Accurate Treatment of Large Supramolecular Complexes by Double-Hybrid Density Functionals Coupled with Nonlocal van der Waals Corrections

Joaquín Calbo,[†] Enrique Ortí,[†] Juan C. Sancho-García,[§] and Juan Aragó*[,†,‡]

[†]Instituto de Ciencia Molecular, Universidad de Valencia, E-46980 Valencia, Spain
[§]Departamento de Química Física, Universidad de Alicante, E-03080 Alicante, Spain

Ⓢ *Supporting Information*

ABSTRACT: In this work, we present a thorough assessment of the performance of some representative double-hybrid density functionals (revPBE0-DH-NL and B2PLYP-NL) as well as their parent hybrid and GGA counterparts, in combination with the most modern version of the nonlocal (NL) van der Waals correction to describe very large weakly interacting molecular systems dominated by noncovalent interactions. Prior to the assessment, an accurate and homogeneous set of reference interaction energies was computed for the supramolecular complexes constituting the L7 and S12L data sets by using the novel, precise, and efficient DLPNO-CCSD(T) method at the complete basis set limit (CBS). The correction of the basis set superposition error and the inclusion of the deformation energies (for the S12L set) have been crucial for obtaining precise DLPNO-CCSD(T)/CBS interaction energies. Among the density functionals evaluated, the double-hybrid revPBE0-DH-NL and B2PLYP-NL with the three-body dispersion correction provide remarkably accurate association energies very close to the chemical accuracy. Overall, the NL van der Waals approach combined with proper density functionals can be seen as an accurate and affordable computational tool for the modeling of large weakly bonded supramolecular systems.
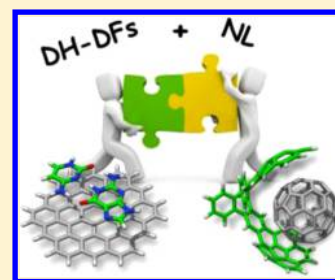
■ INTRODUCTION

Noncovalent interactions (NCIs) are considered to be weak but crucial forces governing the tridimensional organization and reactivity of molecular biological systems, such as proteins and nucleic acids, and their understanding is the key for the development of drug design, crystallinity, and design of novel materials by controlling their self-assembly. NCIs can be seen as a combination of electrostatic and dispersion forces; the latter, which mainly arise from long-range electron correlation effects, can be properly captured by highly correlated wave function methods. Coupled-cluster theory with singles, doubles, and perturbatively connected triple excitations [CCSD(T)] in combination with large basis sets has become the "gold-standard" method to accurately deal with these weak but important forces in supramolecular systems.[1−3] Its use is, however, restricted to small- or medium-size molecular systems due to its unfavorable computational $O(N^7)$ scaling, where N is related to the molecular size. In this sense, the longstanding quest for efficient highly correlated wave function methods for applications in large (real-life) systems is still an active research field.

Very recently, Riplinger et al. have successfully coupled the domain-based local pair natural orbital (DLPNO) approach with the CCSD(T) method providing a highly correlated near-linear-scaling methodology known as DLPNO-CCSD(T).[4] The DLPNO-CCSD(T) method has proved to offer promising average errors in relative energies of approximately 0.5 kcal/mol for a challenging test set of medium-size organic molecules,

and it is thus increasingly being applied for benchmarking calculations in large systems, where the canonical CCSD(T) approach is unaffordable.[5] This methodology therefore not only opens the door for rigorous studies in large molecular systems but also offers the opportunity to benchmark and parametrize less-costly quantum chemical methods in the context of NCIs.

On the other hand, density functional theory (DFT) has become the standard tool for a vast array of quantum chemistry applications owing to its more favorable computational cost ranging from $O(N^3)$ to $O(N^5)$. Nevertheless, common density functionals (DFs) are not capable of totally capturing the long-range correlation phenomena required for the adequate description of NCIs.[6−8] Over the last years, much effort within the DFT framework has been made to develop approximations that allow the accurate treatment of dispersion forces between molecular entities.[9,10] Among the most modern approaches, the atom-pairwise dispersion-corrected DFT approach, developed by Grimme et al., known in its current version as DFT-D3, is a manner of dealing with NCIs with a reasonable compromise between computational cost and accuracy.[11−14] A less popular but seamless and general approach known as van der Waals density functional theory (vdW-DFT) has recently received a great deal of attention due to its low degree of empiricism.[15,16] The vdW-DFT approach accounts for the long-range electron correlation effects by means of an explicit nonlocal (NL)

correlation functional that depends on the electron density at two different points in space (r and r') and is added to the general expression of the exchange−correlation energy functional. In the modern and efficient formulation developed by Vydrov and Van Voorhis (VV10),[15] commonly known as DFT-NL, the NL correction can be coupled easily to any standard DF.

Compared with the extended DFT-D3 approach, the NL correction has been used with a relatively scarce number of standard DFs. Hujo and Grimme assessed the performance of different generalized gradient approximation (GGA) and hybrid density functionals in combination with the NL correction.[16] More recently, the NL correction was merged with more sophisticated double-hybrid density functionals (DHDFs).[17,18] These studies have proven that DFT-NL can be considered as a robust electronic structure method capable of dealing with the challenging problems dominated by intermolecular NCIs independently of the functional of choice. However, despite the great success of the NL approximation combined with standard DFs for reliable studies in supramolecular chemistry, its performance has been mainly evaluated in small- and medium-size supramolecular systems where benchmark energies, usually obtained at CCSD(T) level with extrapolation to infinite basis set, are available. Therefore, investigating the behavior of the DFT-NL correction in very large systems would provide valuable information about the reliability of this methodology for treating NCIs in real-life problems.

In this contribution, we first provide an accurate and homogeneous set of reference interaction energies for large supramolecular complexes by using the novel DLPNO-CCSD(T) approach at the complete basis set limit (CBS). For that purpose, we have employed the L7[19] and S12L[20] data sets, which are formed by large supramolecular complexes of different nature (vide infra). Second, we use the DLPNO-CCSD(T) interaction energies to assess the performance in large supramolecular complexes of the NL correction combined with two representative double-hybrid density functionals (revPBE0-DH-NL and B2PLYP-NL), which behaved remarkably accurate for the S22 and S66 training sets.[17] A comparison with their corresponding hybrid (revPBE0-NL and B3LYP-NL) and GGA (revPBE-NL and BLYP-NL) analogues is also performed to assess the performance across the hierarchy of DFT-based methods.[21]

■ THEORETICAL MODELS AND COMPUTATIONAL DETAILS

Reference interaction energies for the L7[19] and S12L[22] training sets were computed with the DLPNO-CCSD(T) method.[4] As mentioned above, DLPNO-CCSD(T) can be seen as a highly correlated wave function method whose accuracy is comparable to the canonical CCSD(T) counterpart but with a dramatically reduced computational cost. Default values for all internal thresholds of the methodology (TCutPairs, TCutPNO, and TCutMKN) were used as suggested by the developers.[5] For further details, the reader is referred to the original literature.[4,23] The Ahlrich's def2-TZVPP basis set[24] was employed for the DLPNO-CCSD(T) calculations. Interaction energies at the DLPNO-CCSD(T)/def2-TZVPP level were extrapolated to the CBS according to the hybrid scheme proposed by Sherrill et al.[25] and Jurecka et al.,[26] and further successfully applied by Liakos et al. within the LPNO framework (eq 1).[27,28]

$$
\begin{aligned}
E^{\text{DLPNO-CCSD(T)/CBS}} &= (E^{\text{HF/CBS}} + E_{\text{corr}}^{\text{MP2/CBS}}) \\
&\quad + (E_{\text{corr}}^{\text{DLPNO-CCSD(T)/def2-TZVPP}} - E_{\text{corr}}^{\text{MP2/def2-TZVPP}}) \\
&= E^{\text{MP2/CBS}} + \Delta E_{\text{corr}}^{\text{DLPNO-CCSD(T)}}
\end{aligned}
\tag{1}
$$

Note that by using this scheme we assume that the residual difference between DLPNO-CCSD(T) and MP2 correlation energies has a less marked dependence on basis set than the MP2 correlation energy itself. The $E^{\text{MP2/CBS}}$ term was calculated by a two-point extrapolation scheme using the Dunning's cc-pVDZ and cc-pVTZ basis sets with ordinal numbers $X = 2$ and $Y = 3$, respectively, where the Hartree−Fock ($E^{\text{HF/CBS}}$) and the correlation ($E_{\text{corr}}^{\text{MP2/CBS}}$) terms were computed as

$$
E^{\text{HF/CBS}} = \frac{E^{\text{HF/}(X)}e^{-\alpha\sqrt{Y}} - E^{\text{HF/}(Y)}e^{-\alpha\sqrt{X}}}{e^{-\alpha\sqrt{Y}} - e^{-\alpha\sqrt{X}}}
\tag{2}
$$

and

$$
E_{\text{corr}}^{\text{MP2/CBS}} = \frac{X^{\beta}E_{corr}^{\text{MP2/}(X)} - Y^{\beta}E_{corr}^{\text{MP2/}(Y)}}{X^{\beta} - Y^{\beta}}
\tag{3}
$$

The α and β exponents were set at 4.42 and 2.46 as reported recently in the literature.[29]

The relevance of the basis set superposition error (BSSE) in medium supramolecular systems has been evidenced, for example, by Janowski et al.[30] in the case of the coronene dimer. Furthermore, the seminal works on the L7[19] and the S12L[22] data sets made use of the BSSE correction, since large basis sets close to the CBS limit remain prohibitive due to the size of the supramolecular systems. In this work, the basis set convergence was analyzed in detail for the supramolecular complex **C3GC**, which can be considered as a representative example of the L7 and S12L test sets (Table S1). The correlation interaction energy computed at MP2/CBS(cc-pVDZ/cc-pVTZ) (−64.69 kcal/mol) provides a small deviation of only 1.9% compared to CBS(cc-pVTZ/cc-pVQZ) (−63.50 kcal/mol) when the counterpoise (CP)[31] correction is included. All the MP2 energies were therefore calculated at the CBS(cc-pVDZ/cc-pVTZ) level and were CP-corrected to reduce the BSSE. Table S2 clearly shows a notable BSSE effect on the systems considered in L7 and S12L data sets at the MP2/CBS level.

Deformation energies (vide infra) were computed at the spin-scaled MP2 method (SCS-MP2).[32] This method has proven to attenuate the overbinding tendency of MP2 in π−π supramolecular complexes similar in nature but smaller in size than those considered herein.[33] All MP2 and SCS-MP2 calculations made use of the resolution of the identity (RI) and the chain-of-spheres (COSX) techniques, for Coulomb and exchange integrals, respectively, to alleviate the computational cost. The corresponding matching auxiliary basis sets (cc-pVDZ/C and cc-pVTZ/C, and cc-pVDZ/JK and cc-pVTZ/JK) were employed throughout.[34] Grids for the COSX approximation were increased from defaults to Gridx6.[35]

Two families of DFs have been employed in this work: the Becke-Lee−Yang−Parr (BLYP) and the revised Perdew−Burke−Ernzerhof (revPBE). For each family, the GGA, the hybrid (H), and the double-hybrid (DH) DFs were used. Note that the GGA is considered as the baseline for further improvements (rungs) within each family. A general expression for the exchange−correlation (xc) density functionals utilized in this work and coupled to the NL approach can be expressed as

Table 1. Composition of the Exchange−Correlation Functionals Used along This Work

functional	type	$E_x[\rho]$	$E_c[\rho]$	$w_{HF}{}^a$	$w_{PT2}{}^b$	b^c	ref
revPBE0-DH-NL	double-hybrid GGA	revPBE	PBE	0.50	0.125	5.7	17
revPBE0-NL	hybrid GGA	revPBE	PBE	0.25	0	4.2	37
revPBE-NL	GGA	revPBE	PBE	0	0	3.6	38, 39
B2PLYP-NL	double-hybrid GGA	B88	LYP, VWN	0.53	0.270	7.8	36
B3LYP-NL	hybrid GGA	B88	LYP, VWN	0.20	0	4.6	40, 41
BLYP-NL	GGA	B88	LYP	0	0	4.0	42, 43

^{a}Weight of the HF-like exchange. ^{b}Weight of the perturbative term. ^{c}Values used for the adjustable parameter b in the E_c^{NL} term.

$$E_{xc}[\rho] = w_{HF}E_x^{HF} + (1 - w_{HF})E_x[\rho] + (1 - w_{PT2})E_c[\rho]$$
$$+ w_{PT2}E_c^{PT2} + E_c^{NL}[\rho] \qquad (4)$$

where $E_x[\rho]$ and $E_c[\rho]$ correspond to the GGA exchange and correlation energy terms, respectively, weighted by the scaling parameter w_i. The E_x^{HF} and E_c^{PT2} terms are, respectively, the exact exchange HF-like energy and the correlation energy obtained at the Møller−Plesset perturbation theory up to second order.[36] Note that the E_x^{HF}, E_c^{PT2}, and E_c^{NL} terms are evaluated with the orbitals arising from the solution of the Kohn−Sham one-electron equations but discarding the last two terms of eq 4. Table 1 presents a detailed description for the composition of all the exchange−correlation functionals used in this work.

The DFs under study were merged with the NL approximation (DFT-NL), such as it was proposed by Vydrov and Van Voorhis.[15] The NL energy functional is added to the exchange−correlation energy functional nonself-consistently.[16,17] In the DFT-NL approach, the NL energy functional depends on the electron density, and the sort-range (b) and long-range (C) adjustable parameters (the reader is referred to the original work for further details, ref 15). The short-range attenuation parameter b used here for the density functionals was carefully fitted using the S22[44,45] and S66[46,47] data sets,[17] whereas the long-range parameter C is fixed to its original value $C = 0.0093$ (its optimization leads to only minor improvements).[16,17] To assess the errors of the different DFT-NL functionals, the mean signed error (ME), the mean absolute error (MAE), the mean absolute relative error (MARE), and the maximum absolute error (MAX) are employed. A negative ME indicates an overbinding trend, while positive ME values signify an underestimation in the interaction energy.

The DFT-NL calculations were performed with the def2-TZVP basis set.[24,48] Although the combination of DFT calculations (with dispersion corrections) and a basis set of triple-ζ quality (cc-pVTZ) has been shown to provide interaction energies with a small BSSE,[49] this error can be large enough for the supramolecular systems investigated herein. Consequently, all the interaction energies calculated were counterpoise corrected (see eq S1). The computational effort for the DFT calculations was significantly reduced in all cases by making use of the RI[50] and the COSX[51] techniques, for Coulomb and exchange integrals, respectively. The COSX approximation was extensively applied by using the overlap fitting procedure described in ref 52. Large grids (Gridx6), especially required for the treatment of noncovalent interactions, have been employed in the COSX procedure. The corresponding matching auxiliary basis sets def2-TZVP/C and def2-TZVP/JK were employed throughout.[34] The quadrature grids needed for numerical integration of DFs were also increased (Grid6) with respect to defaults, as it is strongly

recommended for intermolecular interaction energies as well as the corresponding thresholds for converging energies in the self-consistent field procedure (TightSCF). A larger grid for the NL approximation was also used (Vdwgrid4). For comparison, double-hybrid B2PLYP-D3 calculations were carried out by means of the Grimme's dispersion correction (D3) using the Becke−Johnson damping function.[11,12,53−55] The three-body dispersion correction (E_{ABC})[11] has been evaluated in all cases. All the calculations were carried out by using the ORCA 3.0 program package.[35]

RESULTS AND DISCUSSION

Benchmark Energies for Large Supramolecular Complexes. Two sets of very large supramolecular complexes (known as L7 and S12L) have been specially designed to evaluate the performance of accurate quantum chemical methods with low computational cost for "real-life" applications (see Figures S1 and S2 for the structure of the supramolecular complexes of L7 and S12L, respectively). The L7 set proposed by Sedlak et al.[19] contains seven supramolecular complexes intentionally selected to be mostly dispersion-dominated (aliphatic−aliphatic and $\pi−\pi$ interactions) and their size ranges from 48 to 112 atoms. The reference association energies reported before for L7 were computed at the QCISD(T)/CBS level.[19] The S12L data set includes 12 supramolecular complexes (involving both neutral and charged species) dominated by nonpolar, π-stacking, H-bonding, and electrostatic cation−dipolar interactions. In contrast to the L7 set, the reference interaction energies in the S12L set, proposed in the seminal work of Grimme et al., were estimated from (back-corrected) experimental binding affinities.[20] Nevertheless, the authors admitted some uncertainty about the validity of these gas-phase interaction energies for benchmarking purposes. In a further step, Ambrosetti et al. provided accurate values for the interaction energy of six host−guest complexes of the S12L set at the diffusion quantum Monte Carlo (DQMC) level.[56] More recently, Hesselmann et al. calculated the interaction energies for the complexes of S12L by using intermolecular symmetry-adapted perturbation theory combined with a DFT description of the interacting molecules.[57]

Due to the difference in the methods employed to estimate the reference binding energies in L7 and S12L (theoretical and experimental data), a homogeneous set of benchmark interaction energies computed at the same level of theory is a clear demand for these data sets. Hence, interaction energies at the DLPNO-CCSD(T)/CBS level of theory have been consistently computed for L7 and S12L (Table 2). Hesselmann et al.[57] showed that the deformation energies can be important for compounds of S12L and that they should be taken into account to provide accurate interaction energy values. Additionally, it was proven that the deformation energies can be overestimated by MP2 due to its overbinding tendency in $\pi−\pi$

Table 2. Interaction Energies (in kcal/mol) Computed for the L7 and S12L Data Sets at the DLNPO−CCSD(T)/CBS Level

data set	complex	DLPNO-CCSD(T)/CBS[a]	QCISD(T)/CBS[b]	exptl[c]	DQMC[d]
L7	C2C2PD	−24.81	−24.36		
	C3A	−17.98	−18.19		
	C3GC	−29.86	−31.25		
	CBH	−11.64	−11.06		
	GCGC	−13.21	−14.37		
	GGG	−1.68	−2.40		
	PHE	−22.81	−25.76		
S12L	2a	−30.65		−29.9	−27.2
	2b	−23.04		−20.5	−17.2
	3a	−23.68		−24.3	−
	3b	−23.08		−20.4	−
	4a	−[e]		−27.5	−25.8
	4b	−[e]		−28.7	−
	5a	−33.40		−34.8	−33.4
	5b	−22.98		−21.3	
	6a	−79.79		−77.4	−81.0
	6b	−77.82		−77.0	−
	7a	−123.89[f]		−131.5	−
	7b	−22.74		−22.6	−24.1

[a]This work. [b]Reference interaction energies computed at the QCISD(T) level and extrapolated to the CBS by means of a hybrid scheme as explained in ref 19. For the correlated component of the energy, a relatively small 6-31G*(0.25) basis set was employed. [c]Gas-phase interaction energies were obtained back-correcting experimental values for binding affinities in solution according to ref 20. [d]The stochastic DQMC electronic structure method using Slater−Jastrow trial wave functions was employed to compute the reference interaction energies as explained in ref 56. [e]Calculations for these complexes at the DLPNO-CCSD(T)/def2-TZVPP did not converge due to the high number of interacting electron pairs between the two moieties originated by the huge amount of close carbon−carbon contacts. [f]The $E^{MP2/CBS}$ term was computed according to eqs 2 and 3 by using Ahlrichs basis sets with double and triple-ζ quality (def2-SVP and def2-TZVPP, respectively) due to the inexistent definition of the auxiliary cc-pVDZ/C basis set for Fe. The exponential α and β values were set to 10.39 and 2.40, respectively.[29]

supramolecular complexes (especially in compounds **3a** and **3b**). The authors corrected the deformation energies for S12L using the SCS-MP2 method since it provides a better description for $\pi-\pi$ interactions than MP2 does.[33] Deformation energies (see the Supporting Information for further details) were therefore corrected at SCS-MP2/CBS for S12L (Table S3). Deformation energies for complexes of L7 were supposed to be small and were not corrected.[19] Table S4 collects the different interaction energy terms of the total interaction energy computed for L7 and S12L supramolecular systems.

The combination of the two sets of systems provides a wide range of interaction energies for very large supramolecular complexes ranging from −1.68 kcal/mol in the stacked guanine trimer GGG of the L7 set (Figure S1) to −123.89 kcal/mol in the cucurbit[7]uril@ferrocene-based complex of S12L (Figure S2). Broadly speaking, the DLPNO-CCSD(T) association energies present small differences with respect to the previous values available in the literature, especially for the L7 set, which were obtained from a purely theoretical treatment.[19] The largest discrepancy obtained for the L7 set is found for the PHE complex (an amyloid fragment consisting in three phenyl-alanine residues in mixed H-bonded-stacked conformation, Figure S1), for which a value of −22.81 kcal/mol is calculated at DLPNO-CCSD(T) and of −25.76 kcal/mol at QCISD(T)/CBS. The MAE between the reported interaction energies and our suggested DLPNO-CCSD(T) values is of only 1.07 kcal/mol. The experimental back-corrected values provided for S12L[20] differ from our suggested DLPNO-CCSD(T) energies with a MAE of ∼2 kcal/mol. The largest (absolute) deviation is found for the doubly positive-charged ferrocene/adamantane derivative **7a**, for which a difference of 7.61 kcal/mol (6%) is obtained. This may stem from the important role of the counterions in the stabilization of the doubly charged **7a** complex, which would increase the experimental free interaction energy used in obtaining the values presented in the fifth column of Table 2. For the rest of supramolecular complexes, the proposed DLPNO-CCSD(T) interaction energies remain close to the previously reported values. We therefore consider that the interaction energies computed at
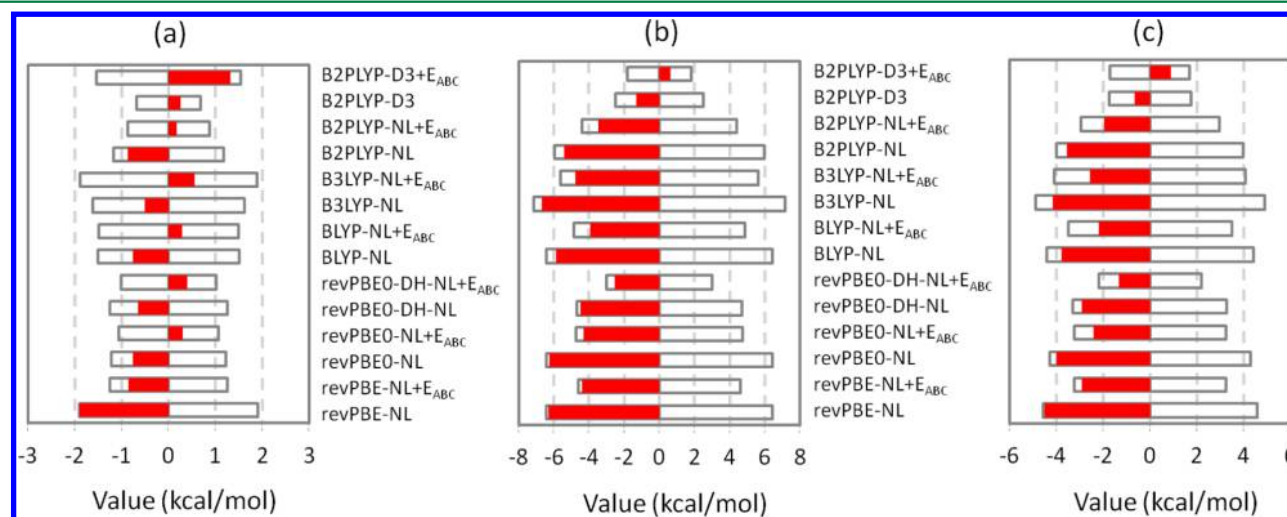
Figure 1. ME (red bar) and MAE (gray-border bar) computed for the different DFs coupled to the NL approach (the B2PLYP-D3 method is included for comparison purposes) for L7 (a), S12L (b), and considering both sets as one (c). The DLNPO−CCSD(T)/CBS interaction energies given in Table 2 are used as a reference. Note that the MAE (gray-bordered bar) has been mirrored.

DLPNO-CCSD(T)/CBS can be used as reference values for benchmark studies, and they are employed in the following discussion if not otherwise indicated.

Performance of NL-Corrected BLYP and revPBE Functional Families. Recently, several double-hybrid functionals merged with the NL approach were tested in the small S22 and S66 data sets, and they were shown to provide a very good performance to deal with NCIs of different nature.[17] However, their behavior when increasing the molecular size has not been analyzed so far. We have evaluated the performance of the B2PLYP-NL and revPBE0-DH-NL functionals as well as their hybrid (B3LYP-NL and revPBE0-NL) and GGA (BLYP-NL and revPBE-NL) analogues in the large L7 and S12L test sets. The B2PLYP-NL and revPBE0-DH-NL functionals have been selected because they exhibited the best performance in the S22 and S66 data sets.[17]

Figure 1 summarizes the ME and MAE errors for the association energies of the L7 and S12L data sets obtained for the different DFs coupled to the NL approach with respect to the reference DLPNO-CCSD(T)/CBS values given in Table 2. The values of the errors (ME, MAE, MARE, and MAX) and of the interaction energies are given in Tables S5−S7. ME, MAE, MARE, and MAX errors with respect to reference values reported in refs 19 and 20 for L7 and S12L, respectively, have also been included for comparison purposes (Tables S8−S10 and Figure S3). Similarly to that performed for the DLPNO-CCSD(T)/CBS interaction energies, deformation energies have been included for all the density functionals for complexes of S12L but not for those of L7. Despite their different flavor according to the Jacob's ladder (GGA, hybrid, and double-hybrid character), the two families of DFT-NL functionals perform strikingly well for the L7 set with MAE errors below 2.00 kcal/mol (Figure 1a and Table S5). Note that it is being admitted now that the standard "chemical accuracy" (1 kcal/mol) is too stringent when interaction energies of very large supramolecular complexes are computed and deviations around 2−3 kcal/mol can be perfectly accepted.[58] A closer inspection of Figure 1a clearly reveals that the double-hybrid B2PLYP-NL and revPBE0-DH-NL functionals exhibit a very good performance within their corresponding family of DFs providing small MAE errors of <1.25 kcal/mol. For example, in the BLYP family, the MAE noticeably decreases from 1.51 kcal/mol (BLYP-NL) and 1.62 kcal/mol (B3LYP-NL) to 1.17 kcal/mol (B2PLYP-NL). A similar behavior is also found for the revPBE family with MAE values of 1.91, 1.22, and 1.25 kcal/mol for revPBE-NL, revPBE0-NL, and revPBE0-DH-NL, respectively. Both double-hybrid functionals show a systematic overestimation of the dispersion forces governing the supramolecular associates with negative ME errors of −0.88 and −0.65 kcal/mol for B2PLYP-NL and revPBE0-DH-NL, respectively. The overestimation can be easily understood since many-body dispersion effects are not included in the DFT-NL functionals. It should be mentioned that in both double-hybrid functionals the MAX error is found for complex **C3GC** with values below 2.60 kcal/mol (Table S5). In terms of the MARE errors (Table S5) and in agreement with MAE values, B2PLYP-NL exhibits the smallest MARE value (6.6%).

Recently, the repulsive three-body approximation (E_{ABC}), usually employed with the pairwise dispersion correction developed by Grimme et al., has been successfully combined with the NL approximation to correct the interaction energies in large supramolecular complexes and the cohesive energies calculated for the anthracene crystal.[22,59] The inclusion of the

E_{ABC} term has a significant effect in both double-hybrid functionals. For revPBE0-DH-NL, the E_{ABC} term gives rise to a decrease of the MAE error (1.25 vs 1.01 kcal/mol) and a change of sign in the ME (−0.65 vs 0.40 kcal/mol) indicating a small underestimation with respect to the reference interaction energies. For B2PLYP-NL, the MAE error decreases from 1.17 to 0.88 kcal/mol upon addition of the E_{ABC} term, a small MARE (7.3%) is obtained, and the ME moves from −0.88 to 0.18 kcal/mol, which is the lowest ME value computed along the series. It should be also noted that the performance of B2PLYP-NL is similar to that found for its B2PLYP-D3 homologous, which can be thus judged also as a high-quality method, presenting a slightly smaller MAE of 0.69 kcal/mol. Nevertheless, B2PLYP-D3 does not benefit from the additional E_{ABC} correction with the corresponding increase of the MAE error up to 1.54 kcal/mol (Figure 1a). It should be also mentioned that the MAX error, once the E_{ABC} correction is introduced in both double-hybrid density functionals, is found for complex **C2C2PD** (Table S5).

Regarding the hybrid and GGA density functionals, it is necessary to stress the good performance of the hybrid revPBE0-NL functional with a MAE of 1.22 kcal/mol, which is even slightly smaller than that found for its double-hybrid analogue. In general, the incorporation of the three-body dispersion correction in the GGA and hybrid functionals has also an improvement in terms of MAE (Figure 1a), with the sole exception of the hybrid B3LYP-NL functional, which slightly worsens upon addition of the E_{ABC} dispersion term (from 1.62 to 1.88 kcal/mol). In terms of performance, the revPBE family incorporating the E_{ABC} term gives in overall interaction energies for the L7 set slightly closer to the reference energy values than the BLYP family with MAEs in the 1.01−1.26 kcal/mol range. Nevertheless, the double-hybrid B2PLYP-NL functional provides the most accurate interaction energies among the two families with ME and MAE going beyond the chemical accuracy (0.18 and 0.88 kcal/mol, respectively).

Moving to the S12L set of complexes, a general overestimation in the association energy is computed for all DFs under study with negative ME values (Figure 1b and Table S6). Due to this overbinding, the inclusion of the three-body E_{ABC} correction improves the results in terms of MAE errors in all cases. Similarly to the L7 data set, the double-hybrid density functionals (and their variants with the three-body dispersion correction, revPBE0-DH-NL + E_{ABC} and B2PLYP-NL + E_{ABC}) exhibit the best performance with the smallest MAE and MARE along the two density functional families. For instance, the MAE (ME) values are found to be 4.71 (−4.47) and 3.01 (−2.53) kcal/mol for revPBE0-DH-NL and revPBE0-DH-NL + E_{ABC}, respectively. A similar behavior is predicted for B2PLYP-NL and B2PLYP-NL + E_{ABC} in terms of MAE errors (5.95 and 4.39, respectively). MARE values below 10% are only found for revPBE0-DH-NL + E_{ABC} and B2PLYP-NL + E_{ABC}. The general overestimation and the improvement with the E_{ABC} correction is in concordance with previous results where the overbinding and the role of the three-body E_{ABC} correction in large supramolecular complexes were also discussed.[22] It is necessary to mention that we neglect at this stage the many-body corrections higher than E_{ABC} terms, which might still contribute in systems with large and permanent multipole moments, but are, however, beyond the scope of the present study. Note that B2PLYP-D3 with the addition of the E_{ABC}

term behaves slightly better than their analogues, revPBE0-DH-NL + E_{ABC} and B2PLYP-NL + E_{ABC} (Figure 1b).

Among the GGA and hybrid density functionals, the best performance is found for the revPBE-NL + E_{ABC} with ME and MAE errors of −4.38 and 4.62 kcal/mol, respectively. Nevertheless, the rest of GGA and hybrid functionals upon inclusion of the E_{ABC} correction show MAEs below 5.61 kcal/mol (Table S6). Figure S4 displays the absolute errors for each compound of S12L according to the nature of the interaction. Compounds **7a** and **7b** dominated by cation−dipolar and H-bonding interactions, respectively, present the maximum errors for all density functionals, whereas compounds mainly dominated by nonpolar and $\pi-\pi$ interactions (compounds **2a**, **2b**, **3a**, and **3b**) are better described (Figure S4). For further comparison, interaction energies reported by Risthaus et al.[22] for S12L using the GGA PBE functional with the def2-TZVP basis set were extracted and compared to the DLPNO-CCSD(T)/CBS reference values proposed in this work (Table S14). The PBE-NL + E_{ABC} functional outperforms the accuracy of the revPBE-NL + E_{ABC} with MAE and ME values of 3.76 and −2.52 kcal/mol, respectively.[22] Furthermore, the PBE functional coupled to the D3 approximation and the three-body dispersion term behaves remarkably accurate with a MAE value of only 1.87 kcal/mol for the S12L data set, corroborating the accuracy of the D3 approach.

By combining the L7 and S12L test sets (Figure 1c), the relative deviations resemble the picture obtained for S12L due to the larger MAE and ME values computed for this set. Again, the inclusion of the three-body dispersion correction improves the results in all GGA, hybrid, and double-hybrid DFs due to the general overbinding of the interaction energies computed along the different DFs merged with the NL approach. Interestingly, the revPBE0-DH-NL and B2PLYP-NL functionals with the E_{ABC} correction show promising results with MAEs (MAREs) of 2.19 (9.6%) and 2.95 kcal/mol (8.9%), respectively, for the joint L7 + S12L data set containing very large supramolecular complexes of diverse interacting nature. Likewise, the ME values are computed to be as small as −1.32 kcal/mol for revPBE0-DH-NL + E_{ABC} and −1.96 kcal/mol for B2PLYP-NL + E_{ABC}. It is worth noting that the conclusions drawn by using the DLPNO-CCSD(T)/CBS interaction energies presented here or by using the previously reported reference data (refs 19 and 20) are in good accord (compare Figures 1 with S3). Again, both double-hybrid revPBE0-DH-NL + E_{ABC} and B2PLYP-NL + E_{ABC} provide the smallest MAEs (3.08 and 2.51 kcal/mol, respectively) along the series for the joint L7 + S12L data set. Therefore, the rung of double-hybrid density functionals (revPBE0-DH and B2PLYP-NL) coupled to the NL approach, which have demonstrated a great performance in the smaller S22 and S66 test sets,[17] is now consolidated as accurate functionals for challenging and "real-life" supramolecular systems, standing close to the chemical accuracy.

■ CONCLUSIONS

In this contribution, we provide an accurate and homogeneous set of reference interaction energies for the L7 and S12L sets calculated by using the DLPNO-CCSD(T) method extrapolated to the CBS limit. The correction of the basis set superposition error and the incorporation of the deformation energies (especially for the S12L set) have been crucial for obtaining accurate interaction energies. These interaction energy values are taken as a reference to assess the performance of the double-hybrid density functionals, revPBE0-DH-NL and

B2PLYP-NL, as well as their hybrid and GGA counterparts, which include the NL van der Waals correction formulated by Vydrov and Voorhis to describe noncovalent interactions. Among the density functionals evaluated, the double-hybrid revPBE0-DH-NL and B2PLYP-NL with the three-body dispersion correction behave remarkably accurate with MAEs in the range of the chemical accuracy for the large weakly bound molecular systems of the L7 and S12L sets. The performance of the double hybrid density functionals is irrespective of the reference interaction energies employed. These results therefore confirm that the NL van der Waals approach combined with modern density functionals can be seen as an accurate and affordable computational tool for the modeling of "real-life" supramolecular assemblies governed by noncovalent interactions.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Chemical structures of the supramolecular complexes and interaction energies computed at DLPNO-CCSD(T)/CBS and with the different DFT-NL functionals for the weakly bonded molecular complexes of the L7 and S12L sets. Absolute errors according to the nature of the interaction computed for S12L. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: juan.arago@uv.es.

Present Address

‡Department of Chemistry and Centre for Scientific Computing, University of Warwick, Coventry CV4 7AL, United Kingdom.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ 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) Riley, K. E.; Hobza, P. Noncovalent Interactions in Biochemistry. *WIREs Comput. Mol. Sci.* **2011**, *1*, 3−17.

(3) Hobza, P. Calculations on Noncovalent Interactions and Databases of Benchmark Interaction Energies. *Acc. Chem. Res.* **2012**, *45*, 663−672.

(4) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.

(5) Sparta, M.; Neese, F. Chemical Applications Carried Out by Local Pair Natural Orbital Based Coupled-Cluster Methods. *Chem. Soc. Rev.* **2014**, *43*, 5032−5041.

(6) Kristyán, S.; Pulay, P. Can (Semi)local Density Functional Theory Account for the London Dispersion Forces? *Chem. Phys. Lett.* **1994**, *229*, 175−180.

(7) Hobza, P.; šponer, J.; Reschel, T. Density Functional Theory and Molecular Clusters. *J. Comput. Chem.* **1995**, *16*, 1315−1325.

(8) Allen, M. J.; Tozer, D. J. Helium Dimer Dispersion Forces and Correlation Potentials in Density Functional Theory. *J. Chem. Phys.* **2002**, *117*, 11113−11120.

(9) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211−228.

(10) Grimme, S.; Antony, J.; Schwabe, T.; Muck-Lichtenfeld, C. Density Functional Theory with Dispersion Corrections for Supramolecular Structures, Aggregates, and Complexes of (Bio)organic Molecules. *Org. Biomol. Chem.* **2007**, *5*, 741−758.

(11) 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.

(12) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

(13) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.

(14) Grimme, S. Accurate Description of van der Waals Complexes by Density Functional Theory Including Empirical Corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.

(15) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals Density Functional: The Simpler the Better. *J. Chem. Phys.* **2010**, *133*, 244103.

(16) Hujo, W.; Grimme, S. Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 3866−3871.

(17) Aragó, J.; Ortí, E.; Sancho-García, J. C. Nonlocal van der Waals Approach Merged with Double-Hybrid Density Functionals: Toward the Accurate Treatment of Noncovalent Interactions. *J. Chem. Theory Comput.* **2013**, *9*, 3437−3443.

(18) Yu, F. Spin-Component-Scaled Double-Hybrid Density Functionals with Nonlocal van der Waals Correlations for Noncovalent Interactions. *J. Chem. Theory Comput.* **2014**, *10*, 4400−4407.

(19) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364−3374.

(20) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem.—Eur. J.* **2012**, *18*, 9955−9964.

(21) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange−correlation energy. *AIP Conf. Proc.* **2001**, *577*, 1−20.

(22) Risthaus, T.; Grimme, S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 1580−1591.

(23) Riplinger, C.; Neese, F. An Efficient and Near Linear Scaling Pair Natural Orbital Based Local Coupled Cluster Method. *J. Chem. Phys.* **2013**, *138*, 034106.

(24) 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.

(25) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the Ab Initio Limit for π−π Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.

(26) Jurečka, P.; Hobza, P. On the Convergence of the (ΔECCSD-(T)−ΔEMP2) Term for Complexes with Multiple H-bonds. *Chem. Phys. Lett.* **2002**, *365*, 89−94.

(27) Liakos, D. G.; Hansen, A.; Neese, F. Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods. *J. Chem. Theory Comput.* **2010**, *7*, 76−87.

(28) Liakos, D. G.; Neese, F. Improved Correlation Energy Extrapolation Schemes Based on Local Pair Natural Orbital Methods. *J. Phys. Chem. A* **2012**, *116*, 4801−4816.

(29) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated ab initio Methods? *J. Chem. Theory Comput.* **2011**, *7*, 33−43.

(30) Janowski, T.; Ford, A. R.; Pulay, P. Accurate Correlated Calculation of the Intermolecular Potential Surface in the Coronene Dimer. *Mol. Phys.* **2009**, *108*, 249−257.

(31) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−566.

(32) Grimme, S. Improved Second-order Møller−Plesset Perturbation Theory by Separate Scaling of Parallel- and Antiparallel-spin Pair Correlation Energies. *J. Chem. Phys.* **2003**, *118*, 9095−9102.

(33) Antony, J.; Grimme, S. Is Spin-Component Scaled Second-Order Møller−Plesset Perturbation Theory an Appropriate Method for the Study of Noncovalent Interactions in Molecules? *J. Phys. Chem. A* **2007**, *111*, 4862−4868.

(34) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97*, 119−124.

(35) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73−78.

(36) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(37) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(38) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0Model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.

(39) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1998**, *80*, 890−890.

(40) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(41) Barone, V.; Adamo, C. Theoretical Study of Direct and Water-Assisted Isomerization of Formaldehyde Radical Cation. A Comparison between Density Functional and Post-Hartree-Fock Approaches. *Chem. Phys. Lett.* **1994**, *224*, 432−438.

(42) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(43) Lee, C.; Yang, W.; 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.

(44) Jurecka, P.; Sponer, J.; Cerny, 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.

(45) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis Set Convergence of the Coupled-Cluster Correction, δMP2CCSD(T): Best Practices for Benchmarking Non-Covalent Interactions and the Attendant Revision of the S22, NBC10, HBC6, and HSG Databases. *J. Chem. Phys.* **2011**, *135*, 194102.

(46) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.

(47) Řezáč, J.; Riley, K. E.; Hobza, P. Extensions of the S66 Data Set: More Accurate Interaction Energies and Angular-Displaced Nonequilibrium Geometries. *J. Chem. Theory Comput.* **2011**, *7*, 3466−3470.

(48) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571−2577.

(49) Aragó, J.; Sancho-García, J. C.; Ortí, E.; Beljonne, D. Ab Initio Modeling of Donor−Acceptor Interactions and Charge-Transfer Excitations in Molecular Complexes: The Case of Terthiophene−

Tetracyanoquinodimethane. *J. Chem. Theory Comput.* **2011**, *7*, 2068−2077.

(50) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240*, 283−290.

(51) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, Approximate and Parallel Hartree−Fock and Hybrid DFT Calculations. A 'Chain-of-Spheres' Algorithm for the Hartree−Fock Exchange. *Chem. Phys.* **2009**, *356*, 98−109.

(52) Izsák, R.; Neese, F. An Overlap Fitted Chain of Spheres Exchange Method. *J. Chem. Phys.* **2011**, *135*, 144105.

(53) Johnson, E. R.; Becke, A. D. A post-Hartree−Fock Model of Intermolecular Interactions. *J. Chem. Phys.* **2005**, *123*, 024101.

(54) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124*, 174104.

(55) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction. *J. Chem. Phys.* **2005**, *122*, 154104.

(56) Ambrosetti, A.; Alfè, D.; DiStasio, R. A.; Tkatchenko, A. Hard Numbers for Large Molecules: Toward Exact Energetics for Supramolecular Systems. *J. Phys. Chem. Lett.* **2014**, *5*, 849−855.

(57) Heßelmann, A.; Korona, T. Intermolecular Symmetry-Adapted Perturbation Theory Study of Large Organic Complexes. *J. Chem. Phys.* **2014**, *141*, 094107.

(58) Steinmetz, M.; Hansen, A.; Ehrlich, S.; Risthaus, T.; Grimme, S. *Accurate Thermochemistry for Large Molecules with Modern Density Functionals*; Springer: Berlin Heidelberg, 2014; pp 1−23.

(59) Sancho-García, J. C.; Aragó, J.; Ortí, E.; Olivier, Y. Obtaining the Lattice Energy of the Anthracene Crystal by Modern Yet Affordable First-Principles Methods. *J. Chem. Phys.* **2013**, *138*, 204304.