Article

A Cost Effective Scheme for the Highly Accurate Description of Intermolecular Binding in Large Complexes

Jiří Czernek [1,*] **, Jiří Brus** [1] **and Vladimíra Czerneková** [2]

1   Institute of Macromolecular Chemistry, Czech Academy of Sciences, Heyrovsky Square 2,
    162 00 Prague, Czech Republic
2   Institute of Physics, Czech Academy of Science, Na Slovance 2, 182 21 Prague, Czech Republic
*   Correspondence: czernek@imc.cas.cz; Tel.: +420-296809-290

Abstract: There has been a growing interest in quantitative predictions of the intermolecular binding energy of large complexes. One of the most important quantum chemical techniques capable of such predictions is the domain-based local pair natural orbital (DLPNO) scheme for the coupled cluster theory with singles, doubles, and iterative triples [CCSD(T)], whose results are extrapolated to the complete basis set (CBS) limit. Here, the DLPNO-based focal-point method is devised with the aim of obtaining CBS-extrapolated values that are very close to their canonical CCSD(T)/CBS counterparts, and thus may serve for routinely checking a performance of less expensive computational methods, for example, those based on the density-functional theory (DFT). The efficacy of this method is demonstrated for several sets of noncovalent complexes with varying amounts of the electrostatics, induction, and dispersion contributions to binding (as revealed by accurate DFT-based symmetry-adapted perturbation theory (SAPT) calculations). It is shown that when applied to dimeric models of poly(3-hydroxybutyrate) chains in its two polymorphic forms, the DLPNO-CCSD(T) and DFT-SAPT computational schemes agree to within about 2 kJ/mol of an absolute value of the interaction energy. These computational schemes thus should be useful for a reliable description of factors leading to the enthalpic stabilization of extended systems.

Keywords: noncovalent interactions; intermolecular binding; CCSD(T); DLPNO; DFT-SAPT

1. Introduction

A quantitative description of intermolecular noncovalent interactions is the key factor in understanding properties of gaseous and condensed phases [1], molecular recognition [2], some chemical transformations [3], and supramolecular structures [4]. Hence, intermolecular interactions are intensely studied by both experimental and theoretical methods, as reviewed in reference [5]. From among of these theoretical methods, of particular importance is the symmetry adapted perturbation theory (SAPT) of intermolecular interactions [6], since it can be combined with the density-functional theory (DFT) treatment of monomers [7] to accurately characterize the nature of noncovalent bonding even for very large (containing over 100 atoms) complexes [8–10]; this technique is denoted here as DFT-SAPT. Another particularly important group of theoretical methods for noncovalent interactions are the highly correlated ab initio approaches [11], because they reliably describe the strength of all types of noncovalent bonding, as exemplified in references [12–18] (in the following, the strength of an intermolecular interaction will be expressed by the interaction energy, ΔE, in kJ/mol). Within the highly correlated approaches, the coupled cluster theory with singles, doubles, and iterative triples [CCSD(T)] is of a special significance. Namely, results of simpler methods for the ΔE prediction are frequently evaluated against the CCSD(T) ΔE value extrapolated to its complete basis set limit (CBS), as such value is considered to be sufficiently accurate for practically all applications. Since the canonical CCSD(T)/CBS computations are unfeasible for larger complexes (at present the

limit is about 60 atoms [19]), reduced-scaling variants of the CCSD(T) method are used, which were most recently surveyed in reference [20]. Currently the domain-based local pair natural orbital (DLPNO) variant [21–25] of the CCSD(T) is the most important due to its particularly favorable scaling with the system size [26,27]. The DLPNO-CCSD(T)/CBS calculations can be employed to obtain the benchmark ΔE values for very large complexes (see references [28–30] and work cited therein, and also the related investigation of chemical reactivity [31]). Nevertheless, various simplified procedures for the DLPNO-CCSD(T)/CBS ΔE estimation are of interest. Examples include an application of the multiplicative CBS extrapolation protocol in reference [32]; using smaller basis sets to obtain the DLPNO-CCSD(T) correction term as described in reference [28]; and extrapolating energies computed with certain reductions in the DLPNO correlation space in reference [33]. These and other simplified procedures aim at reducing the computational cost of the underlying calculations while retaining the quality of a CBS-extrapolated result. The present investigation is also oriented toward this objective. Namely, a general goal of this work is to devise the DLPNO-CCSD(T) based prediction scheme for obtaining completely reliable ΔE values at a relatively low computational cost. In the initial step toward achieving this goal, a diverse testing set of 27 complexes is selected and its reference canonical CCSD(T)/CBS data are obtained together with a careful characterization of the intermolecular binding by means of the DFT-SAPT, as described in Section 2.1. In the next step, the basis set incompleteness error and settings that affect the quality of the DLPNO-CCSD(T) results are studied, and a robust procedure is described that is based on the focal-point analysis [34] of the ΔE data computed using two larger basis sets (see Section 2.2). This procedure is validated using the aforementioned testing set and also for systems from the renowned S22 collection [35]. In the subsequent step, which is detailed in Section 2.3, an even cheaper method is presented. It is based on the fitting of the benchmark values of absolute energies from the S22 dataset to their counterparts, computed using two smaller basis sets. The parameters thus obtained serve for an estimation of the ΔE values of complexes from the aforementioned testing set, which enables establishment of the accuracy limits of this computationally cheap approach. Since the fitting procedure is found to lead to quite reliable results, it is applied to large systems in the final step of this investigation (see Sections 2.4 and 3). Specifically, for two polymorphic forms of poly(3-hydroxybutyrate) (PHB) that were previously characterized experimentally [36,37], dimeric models of PHB chains are considered. Moreover, three systems from the L7 dataset of large complexes [38] are examined: the parallel-displaced dimer of coronene (C2C2PD), the guanine trimer (GGG), and the tetramer consisting of two guanine–cytosine pairs (GCGC), as their ΔE values predicted by the CCSD(T) and the Quantum Monte Carlo (QMC) [39] approaches are a matter of the ongoing debate (see, in particular, references [40–43]).

2. Results

2.1. The Reference Interaction Energies

First the benchmark ΔE values had to be established for their use in a development of simpler predicting scheme(s). Thus, for a total of 47 complexes (the S22 dataset and 27 systems that are specified in Materials and Methods section), the canonical CCSD(T)/CBS interaction energies, $\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}}$, were obtained. Throughout this work, the standard correlation correlation-consistent polarized-valence basis sets augmented with one set of diffuse basis functions were used (related double-zeta, triple-zeta, quadruple-zeta, and quintuple-zeta basis set is abbreviated as aDZ, aTZ, aQZ, and a5Z, respectively), and the counterpoise correction [44] was applied to reduce the basis set superposition error. In order to estimate a value of $\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}}$, the focal-point method expressed by Equation (1) was applied to underlying energies (the basis set used to obtain the respective term is specified in the superscript).

$$\Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}} = \Delta E_{\mathrm{HF}}^{\mathrm{a5Z}} + \Delta E_{\mathrm{MP2}}^{\mathrm{a5Z}} + \Delta E_{\mathrm{post-MP2}}^{\mathrm{aTZ}} \tag{1}$$

The respective portions of the total interaction energy are the Hartree–Fock component, ΔE_{HF}, the second-order Møller–Plesset (MP2) correlation energy component, ΔE_{MP2}, and the correction for higher-order correlation energy contributions that is denoted as $\Delta E_{post-MP2}$ and approximated as a difference of the corresponding CCSD(T) and MP2 correlation energies (see the review [45] for discussion). The procedure from Equation (1) was carefully checked for the S22 dataset using the values from reference [46] (so called S22B set). Importantly, the fit of present $\Delta E_{CCSD(T)}^{CBS}$ results to their counterparts from the S22B collection is almost perfect (see Figure S1; the raw data are shown in Table S1). The highest absolute and relative differences between the two data sets are as low as 0.92 kJ/mol and 2.8%, respectively (they accordingly occur for the dimer of formic acid and for the benzene···methane complex), while the mean absolute deviation is only 0.32 kJ/mol. Hence, the method expressed by Equation (1) was applied also to the 27 testing systems. All underlying absolute energies are provided in Supporting Information inside the Excel spreadsheet 'energies1.xlsx' for $\Delta E_{CCSD(T)}^{CBS}$ estimation using Equation (1).

For all complexes investigated in this work, which are listed in Table 1, the physical nature of intermolecular bonding was described by means of the SAPT-DFT calculations. It is thus important to ascertain the accuracy of the employed SAPT-DFT/CBS computational protocol, whose details are given in Section 4. This check was performed for a subset of 18 dimers. This assembly is called 'Set3x6', because it can be divided into three groups, with six complexes in each group, on the basis of the dispersion-to-polarization ratio [47]. Namely, the respective Set3x6 groups contain electrostatics-dominated, mixed, and dispersion-dominated dimers (see Table 1). For all three groups, an agreement between the SAPT-DFT/CBS and $\Delta E_{CCSD(T)}^{CBS}$ data is fairly good and uniform (see Figure S2 and Table S2). The biggest absolute and relative discrepancy between these two data sets is 1.75 kJ/mol and 13.7% exhibited by anisole···CO_2 and the acetylene dimer, respectively, and the root-mean-square deviation is 0.84 kJ/mol, which is much less than 2.05 kJ/mol reported for a similar comparison for the S22 dataset [45] (in that case, smaller basis sets were used). This indicates that all SAPT-DFT/CBS results presented in this work are fully reliable.

Table 1. The DFT-SAPT analysis of intermolecular clusters considered in this work (all values are in kJ/mol, and "1-Nap" stands for 1-naphtol). Parentheses are used in cases when the literature data contain error estimates.

Type of System	Description	Components of the DFT-SAPT Energy					Best Estimate of ΔE
		E_{pol}	E_{exch}	E_{ind}	E_{disp}	E_{total}	
dispersion-dominated complex from Set3x6	aniline:methane	−6.5	18.9	−1.9	−17.3	−6.8	−6.84 [a]
	anisole:methane	−7.1	19.9	−1.6	−18.5	−7.3	−7.39 [a]
	1-Nap:methane	−9.0	25.3	−1.8	−23.4	−8.9	−9.14 [a]
	1-Nap:CO	−9.4	24.0	−3.4	−20.2	−9.0	−8.37 [a]
	1-Nap:CO_2	−13.1	28.6	−3.2	−24.6	−12.3	−12.68 [a]
	anisole:anisole	−32.8	72.8	−8.6	−57.6	−26.1	−27.16 [a]
mixed-interactions complex from Set3x6	anisole:ammonia	−16.1	24.3	−3.7	−15.7	−11.2	−12.00 [a]
	1-Nap:ethyne	−25.1	35.0	−10.1	−17.4	−17.5	−16.96 [a]
	HCl:HCl	−11.3	17.6	−6.3	−9.2	−9.2	−7.94 [a]
	benzene:water	−13.5	19.2	−5.1	−14.7	−14.0	−13.43 [a]
	anisole:CO_2	−20.5	28.6	−3.5	−18.8	−14.1	−15.86 [a]
	ethyne:ethyne	−9.0	11.6	−2.9	−6.9	−7.3	−6.26 [a]
electrostatics-dominated complex from Set3x6	1-Nap:ammonia	−70.0	80.6	−28.7	−22.8	−40.9	−40.52 [a]
	HCl:water	−41.5	50.7	−17.6	−14.3	−22.8	−22.47 [b]
	HCN:HF	−42.6	42.9	−18.5	−11.8	−30.1	−31.09 [c]
	NCH:FH	−15.4	11.5	−3.4	−5.2	−12.4	−12.34 [c]
	HCN:HCN	−25.2	20.3	−6.8	−7.8	−19.5	−19.83 [d]
	1-Nap:water	−46.2	50.0	−15.8	−16.6	−28.6	−29.86 [a]

Table 1. *Cont.*

Type of System	Description	Components of the DFT-SAPT Energy					Best Estimate of ΔE
		E_{pol}	E_{exch}	E_{ind}	E_{disp}	E_{total}	
furan:toluene stacked complex from reference [48]	configuration #1	−9.0	26.5	−2.9	−29.2	−14.5	−14.43
	configuration #2	−8.9	27.6	−3.1	−29.8	−14.2	−13.94
	configuration #3	−7.3	26.5	−2.8	−29.3	−12.9	−12.62
	configuration #4	−7.7	25.9	−3.1	−28.3	−13.2	−12.82
	configuration #5	−7.6	25.6	−2.8	−27.8	−12.5	−12.06
	configuration #6	−6.7	23.7	−2.6	−25.9	−11.5	−11.00
	configuration #7	−5.3	20.8	−2.2	−23.5	−10.1	−9.65
miscellaneous	anthracene: cyclopropenium	−58.5	81.8	−60.3	−48.3	−85.3	(−89.96 ±0.84) [e]
	pyridine:pyridine	−12.0	28.2	−3.3	−29.8	−16.8	−15.82 [a]
large	α-PHB model	−14.7	17.5	−3.8	−25.1	−26.1	−24.03 [f]
	β-PHB model	−8.1	29.1	−4.0	−28.4	−11.4	−10.23 [f]
	C2C2PD	−37.1	114.3	−12.8	−147.1	−82.7	(−87.82 ±2.51) [g]
		−30.8 [h]	107.2 [h]	−10.7 [h]	−154.3 [h]	−88.6 [h]	
	GCGC	−37.4	104.1	−9.5	−111.9	−54.7	(−56.90 ±1.67) [g]
		−34.6 [h]	97.5 [h]	−8.6 [h]	−115.1 [h]	−60.9 [h]	
	GGG	11.3	27.8	−6.2	−39.9	−6.9	(−8.79 ±0.84) [g]
		12.0 [h]	25.7 [h]	−5.6 [h]	−41.0 [h]	−8.9 [h]	

[a] CCSD(T)/CBS value obtained using Equation (1) in this work; [b] obtained as described at page 7109 of reference [49]; [c] obtained as described at page 3 of reference [50]; [d] obtained as described in Table 8 of reference [51]; [e] obtained as described in Table 1 of reference [16]; [f] DLPNO-CCSD(T)/CBS value obtained using Equation (2) in this work; [g] obtained as described in Table 1 of reference [43]; [h] obtained as described in the supporting information to reference [52].

2.2. Comparing the Canonical and DLPNO-Based CCSD(T) Data

As already mentioned, for larger complexes, the canonical $\Delta E_{CCSD(T)}^{CBS}$ computations would be impractical, and a local electron-correlation scheme, such as the DLPNO method, would need to be applied. Using the DLPNO approximation, there are two ways of estimating the $\Delta E_{CCSD(T)}^{CBS}$ value. In the first approach, the DLPNO-CCSD(T) energies are obtained for a series of some correlation-consistent basis sets and extrapolated to the CBS limit. This approach may very quickly reach computational restrictions if applied to large systems. Nevertheless, it was used for the 49 dimers described in the preceding part, and its results are used for comparison purposes (see below). The second approach to the $\Delta E_{CCSD(T)}^{CBS}$ estimation applies a composite scheme that is analogous to the focal point analysis from Section 2.1. The present composite approach is expressed by Equation (2) (the right arrow indicates an extrapolation of the respective energy term to its CBS limit by applying the two-point formula from reference [53]). It is implemented in the Excel spreadsheet 'energies2.xlsx' (see Supporting Information), where all the underlying absolute energies can be found.

$$\Delta E_{CCSD(T)}^{CBS} = \Delta E_{HF}^{aQZ} + \Delta E_{MP2}^{aTZ \rightarrow aQZ} + \Delta E_{post-MP2}^{aTZ \rightarrow aQZ} \tag{2}$$

It should be noted that the $\Delta E_{post-MP2}$ term was taken as a difference of the pertinent DLPNO-CCSD(T) and DLPNO-MP2 [54] correlation energies. These energies were obtained with tight thresholds of the DLPNO approximation (see Section 4 for details). A contribution of triple excitations to the correlation energy was approximated by the non-iterative calculations, which are sometimes denoted as (T_0), instead of using the iterative scheme that would provide results denoted as (T_1) [24]. This choice was made on the basis of a significant increase of computational time of the DLPNO-CCSD(T_1) calculations with respect to their DLPNO-CCSD(T_0) counterparts, and by negligible differences between the two sets of results for dimers from the Set3x6 (see Figure S3 and Table S3).

The $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ data obtained using Equations (1) and (2) for the aforementioned set of 49 complexes span a large interval of values from ca. 2 to ca. 89 kJ/mol, and are listed in Table S4. Clearly, the DLPNO-based interaction energies closely match their canonical counterparts (see Figure 1). The maximum absolute difference between these data points is 2.29 kJ/mol exhibited by the stacked adenine⋯thymine (AT) pair. It should be noted that there is also a small uncertainty in the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ values (see a related discussion for the S22 set in reference [55]). The linear regression is {y} = 0.9971 × {x} + 0.3766 kJ/mol (shorthand notation is used with {y} for the DLPNO-based and {x} for the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ data, respectively), with adjusted $R^2 = 0.9993$ and the standard deviation of 0.59 kJ/mol. The maximum residual of this fit is 1.77 kJ/mol, and expectedly occurs for the stacked AT pair. However, such discrepancy amounts to only ca. 3.8% of the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ value of ca. 48.6 kJ/mol obtained for this complex. The highest relative error is found for the stacked indole⋯benzene dimer. In this case, the residual is 1.69 kJ/mol, which is ca. 8.9% of the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ value of ca. 19.0 kJ/mol (see Table S4).
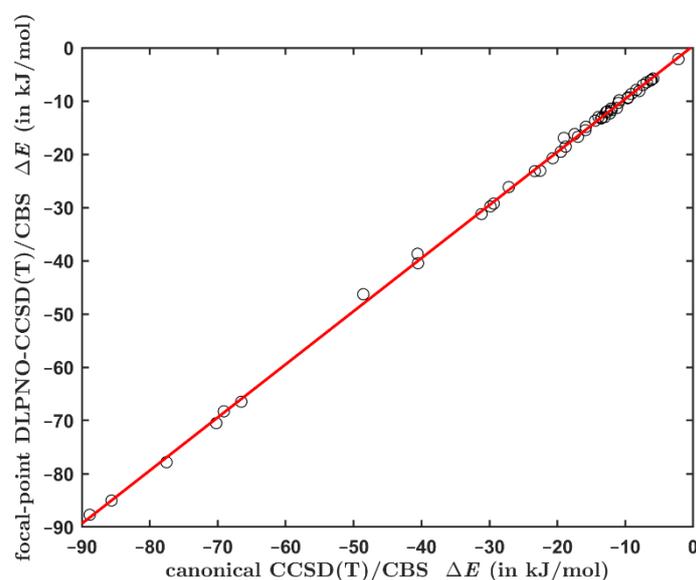

Figure 1. Comparison of the interaction energies computed for the set of 49 dimers by an application of protocols from Equations (1) and (2). The regression line is specified in the text and shown in red.

The interaction energies computed using the focal point method from Equation (2) were also checked against yet another set of results. Namely, the total DLPNO-CCSD(T) energies were calculated while employing the aTZ; aQZ; a5Z series of basis sets, and extrapolated to their CBS limit using the mixed Gaussian/exponential form from reference [54]. The underlying energies are provided in the Excel spreadsheet 'energies3.xlxs' together with an analytical solution to the set of three equations from reference [56], which is used to compute the pertinent $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ value (see Supporting Information). Figure 2 presents an excellent agreement between the two sets of DLPNO-based CCSD(T)/CBS interaction energies (their values are collected in Table S4). This result shows that there is only a negligible basis set incompleteness error in the $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ results obtained by the focal point method expressed by Equation (2). As a consequence, the DLPNO-CCSD(T)/a5Z calculations should not be needed for an accurate estimation of the $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ data.
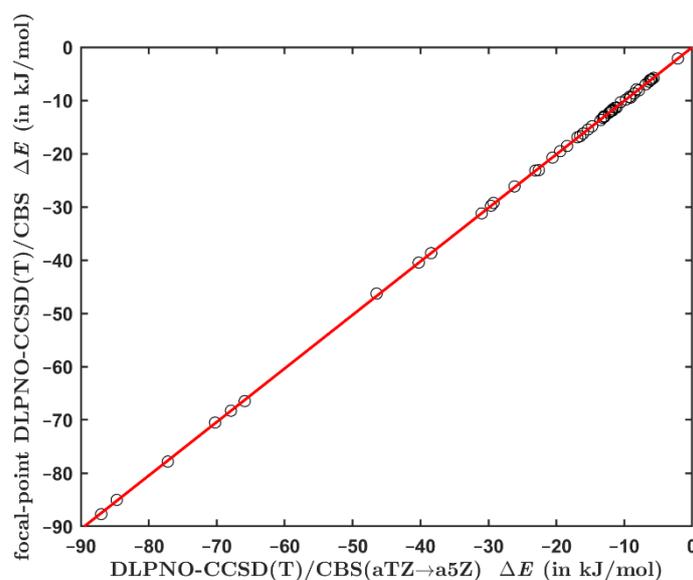
Figure 2. Comparison of the interaction energies computed for the set of 49 dimers by an application of the procedure from reference [56] and the protocol from Equation (2). The regression line is shown in red.

2.3. The Fittings Scheme for Smaller Basis Sets

While the composite procedure from the previous part (Equation (2)) is clearly successful in reliably predicting the interaction energies, it also requires results obtained using the aQZ basis set, which would be impractical for very large systems. Hence, an attempt was made to devise some less costly computational protocol. It is stressed that a direct application of Equation (2) together with smaller basis sets (aDZ and aTZ, for instance) cannot be expected to lead to an accurate estimate of the $\Delta E_{CCSD(T)}^{CBS}$ [57]. Instead, all three terms on the right-hand side of Equation (2) need to be extrapolated to their CBS limits, which in the following will be designated ΔE_{HF}^{CBS*}, ΔE_{MP2}^{CBS*}, and $\Delta E_{post-MP2}^{CBS*}$, in order to assume their sum, $\Delta E_{CCSD(T)}^{CBS*}$, to be close to a true value of the $\Delta E_{CCSD(T)}^{CBS}$ term [41]. A particular attention has to be paid to the choice of respective extrapolation schemes for ΔE_{HF}^{CBS*}, ΔE_{MP2}^{CBS*}, and $\Delta E_{post-MP2}^{CBS*}$ contributions, as each of the underlying energy components converges differently with increasing the basis set size and quality (see the most recent study [58] and references cited therein). For an extrapolation of the HF and correlation energies, in frequent use are the exponents tabulated in reference [59] that were obtained by fitting the absolute energies of 21 small molecules computed using a number of pairs of various basis sets. The pertinent exponents from Table 3 of reference [59] were applied to extrapolate the energies computed using the (aDZ, aTZ) basis sets for systems investigated here, but this led to inaccurate ΔE values (not shown) in some cases. Hence, all relevant energies for S22 dataset were obtained (58 values for each component of the total energy and basis set, because of the symmetry in 7 out of 22 clusters) and fitted as follows. For the HF energy, the coefficient α minimizes (in the least-squares sense) differences between E_{HF}^{a5Z} data and the functional form given by Equation (3):

$$E_{HF}^{fit}\left(\alpha; \ E_{HF}^{aDZ}, E_{HF}^{aTZ}\right) = \frac{\exp\left(\alpha\sqrt{3}\right)E_{HF}^{aDZ} - \exp\left(\alpha\sqrt{2}\right)E_{HF}^{aTZ}}{\exp\left(\alpha\sqrt{3}\right) - \exp\left(\alpha\sqrt{2}\right)} \tag{3}$$

The fit gives $\alpha = -4.473$ (rounding to four digits is performed on the basis of an estimated covariance that is not shown). The correlation energies were treated in an analogous way. Specifically, an optimal value of the coefficient β for the MP2 correlation

energy contribution was obtained by minimizing differences between $E_{\text{MP2}}^{\text{a5Z}}$ data and their $E_{\text{MP2}}^{\text{fit}}$ counterparts from Equation (4):

$$E_{\text{MP2}}^{\text{fit}}\left(\beta;\ E_{\text{MP2}}^{\text{aDZ}}, E_{\text{MP2}}^{\text{aTZ}}\right) = \frac{E_{\text{MP2}}^{\text{aDZ}}2^{\beta} - E_{\text{MP2}}^{\text{aTZ}}3^{\beta}}{2^{\beta} - 3^{\beta}} \tag{4}$$

This value is $\beta = 2.796$. As for the post-MP2 correlation energy term, numerical tests revealed that yet another coefficient would be needed, which is denoted as γ and minimizes differences between $\Delta E_{\text{post}-\text{MP2}}^{\text{aTZ}}$ data and the functional form expressed by Equation (5):

$$E_{\text{post}-\text{MP2}}^{\text{fit}}\left(\gamma;\ E_{\text{post}-\text{MP2}}^{\text{aDZ}}, E_{\text{post}-\text{MP2}}^{\text{aTZ}}, \beta\right) = \frac{2^{\gamma} E_{\text{CCSD(T)}}^{\text{aDZ}} - 3^{\gamma} E_{\text{CCSD(T)}}^{\text{aTZ}}}{2^{\gamma} - 3^{\gamma}} - \frac{2^{\beta} E_{\text{MP2}}^{\text{aDZ}} - 3^{\beta} E_{\text{MP2}}^{\text{aTZ}}}{2^{\beta} - 3^{\beta}} \tag{5}$$

An optimal value of this coefficient is found to be $\gamma = 2.741$ for the parameter β kept constant at $\beta = 2.796$ (see above). The data sets that were actually used for fitting are included in Supporting Information. For the testing set of aforementioned 27 dimers, the $\Delta E_{\text{CCSD(T)}}^{\text{CBS}*}$ values were obtained through Equations (3)–(5) and compared to their $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ counterparts, which are described in Section 2.1. The data are collected in Table S4 and graphically presented in Figure 3, and illustrate a good performance of present fitting scheme for cost-effective estimation of the ΔE. In Figure 3 the highest absolute and relative differences are marked. They amount to 1.72 kJ/mol and 17.9%, respectively, and are accordingly exhibited by the cyclopropenium cation$\cdots$anthracene complex with large interaction energy of almost –90 kJ/mol, and by the highly challenging configuration of furan$\cdots$toluene dimer (see Discussion). The linear regression model is $\Delta E_{\text{CCSD(T)}}^{\text{CBS}*} = 0.9987 \times \Delta E_{\text{CCSD(T)}}^{\text{CBS}} + 0.3789$ kJ/mol with adjusted $R^2 = 0.9998$ and the standard deviation of 0.45 kJ/mol. However, it should be mentioned that if the same testing set is treated using the procedure from Equation (2) that employs also the aQZ data, a significantly better agreement between the two sets of $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ values (namely, those given by Equations (1) and (2)) is obtained. Specifically, the highest absolute and relative differences become as low as 0.71 kJ/mol and 4.2%, respectively (they occur for the same systems as in the case of an application of the computationally cheaper procedure). The procedure expressed by Equation (2) should thus be used whenever permitted by the size of an investigated system.
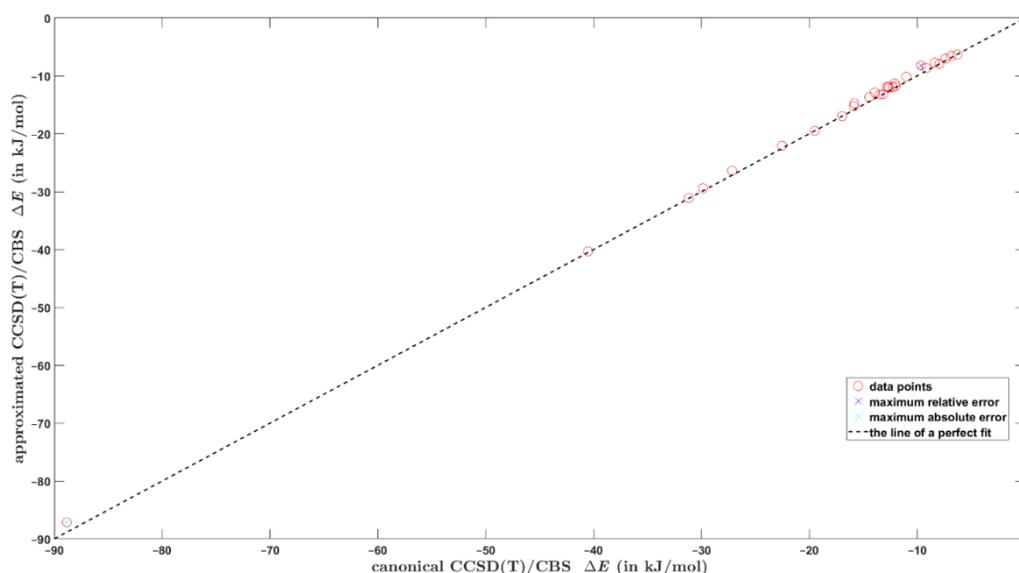

Figure 3. Comparison of the interaction energies computed for the testing set of 27 dimers by two methods that are discussed in the text. The raw data are available from Table S5.

2.4. Testing Large Systems

The procedure from Section 2.3 was of course devised with applications to extended systems in mind. Thus, it needs to be validated by checking its accuracy also for intermolecular complexes that are significantly larger than those listed in Table S4, because the local approximation error growths with the system size (see reference [60] and work cited therein), and this problem might be exacerbated by using relatively small (aDZ, aTZ) basis sets. Since the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ data are not available, the SAPT-DFT/CBS calculations were applied to obtain reference values of the interaction energy for models of two polymer chains (see Materials and Methods). Additionally, the DLPNO-based focal point method from Section 2.2 was used for comparison purposes. The same methods were also applied to three challenging systems from the L7 dataset. Namely, C2C2PD and GCGC were chosen due to known differences in their ΔE values as obtained from the CCSD(T) and QMC computations [43], while GGG was included because of an exceedingly high amount of the dispersion contribution to the stabilization of this complex, leading to the dispersion-to-polarization ratio of about six (see Table 1). Table 2 summarizes results provided by the DLPNO-based methods together with their SAPT-DFT counterparts (the benchmark data for C2C2PD, GCGC and GGG from reference [43] are shown in Table 1 together with their error bars). It is evident that there are no apparent outliers among these results. Figure 4 compares the best estimates of the ΔE values to their counterparts obtained using the cost-effective procedure expressed by Equations (3)–(5). Namely, results from reference [43] are employed in the ordinate for C2C2PD, GCGC, and GGG. For the models of PHB polymorphs, interaction energies obtained by an application of Equation (2) are used together with an uncertainty estimate of ± 2 kJ/mol, which is discussed in the subsequent part.

Table 2. The negative of interaction energies (in kJ/mol) extrapolated to the complete basis set limit for large complexes.

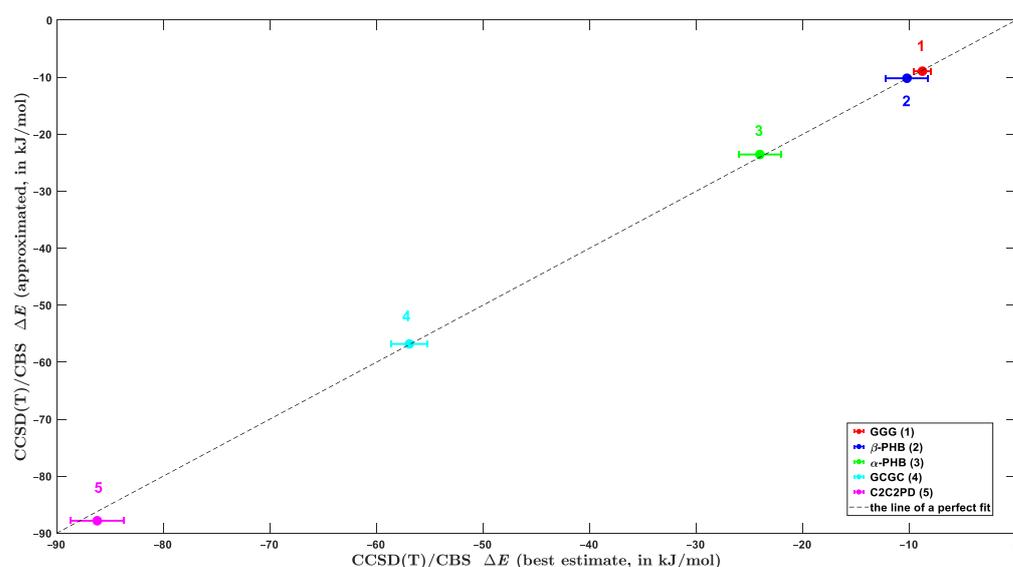| Method	α-PHB	β-PHB	GGG	GCGC	C2C2PD
extrapolations using Equations (3)–(5)	23.56	10.18	8.95	56.79	87.81
the focal point analysis using Equation (2)	24.03	10.23	8.87	56.44	86.19
SAPT-DFT	26.14	11.43	6.88	54.69	82.68

Figure 4. Comparison of the interactions energies of five large complexes that are discussed in the text. The error bars that are plotted for systems number 1, 4, and 5 are given in Table 1. For systems number 2 and 3, the error bars of ± 2 kJ/mol are plotted.

3. Discussion

The DLPNO-based methods from Sections 2.2 and 2.3 showed a good performance in predicting absolute values of the CBS-extrapolated binding energies for a variety of molecular clusters. In particular, if the (aTZ, aQZ) data were used in approach expressed by Equation (2), the ensuing $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ should lie within about 2.0 kJ/mol of a true absolute value of the interaction energy. Thus, they would be expected to lead to the ΔE result of at least "bronze standard" quality [61]. Moreover, an application of the refitted exponents (Equations (3)–(5)) to the (aDZ, aTZ) data was found to work remarkably well for larger systems (see Table 2). Nevertheless, in areas of, for instance, conformational analysis (see the most recent investigation [62] and references cited therein), computer-assisted rotational spectroscopy [63], developing new structural descriptors [48], or for anticipated applications in modeling of polymers, the relative energies of various configurations of an investigated system also need to be obtained with a high accuracy. A stringent test was performed here using seven stacked orientations of the furan···toluene dimer (they are numbered as they consecutively appear in the supporting information to reference [48]). It can be verified by a visual inspection that the complexes are fairly different from each other. However, their interaction energies span a narrow interval of ca. 4 kJ/mol (see Figure 5; raw values as provided by the respective methods are available from Table S6). As follows from the SAPT energy decomposition (see Table 1), the dispersion interaction completely dominates the intermolecular bonding in these clusters. It should be noted that in one case (namely, for the configuration #7) the dispersion-to-polarization ratio is as high as 4.5. This is the structure with highest differences between the DLPNO-based and canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ results described in Section 2.3. Four computational approaches were applied to these furan···toluene dimers in addition to the benchmarking canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ calculations, and key statistical parameters describing the level of agreement between the relevant datasets are shown in Table S7. The $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ results that were obtained from the focal-point analysis of the (aTZ, aQZ) data are designated 'DLPNO-CCSD(T)/CBS' in Figure 5. They are accurate in both absolute and relative terms, as are the SAPT-DFT/CBS values for this challenging set of complexes. An inspection of Figure 5 reveals that the ΔE results that were not extrapolated to the CBS limit, namely, the DLPNO-CCSD(T)/cc-pVQZ data from reference [48], are incorrectly ordered. Moreover, they are heavily shifted with respect to the canonical $\Delta E_{\text{CCSD(T)}}^{\text{CBS}}$ values (on average by 2.88 kJ/mol), but this shift was already discussed in reference [48]. The ordering of interaction energies estimated through Equations (3)–(5) is also incorrect (see Figure 5), and a significant underestimation of the ΔE occurs for the aforementioned configuration #7. These results for the furan···toluene dimers indicate that the relative accuracy of the scheme expressed by Equation (2) should be around a half of kJ/mol. Consequently, this approach is well-suited for checking results of less demanding methods, for example, variants of DFT that were tailored for noncovalent interactions [64]. The computationally much cheaper procedure that applies Equations (3)–(5) can be expected to have the relative accuracy of about one kJ/mol. This level of accuracy is unlikely to be limiting in applications to modeling large systems, though, as for them there would probably be more significant uncertainties due to the geometry and media effects [65]. Nevertheless, further testing of both procedures (those expressed by Equation (2) and by Equations (3)–(5)) is desirable.
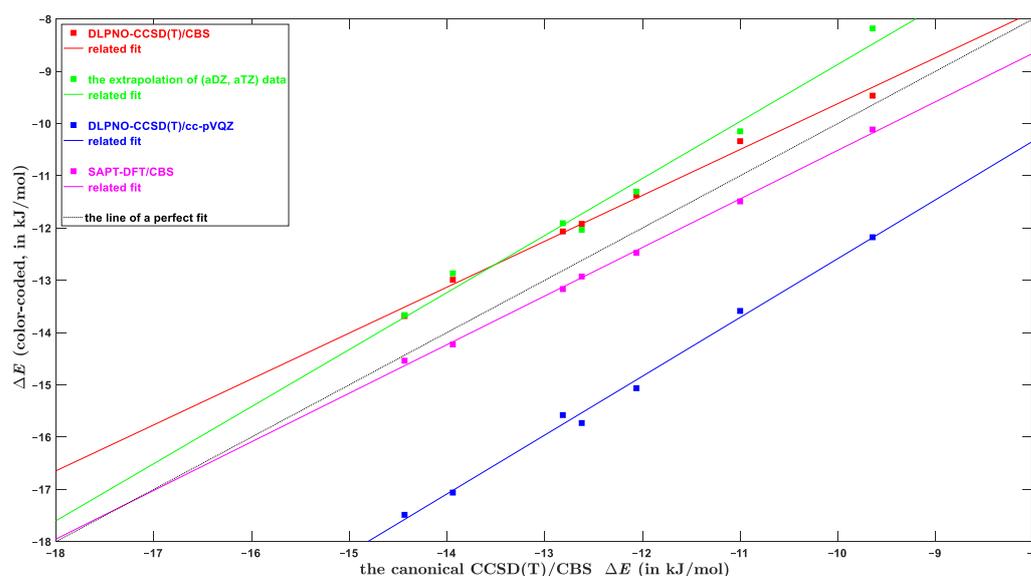
Figure 5. Comparison of the interaction energies of seven configurations of furan⋯toluene dimer.

4. Materials and Methods

The following 11 dimers were considered in their MP2/aTZ geometry from reference [66]: aniline⋯methane, anisole⋯methane, 1-naphthol⋯methane, 1-naphthol⋯CO, 1-naphthol⋯CO_2, anisole⋯anisole, anisole⋯ammonia, 1-naphtol⋯acetylene, HCl⋯HCl, benzene⋯water, 1-naphtol⋯ammonia, and HCl⋯water. The MP2/aTZ geometries, which are provided in 'geometries.tar' file in Supporting Information, were obtained for the following 7 dimers: anisole⋯CO_2 (the optimization started from a structure analogous to the one showed in Figure 1 of reference [67]), acetylene⋯acetylene in the T-shaped configuration (see reference [68]), HCN⋯HF (see reference [50]), NCH⋯FH (see reference [50]), HCN⋯HCN (see reference [49]), and 1-naphthol⋯water (the optimization started from a structure featuring a classical hydrogen bond between the hydroxyl oxygen of 1-naphtol and one of the protons of water). These 18 structures form the aforementioned 'Set3x6' testing suite. Each structure optimized at the MP2/aTZ level was verified to be a minimum of the potential energy surface by inspecting the predicted harmonic vibrational frequencies. Pertinent calculations were performed using the Gaussian 16, revision C.01 suite of codes [69] with default settings.

The dimeric models of the α and β polymorphs of PHB were prepared using coordinates from references [36,37], respectively. Their structures are included in 'structures.zip' file in Supporting Information.

The configurations 1–7 of the furan⋯toluene dimer were taken from 'ja9b00936_si_002.xyz' file of supporting materials to reference [48]. Coordinates of the S22 dataset, the stacked pyridine⋯pyridine, C2C2PD, GCGC, and GGG were downloaded from the BEGDB website [70]. Coordinates of the cyclopropenium⋯anthracene dimer were taken from supporting materials to reference [16].

The SAPT-DFT, canonical CCSD(T) and MP2 calculations were carried out in Molpro 2021.2 [71]. The same procedures and notation as in our most recent work [66] were used for the SAPT-DFT analysis.

The MP2/a5Z energies for an application of Equation (1) were obtained in the resolution-of-the-identity integral approximation [72,73] while using the relevant auxiliary basis sets [74]. The program package Turbomole, version 7.1 [75] was used for these and for related HF/a5Z calculations.

The DLPNO-based computations were carried out in ORCA 5.0 [76]. The underlying HF calculations used 'VeryTightSCF' accuracy settings. The default 'augmented Hessian Foster–Boys' localization scheme was adopted. The truncation of the electron-correlation space was performed by applying the 'TightPNO' set of parameters.

The Levenberg–Marquardt algorithm as implemented in 'lsqcurvefit' function of MATLAB® Optimization Toolbox™ was used to perform the nonlinear fitting of models expressed by Equations (3)–(5). The values of α, β, γ parameters thus found (see Equations (3)–(5)) were checked using 'e04fcf' routine of the NAG® Fortran Library; the respective data files are included in Supporting Information.

References

1. Bernstein, E. *Intra- and Intermolecular Interactions between Non-covalently Bonded Species*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2020.
2. Jin, M.Y.; Zhen, Q.; Xiao, D.; Tao, G.; Xing, X.; Yu, P.; Xu, C. Engineered non-covalent π interactions as key elements for chiral recognition. *Nat. Commun.* **2022**, *13*, 3276. [CrossRef] [PubMed]
3. Jiao, Y.; Chen, X.-Y.; Stoddard, J.F. Weak bonding strategies for achieving regio- and site-selective transformations. *Chem* **2022**, *8*, 414–438. [CrossRef]
4. Jena, S.; Dutta, J.; Tulsiyan, K.D.; Sahu, A.K.; Choudhury, S.S.; Biswal, H.S. Noncovalent interactions in proteins and nucleic acids: Beyond hydrogen bonding and π-stacking. *Chem. Soc. Rev.* **2022**, *51*, 4261–4286. [CrossRef] [PubMed]
5. Puzzarini, C.; Spada, L.; Alessandrini, S.; Barone, V. The challenge of non-covalent interactions: Theory meets experiment for reconciling accuracy and interpretation. *J. Phys. Condens. Matter.* **2020**, *32*, 343002. [CrossRef] [PubMed]
6. Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2020**, *10*, e1452. [CrossRef]
7. Shahbaz, M.; Szalewicz, K. Evaluation of methods for obtaining dispersion energies used in density functional calculations of intermolecular interactions. *Theor. Chem. Acc.* **2019**, *138*, 25. [CrossRef]
8. Carter-Fenk, C.; Lao, K.U.; Herbert, J.M. Predicting and Understanding Non-Covalent Interactions Using Novel Forms of Symmetry-Adapted Perturbation Theory. *Acc. Chem. Res.* **2021**, *54*, 3679–3690. [CrossRef]
9. Sharapa, D.I.; Margraf, J.T.; Hesselmann, A.; Clark, T. Accurate Intermolecular Potential for the C60 Dimer: The Performance of Different Levels of Quantum Theory. *J. Chem. Theory Comput.* **2017**, *13*, 274–285. [CrossRef]
10. Szalewicz, K.; Jeziorski, J. Physical mechanisms of intermolecular interactions from symmetry-adapted perturbation theory. *J. Molec. Model.* **2022**, *28*, 273. [CrossRef]
11. Calvin, J.A.; Peng, C.; Rishi, V.; Kumar, A.; Valeev, E.F. Many-Body Quantum Chemistry on Massively Parallel Computers. *Chem. Rev.* **2021**, *121*, 1203–1231. [CrossRef]
12. Burns, L.A.; Faver, J.C.; Zheng, Z.; Marshall, M.S.; Smith, D.G.A.; Vanommeslaeghe, K.; MacKerrel, A.D.; Merz, K.M.; Sherrill, C.D. The BioFragment Database (BFDb): An open-data platform for computational chemistry analysis of noncovalent interactions. *J. Chem. Phys.* **2017**, *147*, 161727. [CrossRef]
13. Zhang, I.Y.; Grüneis, A. Coupled Cluster Theory in Materials Science. *Front. Mater.* **2019**, *6*, 123. [CrossRef]
14. Kříž, K.; Nováček, M.; Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets 3: Repulsive Contacts. *J. Chem. Theory Comput.* **2021**, *17*, 1548–1561. [CrossRef]
15. Kříž, K.; Řezáč, J. Non-covalent interactions atlas benchmark data sets 4: σ-hole interactions. *Phys. Chem. Chem. Phys.* **2022**, *24*, 14794–14804. [CrossRef]

16. Spicher, S.; Caldeweyher, E.; Hansen, A.; Grimme, S. Benchmarking London dispersion corrected density functional theory for noncovalent ion–π interactions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 11635–11648. [CrossRef]

17. Huang, H.-H.; Wang, Y.-S.; Chao, S.D. A Minimum Quantum Chemistry CCSD(T)/CBS Data Set of Dimeric Interaction Energies for Small Organic Functional Groups: Heterodimers. *ACS Omega* **2022**, *7*, 20059–20080. [CrossRef]

18. Czernek, J.; Brus, J. Parametrizing the Spatial Dependence of 1H NMR Chemical Shifts in π-Stacked Molecular Fragments. *Int. J. Mol. Sci.* **2020**, *21*, 7908. [CrossRef]

19. Gyevi-Nagy, L.; Kállay, M.; Nagy, P.R. Integral-direct and parallel implementation of the CCSD(T) method: Algorithmic developments and large-scale applications. *J. Chem. Theory Comput.* **2020**, *16*, 366–384. [CrossRef]

20. Nagy, P.R.; Gyevi-Nagy, L.; Lőrincz, B.D.; Kállay, M. Pursuing the bases set limit of CCSD(T) non-covalent interaction energies for medium-sized complexes: Case study on the S66 compilation. *Mol. Phys.* **2022**, e2109526. [CrossRef]

21. Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106. [CrossRef]

22. 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. [CrossRef] [PubMed]

23. Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E.F.; Neese, F. Sparse maps–A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, 024109. [CrossRef] [PubMed]

24. Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D.G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101. [CrossRef] [PubMed]

25. Guo, Y.; Riplinger, C.; Liakos, D.G.; Becker, U.; Saitow, M.; Neese, F. Linear scaling perturbative triples correction approximations for open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory [DLPNO-CCSD(T0/T)]. *J. Chem. Phys.* **2020**, *152*, 024116. [CrossRef]

26. Liakos, D.G.; Sparta, M.; Kesharwani, M.K.; Martin, J.M.L.; Neese, F. Exploring the Accuracy Limits of Local Pair Natural Orbital Coupled-Cluster Theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539. [CrossRef]

27. Liakos, D.G.; Guo, Y.; Neese, F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *J. Phys. Chem. A* **2020**, *124*, 90–100. [CrossRef]

28. Chen, J.-L.; Sun, T.; Wang, Y.-B.; Wang, W. Toward a less costly but accurate calculation of the CCSD(T)/CBS noncovalent interaction energy. *J. Comput. Chem.* **2020**, *41*, 1252–1260. [CrossRef]

29. Beck, M.E.; Riplinger, C.; Neese, F. Unraveling individual host–guest interactions in molecular recognition from first principles quantum mechanics: Insights into the nature of nicotinic acetylcholine receptor agonist binding. *J. Comput. Chem.* **2021**, *42*, 293–302. [CrossRef]

30. Villot, C.; Ballesteros, F.; Wang, D.; Lao, K.U. Coupled Cluster Benchmarking of Large Noncovalent Complexes in L7 and S12L as Well as the C60 Dimer, DNA–Ellipticine, and HIV–Indinavir. *J. Phys. Chem. A* **2022**, *126*, 4326–4341. [CrossRef]

31. Sandler, I.; Chen, J.; Taylor, M.; Sharma, S.; Ho, J. Accuracy of DLPNO-CCSD(T): Effect of Basis Set and System Size. *J. Phys. Chem. A* **2021**, *125*, 1553–1563. [CrossRef]

32. Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Sponer, J. Quantum Chemical Benchmark Study on 46 RNA Backbone Families Using a Dinucleotide Unit. *J. Chem. Theory Comput.* **2015**, *11*, 4972–4991. [CrossRef]

33. Altun, A.; Neese, F.; Bistoni, G. Extrapolation to the Limit of a Complete Pair Natural Orbital Space in Local Coupled-Cluster Calculations. *J. Chem. Theory Comput.* **2020**, *16*, 6142–6149. [CrossRef]

34. East, A.L.L.; Allen, D.L. The heat of formation of NCO. *J. Chem. Phys.* **1993**, *99*, 4638–4650. [CrossRef]

35. Jurečka, P.; Šponer, J.; Černý, 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. [CrossRef]

36. Wang, H.; Tashiro, K. Reinvestigation of Crystal Structure and Intermolecular Interactions of Biodegradable Poly(3-Hydroxybutyrate) α-Form and the Prediction of Its Mechanical Property. *Macromolecules* **2016**, *49*, 581–594. [CrossRef]

37. Phongtamrug, S.; Tashiro, K. X-ray Crystal Structure Analysis of Poly(3-hydroxybutyrate) β-Form and the Proposition of a Mechanism of the Stress-Induced α-to-β Phase Transition. *Macromolecules* **2019**, *52*, 2995–3009. [CrossRef]

38. 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. [CrossRef]

39. Kent, P.R.C.; Annaberdiyev, A.; Benali, A.; Bennett, M.C.; Landinez Borda, E.J.; Doak, P.; Hao, H.; Jordan, K.D.; Krogel, J.T.; Kylänpää, I.; et al. QMCPACK: Advances in the development, efficiency, and application of auxiliary field and real-space variational and diffusion quantum Monte Carlo. *J. Chem. Phys.* **2020**, *152*, 174105. [CrossRef]

40. Benali, A.; Shin, H.; Heinonen, O. Quantum Monte Carlo benchmarking of large noncovalent complexes in the L7 benchmark set. *J. Chem. Phys.* **2020**, *153*, 194113. [CrossRef]

41. Morales-Silva, M.A.; Jordan, K.D.; Shulenburger, L.; Wagner, L.K. Frontiers of stochastic electronic structure calculations. *J. Chem. Phys.* **2021**, *154*, 170401. [CrossRef]

42. Ballesteros, F.; Dunivan, S.; Lao, K.U. Coupled cluster benchmarks of large noncovalent complexes: The L7 dataset as well as DNA–ellipticine and buckycatcher–fullerene. *J. Chem. Phys.* **2021**, *154*, 154104. [CrossRef] [PubMed]

43. Al-Hamdani, Y.S.; Nagy, P.R.; Zen, A.; Barton, D.; Kállay, M.; Bradenburg, J.G.; Tchatkenko, A. Interactions between large molecules pose a puzzle for reference quantum mechanical methods. *Nat. Commun.* **2021**, *12*, 3927. [CrossRef] [PubMed]

44. 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. [CrossRef]

45. 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. [CrossRef]

46. Marshall, M.S.; Burns, L.A.; Sherrill, C.D. Basis set convergence of the coupled-cluster correction: 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. [CrossRef] [PubMed]

47. 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. [CrossRef]

48. Bootsma, A.N.; Doney, A.C.; Wheeler, S.E. Predicting the Strength of Stacking Interactions between Heterocycles and Aromatic Amino Acid Side Chains. *J. Am. Chem. Soc.* **2019**, *141*, 11027–11035. [CrossRef]

49. Liu, Y.; Li, J.; Felker, P.M.; Bačić, Z. HCl–H2O dimer: An accurate full-dimensional potential energy surface and fully coupled quantum calculations of intra- and intermolecular vibrational states and frequency shifts. *Phys. Chem. Chem. Phys.* **2021**, *23*, 7101–7114. [CrossRef]

50. Sexton, T.M.; Van Benschoten, W.Z.; Tschumper, G.S. Dissociation energy of the HCN···HF dimer. *Chem. Phys. Lett.* **2020**, *748*, 137382. [CrossRef]

51. Hoobler, P.R.; Turney, J.M.; Agarwal, J.; Schaefer III, J.F. Fundamental Vibrational Analyses of the HCN Monomer, Dimer and Associated Isotopologues. *ChemPhysChem* **2018**, *19*, 3257–3265. [CrossRef]

52. Carter-Fenk, K.; Lao, K.U.; Liu, K.-Y.; Herbert, J.M. Accurate and Efficient ab Initio Calculations for Supramolecular Complexes: Symmetry-Adapted Perturbation Theory with Many-Body Dispersion. *J. Phys. Chem. Lett.* **2019**, *10*, 2706–2714. [CrossRef]

53. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A.K. Basis-set convergence in correlated calculations on Ne, N2, and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243–252. [CrossRef]

54. Pinski, P.; Riplinger, C.; Valeev, E.F.; Neese, F. Sparse maps–A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108. [CrossRef]

55. Vogiatzis, K.D.; Klopper, W. Accurate non-covalent interactions with basis-set corrections from interference-corrected perturbation theory: Comparison with the S22B database. *Mol. Phys.* **2013**, *111*, 2299–2305. [CrossRef]

56. Peterson, K.A.; Woon, D.E.; Dunnig, T.J., Jr. Benchmark calculations with correlated wave functions. *J. Chem. Phys.* **1994**, *100*, 7410–7415. [CrossRef]

57. Takatani, T.; Hohenstein, E.G.; Malagoli, M.; Marshall, M.C.; Sherrill, C.D. Basis set consistent revision of the S22 test set of noncovalent interaction energies. *J. Chem. Phys.* **2010**, *132*, 144104. [CrossRef]

58. Ye, H.-Z.; Berkelbach, T.C. Correlation-Consistent Gaussian Basis Sets for Solids Made Simple. *J. Chem. Theory Comput.* **2022**, *18*, 1595–1606. [CrossRef]

59. 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. [CrossRef]

60. Altun, A.; Ghosh, S.; Riplinger, C.; Neese, F.; Bastoni, G. Addressing the System-Size Dependence of the Local Approximation Error in Coupled-Cluster Calculations. *J. Phys. Chem. A* **2021**, *125*, 9932–9939. [CrossRef]

61. Kesharwani, M.K.; Karton, M.; Sylvetsky, N.; Martin, J.M.L. The S66 Non-Covalent Interactions Benchmark Reconsidered Using Explicitly Correlated Methods Near the Basis Set Limit. *Austr. J. Chem.* **2018**, *71*, 238–248. [CrossRef]

62. Ehlers, S.; Grimme, S.; Hansen, A. Conformational Energy Benchmark for Longer n-Alkane Chains. *J. Phys. Chem. A* **2022**, *126*, 3521–3535. [CrossRef]

63. Li, X.; Spada, L.; Alessandrini, S.; Zheng, Y.; Lengsfeld, K.G.; Grabow, J.-U.; Feng, G.; Puzzarini, C.; Barone, V. Stacked but not Stuck: Unveiling the Role of $\pi \rightarrow \pi^*$ Interactions with the Help of the Benzofuran–Formaldehyde Complex. Angew. *Chem. Int. Ed.* **2022**, *61*, 264–270. [CrossRef] [PubMed]

64. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215. [CrossRef] [PubMed]

65. Al-Hamdani, Y.S.; Tkatchenko, A. Understanding non-covalent interactions in larger molecular complexes from first principles. *J. Chem. Phys.* **2019**, *150*, 010901. [CrossRef]

66. Czernek, J.; Brus, J.; Czerneková, V. A computational inspection of the dissociation energy of mid-sized organic dimers. *J. Chem. Phys.* **2022**, *156*, 204303. [CrossRef] [PubMed]

67. Becucci, M.; Mazzoni, F.; Pietraperzia, G.; Řezáč, J.; Nachtigallová, D.; Hobza, P. Non-covalent interactions in anisole–(CO2)n (n = 1, 2) complexes. *Phys. Chem. Chem. Phys.* **2017**, *19*, 22749–22758. [CrossRef]

68. Leforestier, C.; Tekin, A.; Jansen, G.; Herman, M. First principles potential for the acetylene dimer and refinement by fitting to experiments. *J. Chem. Phys.* **2011**, *135*, 234306. [CrossRef]

69. Frish, M.J.; Trucks, J.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16*; Revision C.01; Gaussian, Inc.: Wallingford, CT, USA, 2019.

70. The Benchmark Energy & Geometry Database (BEGDB). Available online: http://www.begdb.org/ (accessed on 5 October 2022).
71. Werner, H.J.; Knowles, P.J.; Manby, F.R.; Black, J.A.; Doll, K.; Hesselmann, A.; Kats, D.; Kohn, A.; Korona, T.; Kreplin, D.A.; et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 144107. [CrossRef]
72. Vahtras, O.; Almlöf, J.; Feyereisen, M.W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518. [CrossRef]
73. Weigend, F.; Häser, M. RI-MP2: First derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340. [CrossRef]
74. Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152. [CrossRef]
75. Balasubramani, S.G.; Chen, G.P.; Coriani, S.; Diedenhofen, M.; Frank, M.S.; Franzke, Y.J.; Furche, F.; Grotjahn, R.; Harding, M.E.; Hättig, C.; et al. TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, 184107. [CrossRef]
76. Neese, F. Software update: The ORCA program system–Version 5.0. Wiley Interdiscip. *Rev. Comput. Mol. Sci.* **2022**, *12*, e1606. [CrossRef]