hand, 32% global E_c^{PT2} in XYG3 is incomplete to embrace the full correlation effect in the long range. A tendency for XYG3 to still underestimate the pure DISPs has been observed, for example, in the dimers of alkanes²¹ (C_nH_{2n+2})₂, which could sum up to a significant nonpolar effect in, for example, protein folding.

Other types of DH DFAs were found to suffer more seriously from the incomplete treatment of the long-range DISPs.^{21,26,27} It was therefore suggested to add to the functionals an a posteriori force field (FF)-like dispersion correction (e.g., -D or -D3)^{26,27} to impose the correct long-range R^{-6} interatomic dependence. More recently, the DFT-D or DFT-D3 method was made an integrated method where the parameters in the corresponding DFA (i.e., electronic) part were optimized in conjunction with the optimization of the parameters in the FF-like dispersion correction terms, with a hope that the medium-to long-range correlations were completely taken over by the classic FF terms and not mimicked by the DFA part to avoid double-counting.²¹ While such a DFT-D or -D3 scheme can be implemented efficiently without additional computational cost and was found to be quite accurate, it includes many empirical parameters and suffers from some inherent limitations.^{21,26,27} Especially, the many-body correlation effects and anisotropy effects in the long-range DISPs, as well as the orbital dependence in the medium range, are more subtle, which are difficult to approximate in the pairwise additive FF models.^{25–27,45–47} We believe that the classic FF terms are not the optimal partner to the DH DFAs where the physically sound PT2 term has already been enfolded.

Range-separated schemes provide a particularly powerful route in combining the strengths of (semi)local DFAs and the nonlocal orbital-dependent terms from first-principles.^{32,47–50} In this Letter, we introduce a portion of long-range PT2 correction upon XYG3

$$E_c^{lrPT2} = \sum_{i < j} \sum_{a < b} \frac{|\langle \varphi_i \varphi_j | \hat{v}_{ee}^{lr} | \varphi_a \varphi_b \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \quad (3)$$

where $\hat{v}_{ee}^{lr}(r_{12}) = \text{erf}(\omega r_{12})/r_{12}$ is the long-range electron–electron repulsion operator. The adjustment function is the most widely used Gauss error function, where the parameter ω

determines the distance at which the asymptotic limit $\hat{v}_{ee}^{lr} \rightarrow \hat{v}_{ee}$ is reached.

Combining the correction term (eq 3) with XYG3 (eq 1), we propose the long-range-corrected XYG3, namely, lrc-XYG3, as

$$E_{xc}^{lrc-XYG3} = c_1 E_x^{HF} + (1 - c_1) E_x^{LDA} + c_2 \Delta E_x^{B88} + (1 - c_3) E_c^{LYP} + c_3 E_c^{PT2} + (c_1 - c_3) E_c^{lrPT2} \quad (4)$$

The scaling factor of E_c^{lrPT2} is predefined to be $c_1 - c_3 = 0.4822$ to avoid introducing a new mixing parameter to the functional. Hence, about 80% delocalized PT2 correlation is restored, which integrates with the same amount of delocalized HF exchange in the long range. This is justified by the fact that a full portion of the HF exchange plus the PT2 correlation usually overestimates the C_6 coefficients.^{1,26,27,49} Analogous to that in B2-P3LYP,⁴⁷ $\omega = 0.2 \text{ Bohr}^{-1}$ is found to be appropriate for lrc-XYG3, which turns on the full portion of the electron–electron interaction around 10 Bohr (or $\sim 5 \text{ \AA}$). A smaller value of ω (e.g., 0.1 Bohr^{-1}) leads to marginal improvement of lrc-XYG3 over XYG3, whereas a larger value of ω (e.g., 0.3 Bohr^{-1} or higher) results in undesirable double-counting of the PT2 correlation contributions in the short and middle ranges, deteriorating the overall performance of lrc-XYG3.

In comparison, only the exchange part was treated by the range-separated hybrids in the long-range-corrected DH functional $\omega B97X-2$ ³² of Chai and Head-Gordon. Even though its correlation part is treated by global hybrids, $\omega B97X-2$ has been shown to display excellent general performance, including the S22⁵¹ set for NCIs. A series of range-separated hybrid DFAs have been put forward recently that include the standard PT2 or random-phase approximation (RPA).^{49,50} These schemes make a range-dependent admixture in both exchange and correlation terms. Benighaus et al. proposed a range-dependent DH (B2-P3LYP),⁴⁷ featured by using a modified two-electron operator $\hat{g}_{ee} = \hat{v}_{ee} + c \hat{v}_{ee}^{lr}$ to replace the standard (\hat{v}_{ee}) operator in evaluation of the PT2 energy. Such an operator, as a sum of \hat{v}_{ee} and \hat{v}_{ee}^{lr} , results in a “modified” PT2 correlation,^{47,48} which, however, contains a piece of contribution from the coupling term between \hat{v}_{ee} and \hat{v}_{ee}^{lr} . Distinct from B2-P3LYP,⁴⁷ the present lrc-XYG3 method rectifies the global XYG3 method by adding only a portion of the distance-dependent E_c^{lrPT2} contribution without invoking the coupling term between \hat{v}_{ee} and \hat{v}_{ee}^{lr} . Hence, eq 4 can be reformulated in a way just as in the DFT-D method^{26,27}

$$E_{xc}^{lrc-XYG3} = E_{xc}^{XYG3} + s_{XYG3} E_c^{lrPT2} \quad (5)$$

where $E_{\text{disp}} = s_{\text{DFA}} E_c^{lrPT2}$, with $s_{XYG3} = 0.4822$. The delocalized E_c^{lrPT2} term is not only distance-dependent and atom-dependent but also orbital-specific, which goes beyond the classic FF model by construction and is of electronic origin at the same level of complexity as the original DH functionals.

Goerigk and Grimme have carried out a thorough evaluation on the performance of DH functionals against their extended GMTKN30 database for general main group thermochemistry, kinetics, and NCIs.²¹ It was concluded that XYG3, along with all D3-corrected DH DFAs, has achieved an accuracy that is superior to the B3LYP-D3 reference, whereas the only exception is NCI, for which B3LYP-D3 is significantly better than XYG3 for data sets such as S22, ADIM6, RG6, HEAVY28, PCONE, and ACONF.²¹ These results are reproduced in Table 1 (see the Supporting Information and ref 21 for more results). It was also argued that D3-corrected XYG3 cannot be made possible without reoptimizing XYG3 because double-counting

Table 1. MADs (in kcal/mol) for B3LYP, B3LYP-D3, XYG3, and lrc-XYG3 against Various Testing Sets

set ^a	subset ^b	av. ΔE ^c	B3LYP	B3LYP-D3	XYG3	lrc-XYG3
CBI	G2RC	50.6	2.63	2.70	1.85	1.85
	BH76RC	21.5	2.28	2.26	1.21	1.22
RBH	BH76	18.5	4.68	5.20	1.14	1.17
	BHPERI	19.4	5.70	2.79	1.87	1.35
ISO	ISO34	14.3	2.29	1.87	1.14	1.12
CONF	SCONF	4.9	1.00	0.45	0.34	0.24
	CYCONF	2.1	0.45	0.24	0.16	0.12
	ACONF	1.8	0.96	0.14	0.20	0.11
	PCONF	1.5	3.93	0.34	0.51	0.13
DISP	IDISP	13.5	17.36	6.63	3.15	1.16
	ADIM6	3.3	4.96	0.43	1.08	0.27
	RG6	0.46	0.79	0.07	0.21	0.08
	HEAVY28	1.3	1.33	0.16	0.27	0.10
HBI	WATER27	82.0	6.18	4.24	1.44	1.95
BIO	S22	7.3	3.72	0.36	0.46	0.41

^aCodes for the set names. CBI: covalent bond interactions; RBH: reaction barrier heights; ISO: isomerization energies; CONF: relative conformational energies; DISP: dispersive interactions; HBI: hydrogen bond interactions; BIO: biologically oriented complexes. ^bAll subsets are taken from the GMTKN30 database.²¹ Geometries and reference values are taken from the Web site.²¹ All calculations are performed at the same level of basis set def2-QZVP,⁵² using a development version of Q-Chem 4.0.⁵³ ^cAveraged absolute energy.

effects cannot be avoided.²¹ As the intent of the present work is to resolve these problems with little impairment of the otherwise excellent performance of XYG3 for other properties, the assessment of lrc-XYG3, based on mean absolute deviation (MAD in kcal/mol), against the aforementioned data sets, as well as others, has been performed (see Figure 1 and Table 1).

S22 is a well-established data set,⁵¹ which consists of 22 NCI systems of biological importance, ranging in size from the water dimer to DNA base pairs with interaction energies varying from 0.5 kcal/mol for a dispersion-predominant complex to 21 kcal/mol for a hydrogen-bonded complex. This set is labeled here as BIO in Table 1. S22 is now frequently used not only in evaluation but also in development of new theoretical methods.^{21,26,27,32,47} S22 is also one of the fitting sets in determining the -D3 parameters.²⁷ MAD achieved by B3LYP-D3 is 0.36 kcal/mol.²¹ XYG3 has a MAD of 0.46 kcal/mol, showing a clear underbinding tendency at the present basis set of def2-QZVP⁵² as was used in ref 21. Previously, we found that MAD for XYG3 with Pople's 6-311+G(3df,2p) is only 0.19 kcal/mol.³⁰ Although the original parameters in XYG3 were optimized with this Pople basis set, the fitting set was only HOFs of the G3/99 data set.^{31,34} Further investigation showed that, while the calculations of HOFs are subject to the largest basis set dependence such that the intended basis set is recommended, the basis set dependence is relatively weaker for other applications,⁵⁴ although the Pople basis sets are found to be more compatible with XYG3 than other basis sets in describing NCIs.^{30,55,56} The lrc-XYG3 functional improves XYG3 marginally in terms of MAD, being 0.41 kcal/mol. However, it works in the right direction, albeit being slightly overbinding. Here, one might tune the s_{XYG3} factor against S22 to cut MAD down to 0.16 kcal/mol; this procedure is reserved for future systematic study. We prefer to keep eq 4 with an intention of fewer fits. In fact, a counterpoise correction or a basis set extrapolation will remove the residual basis set

incompleteness, bringing the lrc-XYG3 results into a better agreement with the reference data.

The WATER27 subset⁵⁷ is utilized to gauge more specifically the functional performance for HBIs. It includes 27 binding energies of $(\text{H}_2\text{O})_n$, $\text{H}^+(\text{H}_2\text{O})_n$, and $\text{OH}^-(\text{H}_2\text{O})_n$ clusters where the interaction energies range from 5 to 218 kcal/mol. Both the medium-range and the long-range correlations are believed to play an important role.^{26,27} B3LYP-D3 reduces the MAD of 6.18 kcal/mol for B3LYP only to 4.24 kcal/mol, which is insufficient due to the incomplete treatment of the orbital-sensitive medium-range correlation. XYG3 is excellent (MAD = 1.44 kcal/mol), while lrc-XYG3 is also satisfactory (MAD = 1.95 kcal/mol) if one considers the large average binding energy (av. | ΔE |) of 82.0 kcal/mol for this subset.²¹

The set (dubbed DISP) is employed to examine the functional performance for DISPs.^{21,22,25–27} It includes four subsets, namely, IDISP of 6 large organic systems for intramolecular interactions, ADIM6 of 6 *n*-alkane dimers from ethane to heptane for intermolecular interactions, RG6 of 5 homonuclear rare gas dimers from Ne to Rn plus Xe–Rn, and HEAVY28 involving 28 NCI complexes between hydride molecules with heavier elements such as Bi, Pb, Sb, Te, Br, I, and so forth. As seen from Table 1, B3LYP fails completely, producing errors that exceed the corresponding av. | ΔE |. For ADIM6 and RG6, which are of purely dispersive nature, B3LYP-D3 is more satisfactory, reducing the error by 1 order of magnitude compare to that of B3LYP. The DISPs in the HEAVY28 subset are more complicated, messed up by other types of weak interactions such as dipole interactions or charge-transfer effects. B3LYP-D3 works also well in such a situation (MAD = 0.16 kcal/mol).²¹ Note that ADIM6 and RG6 are among the fitting sets in determining the -D3 parameters, while hydrides in HEAVY28 are the reference molecules for the determination of the dispersion coefficients.²⁷ Perusing Table 1 for these three intermolecular systems with predominant dispersive character, one finds that XYG3 works much better than B3LYP,²¹ but XYG3 still misses part of the relevant interactions. The lrc term is very effective, rendering lrc-XYG3 one of the best DH functionals on the market for purely DISPs (see the Supporting Information and ref 21 for more results).

In the subset of IDISP, both the short- and medium-range correlations may play an important role along with the long-range correlation (see also the Supporting Information).^{22,25–27} For example, the photodimerization of anthracenes is a reversible [4+4] cycloaddition, which yields a covalently bound polycyclic dimer.²⁵ Here, the short-range correlation is vital for the description of $\pi \rightarrow \sigma$ bond transfer, while the long-range correlation is crucial for the description of intramolecular dispersion in the dimer product. In the case of hydrogenation of a cyclophane, the short-range correlation due to bond breaking and making is also obviously involved. Meanwhile, the distances between the two phenyl planes in cyclophane are below the van der Waals radius of carbon, which shows that this is not a typical dispersion problem. The medium-range correlation is, therefore, imperative,²⁶ which is related to the involved orbitals in a very system-dependent manner. The IDISP subset also includes branching isomerizations and linear to singly folded transformation of alkane chains.^{21,22,26} While in the former cases the medium- and long-range correlation contributions are both critical, in the latter, the intramolecular DISPs are becoming more and more decisive with the elongation of the carbon chain length. All in all, B3LYP fails, having a MAD of 17.4 kcal/mol for IDISP, exceeding the av.

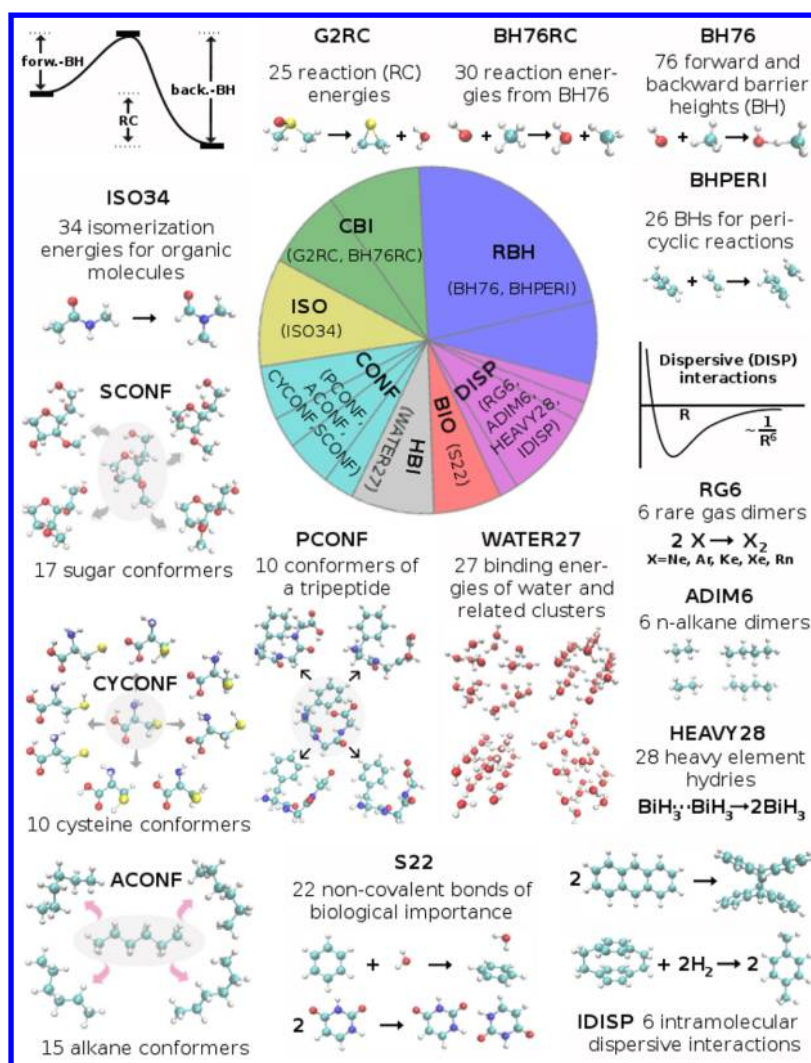


Figure 1. Representatives in the benchmarking subsets.²¹ The choice is made in regard to the robustness and usefulness of the method for complex molecular systems of main group elements.

$|\Delta E|$ by ~ 4 kcal/mol. B3LYP-D3 removes the B3LYP MAD by 62%, which is still far from being satisfactory. XYG3 provides remarkable progress over B3LYP, removing its MAD by 82% but still suffering from the imperfection in the long-range correlation.²¹ The lrc-XYG3 functional is shown here to display a balanced description for all short-, medium-, and long-range correlations, with MAD being only 1.16 kcal/mol for the entire IDISP subset.

For the CONF, also called stereo ISO, relative energies of 10 phenylalanyl-glycyl-glycine tripeptide conformers (PCONF), 15 alkane conformers (ACONF), 17 sugar conformers (SCONF), and 10 cysteine conformers (CYCONF) are taken into account^{21,27} and grouped together as the CONF set in Table 1. For structural isomerization, the ISO34 subset is employed, which encompasses 34 reaction energies where structural isomers undergo skeletal, position, and functional group isomerizations.^{21,24} Note that all isomerization reactions in ISO are accompanied by considerable change in electronic structure involving bond breaking and making, as well as the bond connectivity rearrangements, and a balanced description between short- and medium-range correlations is essential, while in the CONF set, the stereoisomers differ in the three-dimensional orientations of their atoms in space, resulting in

subtle changes in medium- and long-range correlations. All four subsets in CONF are in the fitting set to optimize the -D3 parameters.²⁷ In fact, B3LYP-D3 represents a substantial improvement over B3LYP. XYG3, although with neither specific fitting to a NCI set nor using a dispersion correction term, already displays an impressive accuracy (see Table 1). It is further improved by augmenting the lrc term, leading to errors that are close to those of the coupled-cluster reference values for all four subsets in CONF.²¹ For ISO, B3LYP gives a MAD of 2.29 kcal/mol, as compared to the av. $|\Delta E|$ of 14.3 kcal/mol. All DHs have been shown to cut the B3LYP MAD in more than half, while the -D3 scheme improves marginally.²¹ This observation is confirmed in this work: XYG3 gives MAD = 1.14 kcal/mol, while the MAD for lrc-XYG3 is 1.12 kcal/mol, reinforcing the argument that DHs provide a better description of short- and medium-range correlation effects encountered in ISO.^{15,16,21,26,27}

To further validate the performance of lrc-XYG3 for chemical reactions, two test sets are utilized. One is the RBH set, which comprises Truhlar's well-known BH76²⁰ set that includes 76 data points for the substitution, association, and unimolecular and atom transfer reactions, as well as Grimme's NHPERI²¹ that includes 26 barrier heights of pericyclic reactions. The

other one is dubbed the CBI set, which comprises 25 reaction energies (G2RC) of the selected G2/97 systems,^{21,34} as well as 30 reaction energies (BH76RC) associated with the BH76 set.²⁰ As shown in Table 1, the already high accuracy of XYG3 (MAD = 1.14 kcal/mol) for BH76 is well-preserved in lrc-XYG3 (MAD = 1.17 kcal/mol). For BHPERI, where larger systems are involved,²¹ MAD is reduced from 1.87 kcal/mol for XYG3 to 1.35 kcal/mol for lrc-XYG3. This improvement is in line with the fact that dispersion effects become more important with increasing system size where the lrc scheme provides a suitable therapy, as is also observed from B3LYP (MAD = 5.70 kcal/mol) to B3LYP-D3 (MADs = 2.79 kcal/mol). On the other hand, the two reaction energy subsets grouped in CBI measure mainly the covalent bond interactions. They are expected to be least affected by the introduction of the dispersion terms, which are indeed the cases as shown in Table 1.

In order to judge the overall performance of DFAs for various properties, we adopt a four-point grading system, where A = 4, B = 3, C = 2, D = 1, and F = 0. The genuine score in percentage for each set is defined as

$$\text{Score} = \frac{1}{N} \sum_i^N \frac{\text{Acc}_i}{\text{MAD}_i} \times 100 \quad (6)$$

where N is the number of subsets and Acc_i is the computational accuracy required (or expected) for the i th subset. For CBI, CONF, and BIO, Acc_i is defined as 5% of the av. $|\Delta E|$ of the i th subset (listed in the fourth column of Table 1). For HBI (WATER27) with large av. $|\Delta E|$, a tighter threshold is adopted (i.e., halved to 2.5% of the av. $|\Delta E|$), which, however, is relaxed to 6% for RBH and ISO and 12% for DISP. A higher score, of course, means better performance. If the expected accuracy is reached or surpassed, a maximum score of 100 is granted.

Figure 2 presents the grade report cards for B3LYP, B3LYP-D3, XYG3, and lrc-XYG3 in different subjects (sets). As usual, one earns an A for a 90–100 score, a B for 80–89, a C for 70–79, and a D for 60–69. An F is the failing grade for a score between 0 and 59. Figure 2 reveals that B3LYP is fairly satisfactory for CBI, being a competent method in describing many cases of chemical reactions. It, however, fails miserably in all other subjects with a grade-point-average (GPA) of only 0.3. Admittedly, CBI, along with RBH and ISO, where short-range correlation effects are crucial, contributes only 5 credits out of the total of 15, whereas CONF, DISP, HBI, and BIO contribute the other 10. This GPA just reflects again the lack of capability for B3LYP to properly describe a complex system where medium- and long-range correlation effects are decisive.^{15,16,21–30} The -D3 scheme helps a lot in DISP and BIO, raising its GPA to 1.1, while it is insufficient for the medium-range correlations where the orbital contributions are more involved.^{26,27} On the other hand, XYG3 is good (or even excellent) at CBI, RBH, and HBI, being fair for ISO and BIO, but it still suffers from CONF and DISP. It merely obtains a 1.5 GPA in this school where CONF and DISP (8 credits out of 15) are compulsory. The advantage inherent in the long-range PT2 term is evident in the description of the challenging NCIs, which almost perfectly heals XYG3, and boosts the GPA up to 3.5.

Both the exchange and correlation energies have a short- and long-range part in terms of the distance between two electrons.^{26,27,32,47–50} In many important applications, the medium-range effect, by itself, is also significant.^{26,27} In the

	B3LYP	B3LYP-D3	XYG3	lrc-XYG3
CBI(2)	C	C	A	A
RBH(2)	F	F	B	A
ISO(1)	F	F	C	C
CONF(4)	F	F	F	B
DISP(4)	F	C	F	A
HBI(1)	F	F	A	A
BIO(1)	F	A	C	B
GPA	0.3	1.1	1.5	3.5

	B3LYP	B3LYP-D3	XYG3	lrc-XYG3
CBI(2)	C	C	A	A
RBH(2)	F	F	B	A
ISO(1)	F	F	C	C
CONF(4)	F	F	F	B
DISP(4)	F	C	F	A
HBI(1)	F	F	A	A
BIO(1)	F	A	C	B
GPA	0.3	1.1	1.5	3.5

Figure 2. Grade report cards for B3LYP, B3LYP-D3, XYG3, and lrc-XYG3 in describing covalent bond interactions (CBIs), reaction barrier heights (RBHs), isomerization energies (ISOs), relative conformational energies (CONFs), dispersive interactions (DISPs), hydrogen bond interactions (HBIs), and biologically oriented complexes (BIO). The number in parentheses corresponds to the credit granted to the subject. A grade-point-average (GPA) is calculated on the 4.0 scale.

current implementations, many common DFAs are mostly local (short-ranged), depending only on the density at the given point and the immediate vicinity via derivatives of the density. Among the state-of-the-art DFAs, the DH functionals (e.g., XYG3 in this Letter) provide a remarkable improvement with the introduction of nonlocality not only through the HF exchange but also through the PT2 correlation.^{15,16,21,26,31,32,47}

In this work, we have shown that by adding a distance-dependent PT2 correlation, the long-standing problem that many DHs are insufficient for long-range DISPs can be cured without relying on the classic FF-like dispersion corrections. This has made the dispersion correction of electronic origin at the same level of complexity as the original DH functionals. We have shown here that lrc-XYG3, with fewer fits, can perform equally well for CBIs, RBHs, CONFs, HBIs and DISPs. On the basis of these results, we state that lrc-XYG3 is competitive in providing an unbiased comprehension of complex molecular systems, being robust and useful for a wide range of chemical, biological, and materials systems.

■ ASSOCIATED CONTENT

● Supporting Information

Table S1, the IDISP subset; Table S2, mean absolute deviations (MADs, in kcal/mol) for B3LYP, B3LYP-D3, XYG3, and lrc-XYG3 against the IDISP subset; Table S3, MADs (in kcal/mol) for B2PLYP, B2PLYP-D3, PWPB95, PWPB95-D3, XYG3, and Table S4, lrc-XYG3 against various testing sets; and scores and GPAs achieved by various methods for different subjects. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: xxchem@fudan.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the NSFC (Grants 21133004, 91027044) and the Ministry of Science and Technology (Grants 2013CB834606, 2011CB808505). The authors acknowledge helpful discussion with Dr. Yong Xu and Dr. Li Rao.

■ REFERENCES

- (1) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories. *Chem. Rev.* **2010**, *110*, 5023–5063.
- (2) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. T. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289–320.
- (3) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (4) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1964**, *136*, A1133–A1138.
- (5) Levy, M. Universal Variational Functionals of Electron Densities, 1st-Order Density Matrices, and Natural Spin–Orbitals and Solution of the V-Representability Problem. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6062–6065.
- (6) Slater, J. C. *The Self-Consistent Field for Molecular and Solids, Quantum Theory of Molecular and Solids*; McGraw-Hill: New York, 1974, Vol. 4.
- (7) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (8) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (9) 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* **1988**, *37*, 785–789.
- (10) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- (11) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (12) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-3 and Density Functional Theories for a Larger Experimental Test Set. *J. Chem. Phys.* **2000**, *112*, 7374–7383.
- (13) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. Assessment of Gaussian-3 and Density Functional Theories for Enthalpies of Formation of C₁–C₁₆ Alkanes. *J. Phys. Chem.* **2000**, *104*, 5850–5854.
- (14) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. Systematic Errors in Computed Alkane Energies Using B3LYP and Other Popular DFT Functionals. *Org. Lett.* **2006**, *8*, 3631–3634.
- (15) Zhang, I. Y.; Wu, J. M.; Xu, X. Extending the Reliability and Applicability of B3LYP. *Chem. Commun.* **2010**, *46*, 3057–3070.
- (16) Zhang, I. Y.; Xu, X. Doubly Hybrid Density Functional for Accurate Description of Thermochemistry, Thermochemical Kinetics and Nonbonded Interactions. *Int. Rev. Phys. Chem.* **2011**, *30*, 115–160.
- (17) Jursic, B. S. The Hydrogen Abstraction from Ethane with a Hydrogen Radical as Examined with Ab Initio and Density Functional Theory Methods. *J. Mol. Struct.: THEOCHEM* **1998**, *428*, 49–54.
- (18) Kobayashi, Y.; Kamiya, M.; Hirao, K. The Hydrogen Abstraction Reactions: A Multireference Møller–Plesset Perturbation (MRMP) Theory Study. *Chem. Phys. Lett.* **2000**, *319*, 695–700.
- (19) Lynch, B. J.; Truhlar, D. G. How Well Can Hybrid Density Functional Methods Predict Transition State Geometries and Barrier Heights. *J. Phys. Chem. A* **2001**, *105*, 2936–2941.
- (20) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
- (21) (a) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals—Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309. See also: (b) <http://www.unimuenster.de/Chemie.oc/grimme/en/index.html> (accessed December 2010).
- (22) Grimme, S. Seemingly Simple Stereoelectronic Effects in Alkane Isomers and the Implications for Kohn–Sham Density Functional Theory. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460–4464.
- (23) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A.; De Meijere, A. Many Density Functional Theory Approaches Fail to Give Reliable Large Hydrocarbon Isomer Energy Differences. *Org. Lett.* **2006**, *8*, 3635–3638.
- (24) Grimme, S.; Steinmetz, M.; Korth, M. How to Compute Isomerization Energies of Organic Molecules with Quantum Chemical Methods. *J. Org. Chem.* **2007**, *72*, 2118–2126.
- (25) Grimme, S.; Diedrich, C.; Korth, M. The Importance of Inter- and Intramolecular van der Waals Interactions in Organic Reactions: the Dimerization of Anthracene Revisited. *Angew. Chem., Int. Ed.* **2006**, *45*, 625–629.
- (26) Schwabe, T.; Grimme, S. Double-Hybrid Density Functionals with Long-Range Dispersion Corrections: Higher Accuracy and Extended Applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (27) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (28) Kristyan, S.; Pulay, P. Can (Semi)Local Density Functional Theory Account for the London Dispersion Forces? *Chem. Phys. Lett.* **1994**, *229*, 175–180.
- (29) Hobza, P.; Šponer, J.; Reschel, T. Density Functional Theory and Molecular Clusters. *J. Comput. Chem.* **1995**, *16*, 1315–1325.
- (30) Zhang, I. Y.; Xu, X. XYG3 and XYGJ-OS Performances for Noncovalent Binding Energies Relevant to Biomolecular Structures. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12554–12570.
- (31) Zhang, Y.; Xu, X.; Goddard, W. A. Doubly Hybrid Density Functional for Accurate Descriptions of Nonbond Interactions, Thermochemistry, and Thermochemical Kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4963–4968.
- (32) Chai, J. D.; Head-Gordon, M. Long-Range Corrected Double-Hybrid Density Functionals. *J. Chem. Phys.* **2009**, *131*, 174105.
- (33) Görling, A.; Levy, M. Correlation-Energy Functional and Its High-Density Limit Obtained from a Coupling-Constant Perturbation Expansion. *Phys. Rev. B* **1993**, *47*, 13105–13113.
- (34) Curtiss, L. A.; Raghavachari, A.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.
- (35) Zhang, I. Y.; Wu, J. M.; Luo, Y.; Xu, X. Accurate Bond Dissociation Enthalpies by Using Doubly Hybrid XYG3 Functional. *J. Comput. Chem.* **2011**, *32*, 1824–1838.
- (36) Liu, G.; Wu, J. M.; Zhang, I. Y.; Chen, Z. N.; Li, Y. W.; Xu, X. Theoretical Studies on Thermochemistry for Conversion of 5-Cholomethylfurfural into Valuable Chemicals. *J. Phys. Chem. A* **2011**, *115*, 13628–13641.
- (37) Zhang, I. Y.; Wu, J. M.; Luo, Y.; Xu, X. Trends in R–X Bond Dissociation Energies (R• = Me, Et, *i*-Pr, *t*-Bu, X• = H, Me, Cl, OH). *J. Chem. Theory Comput.* **2010**, *6*, 1462–1469.
- (38) Zhang, I. Y.; Xu, X. Gas-Phase Thermodynamics as a Validation of Computational Catalysis on Surfaces: A Case Study of Fischer–Tropsch Synthesis. *ChemPhysChem* **2012**, *13*, 1486–1494.

- (39) Shen, C.; Zhang, I. Y.; Fu, G.; Xu, X. Pyrolysis of D-Glucose to Acrolein. *Chin. J. Chem. Phys.* **2011**, *24*, 249–252.
- (40) Perdew, J. P.; Zunger, A. Self-Interaction Correction to Density-Functional Approximations for Many-Electron Systems. *Phys. Rev. B* **1981**, *23*, 5048–5079.
- (41) Merkle, R.; Savin, A.; Preuss, H. Singly Ionized First-Row Dimers and Hydrides Calculated with the Fully Numerical Density-Functional Program NUMOL. *J. Chem. Phys.* **1992**, *97*, 9216–9221.
- (42) Zhang, Y.; Yang, W. T. A Challenge for Density Functionals: Self-Interaction Error Increases for Systems with a Noninteger Number of Electrons. *J. Chem. Phys.* **1998**, *109*, 2604–2608.
- (43) Gräfenstein, J.; Kraka, E.; Cremer, D. The Impact of the Self-Interaction Error on the Density Functional Theory Description of Dissociating Radical Cations: Ionic and Covalent Dissociation Limits. *J. Chem. Phys.* **2004**, *120*, 524–539.
- (44) Ciofini, I.; Adamo, C.; Chermette, H. Self-Interaction Error in Density Functional Theory: A Mean-Field Correction for Molecules and Large Systems. *Chem. Phys.* **2005**, *309*, 67–76.
- (45) Klimeš, J.; Michaelides, A. Perspective: Advances and Challenges in Treating van der Waals Dispersion Forces in Density Functional Theory. *J. Chem. Phys.* **2012**, *137*, 120901.
- (46) Tkatchenko, A.; Rossi, M.; Blum, V.; Ireta, J.; Scheffler, M. Unraveling the Stability of Polypeptide Helices: Critical Role of van der Waals Interactions. *Phys. Rev. Lett.* **2011**, *106*, 118102.
- (47) Benighaus, T.; DiStasio, R. A.; Lochan, R. C.; Chai, J. D.; Head-Gordon, M. Semiempirical Double-Hybrid Density Functional with Improved Description of Long-Range Correlation. *J. Phys. Chem. A* **2008**, *112*, 2702–2712.
- (48) Lochan, R.; Jung, Y.; Head-Gordon, M. Scaled Opposite Spin Second Order Møller–Plesset Theory with Improved Physical Description of Long-Range Dispersion Interactions. *J. Phys. Chem. A* **2005**, *109*, 7598–7605.
- (49) Ángyán, J.; Gerber, I.; Savin, A.; Toulouse, J. van der Waals Forces in Density Functional Theory: Perturbational Long-Range Electron-Interaction Corrections. *Phys. Rev. A* **2005**, *72*, 012510.
- (50) Janesko, B. G.; Henderson, T. M.; Scuseria, G. E. Long-Range-Corrected Hybrids Including Random Phase Approximation Correlation. *J. Chem. Phys.* **2009**, *130*, 081105.
- (51) Jurečka, P.; Šponer, J.; Cerný, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes, DNA Base Pairs, and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- (52) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (53) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; et al. M. Advances in Methods and Algorithms in a Modern Quantum Chemistry Program Package. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.
- (54) Zhang, I. Y.; Luo, Y.; Xu, X. Basis Set Dependence of the Doubly Hybrid XYG3 Functional. *J. Chem. Phys.* **2010**, *133*, 104105.
- (55) Burns, L. A.; Vázquez-Mayagoitia, A. V.; Sumpter, B. G.; Sherrill, C. D. Density-Functional Approaches to Noncovalent Interactions: A Comparison of Dispersion Corrections (DFT-D), Exchange–Hole Dipole Moment (XDM) Theory, and Specialized Functionals. *J. Chem. Phys.* **2011**, *134*, 084107.
- (56) Vázquez-Mayagoitia, Á.; Sherrill, C. D.; Aprà, E.; Sumpter, B. G. An Assessment of Density Functional Methods for Potential Energy Curves of Nonbonded Interactions: The XYG3 and B97-D Approximations. *J. Chem. Theory Comput.* **2010**, *6*, 727–734.
- (57) Bryantsev, V. S.; Diallo, M. S.; Van Duin, A. C. T.; Goddard, W. A. Evaluation of B3LYP, X3LYP, and M06-Class Density Functionals for Predicting the Binding Energies of Neutral, Protonated, and Deprotonated Water Clusters. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.