The CHAL336 Benchmark Set: How Well Do Quantum-Chemical Methods Describe Chalcogen-Bonding Interactions?

Nisha Mehta,[†] Thomas Fellowes,[†,‡] Jonathan M. White,[†,‡] and Lars Goerigk[*,†]

†*School of Chemistry, The University of Melbourne, Victoria 3010, Australia*

‡*Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Victoria, 3010, Australia*

E-mail: lars.goerigk@unimelb.edu.au

Phone: +61 3 834 46784

Abstract

We present the CHAL336 benchmark set—the most comprehensive database for the assessment of chalcogen-bonding (CB) interactions. After careful selection of suitable systems and identification of three high-level reference methods, the set comprises 336 dimers each consisting of up to 49 atoms and covers both σ- and π-hole interactions across four categories: chalcogen-chalcogen, chalcogen-π, chalcogen-halogen, and chalcogen-nitrogen interactions. In a subsequent study of DFT methods, we re-emphasize the need for using proper London-dispersion corrections when treating non-covalent interactions. We also point out that the deterioration of results and systematic overestimation of interaction energies for some dispersion-corrected DFT methods does not hint at problems with the chosen dispersion correction, but is a consequence of large density-driven errors. We conclude this work by performing the most detailed DFT

benchmark study for CB interactions to date. We assess 98 variations of dispersion-corrected and -uncorrected DFT methods, and carry out a detailed analysis of 72 of them. Double-hybrid functionals are the most reliable approaches for CB interactions, and they should be used whenever computationally feasible. The best three double hybrids are SOS0-PBE0-2-D3(BJ), revDSD-PBEP86-D3(BJ), and B2NCPLYP-D3(BJ). The best hybrids in this study are ωB97M-V, PW6B95-D3(0), and PW6B95-D3(BJ). We do not recommend using any lower-rung DFT methods nor the popular B3LYP and MP2 approaches, which have been used to describe CB interactions in the past. We hope to inspire a change in computational protocols surrounding CB interactions that leads away from the commonly used, popular methods to the more robust and accurate ones recommended herein. We would also like to encourage method developers to use our set for the investigation and reduction of density-driven errors in new density functional approximations.

1 Introduction

Noncovalent interactions (NCIs) play a crucial role in areas such as supramolecular chemistry, materials science, and chemical biology,[1–4] which is why a large number of experimental and theoretical studies have been devoted to understanding their nature.[1–3,5–40] Group 14-17 atoms covalently bound to electron-withdrawing groups can serve as Lewis-acid centers and be involved in relatively strong NCIs, which have been recognized in numerous chemical and biological systems.[41–44] The electron-withdrawing effect of the covalently bound group induces an electron-deficient area on the Lewis-acid centre opposite the covalent bond; this region is usually called the σ-hole. The σ-hole can form an adduct with a Lewis base (see Fig. 1 for a schematic representation), and such interactions are highly directional. Their strength depends on both the degree of electron deficiency in the σ-hole region as well as the electron density on the Lewis-base center. The resulting interactions are known as tetrel (for the atoms of group 14),[45–47] pnictogen (group 15),[42,43,48–50] chalcogen (group 16)[51–53]

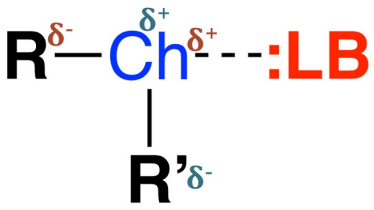

Figure 1: Schematic representation of CB interactions. Here, "Ch" represents the CB donor that is covalently bound to the functional groups R and R', and LB is a Lewis base.

and halogen (group 17)[41,44,54–56] bonding interactions.[57]

This study focuses on chalcogen-bonding (CB) interactions (Fig. 1). Their strength is comparable to that of hydrogen bonds and other σ-hole interactions.[58] CB interactions control the molecular orientations in crystal-structure packing as well as molecular recognition processes in biological systems. For example, S···O and Se···O interactions control the conformation of thiazole and selenazole nucleosides.[59] The structure and biological behavior of some sulfur-containing biological molecules are also controlled by CB interactions.[60,61] CB interactions also play an important role in the stabilization of protein structures and regulation of enzymatic functionalities,[62,63] and they can control the conformational stability of organic molecules.[64–71] Examples of using CB interactions for pharmacological benefits are selenazole nucleosides, where Se···O interactions play an important role in driving antiviral and antitumor activities.[59] Some of the present authors have recently given experimental evidence of CB at oxygen in o-nitro-O-aryl oxime, which is a rare example of oxygen involvement.[72]

Most systems induce CB through the aforementioned σ-hole but an analogous π-hole can also be the cause. A π-hole is a positive region of the electrostatic potential perpendicular to a planar π framework of the molecule.[73] Such positively charged regions can also interact with Lewis bases. π-hole mediated CB interactions have, for instance, been characterized in the H_2O-SO_3,[74–76] NH_3-SO_3,[77–79] $(SO_3)_n$-H_2CO,[80,81] $(SO_3)_n$-CO[82] and SO_3-$(CO)_n$[82] complexes, where n=1 and 2.

The aforementioned examples demonstrate that detailed studies of CB interactions are

of significance to the general chemistry community. Naturally, the interest of computational quantum chemists has also grown recently. However, a closer look at the literature reveals that high level, *ab-initio* calculations have rarely been conducted. Instead, most computational analyses are limited to second-order Møller-Plesset Perturbation Theory (MP2)[83] or low-level Density Functional Theory (DFT) approximations.[51,58,70,73,84] MP2/triple-ζ numbers have been compared with the Coupled-Cluster Singles Doubles and Perturbative Triples (CCSD(T))[85]/triple-ζ level of theory only in some occasions claiming that MP2 was reliable.[86–88] However, we would like to point out that reference energies based on a triple-ζ basis are unlikely be trustworthy due to the well-known slow energy convergence of electron-correlation methods with basis-set size; for a recent discussion on this topic in the context of benchmarking, see Refs 89 and 90. With some exceptions, many DFT based studies are limited to the B3LYP[91,92] hybrid density functional approximation (DFA) with relatively small basis sets, despite evidence that this level of theory is not adequate due to the London-dispersion, basis-set superposition and basis-set incompleteness errors, as well as the underlying exchange-correlation functional itself.[19,20,93,94]

Recently, CB interactions have gained high attention for applications in drug design, ion transport, catalysis, and materials design.[95–102] Therefore, it is very important to thoroughly investigate current quantum-chemical methods in the context of CB interactions before they are used in similar studies. DFT is the methodology of choice in such applications, however, as commented elsewhere there is an increasing gap between DFT developers and users due to the large and ever-growing DFT "zoo".[20,94] As a consequence, the field is full of misconceptions and many users base their computational studies on popular and highly cited methodologies, which however may not guarantee reliability and robustness.[94] In order to determine the most robust and reliable methods, comprehensive databases have been developed that cover different chemical problems simultaneously. Examples of such benchmark databases are the Gn test sets,[103–106] Database 2015B,[107] MGCDB84,[108] and the GMTKN24,[22] GMTKN30,[18] GMTKN55[20] databases for general main-group thermochem-

istry, kinetics and NCIs. Apart from test sets for chemical reactions, those comprehensive databases also include a variety of test sets for intra- and intermolecular NCIs, among which some evaluate directional interactions, such as hydrogen- and halogen-bonding.[11–13,15,16] While halogen-bonding interactions have been analyzed in detail in Refs 15,16 and 109–117, to the best of our knowledge no comprehensive benchmark study on CB interactions has been published with the exception of a 2013 study limited to a few smaller systems focussing on halogen, pnictogen, and CB interactions with neutral and anionic electron donors for some older pure/hybrid DFAs.[118] In addition, the findings for the aforementioned test sets and databases may not be sufficient to also guarantee that any of the recommended DFAs can be safely used for CB interactions. This can be seen from the fact that recent studies have revealed that many DFAs show a particularly large density error for small CB model systems;[119] this error is a consequence of the well-known self-interaction error (SIE) in DFT and can have negative impacts on a DFA's applicability; see Refs 119–126 for more details on the density error.

In the previous paragraph, we have established that there is a gap in the expertise of computational chemists when it comes to the accurate and reliable treatment of CB interactions. Moreover, DFT developers have identified the description of said interactions as a difficult problem. Herein, we close those existing gaps and provide users and developers alike with valuable new insights. We present a new benchmark set that comprises 336 high-level wave-function interaction energies in CB complexes dubbed "CHAL336". It is divided into different categories depending on the nature of the Lewis acid-base pairs as outlined in Section 2. The aim of this study is to gain insights into how well contemporary quantum-chemical methods, mainly DFAs, describe CB interactions. Such endeavor rises and falls with the accuracy of the reference data,[20,90,127] and in Section 4 we identify and assess different high-level strategies targeted at smaller and larger CB dimers. The resulting benchmark set and its reference values are then presented in Section 5. A detailed evaluation of DFT methods follows in Section 6. We carefully selected 98 DFA variations for this task and do

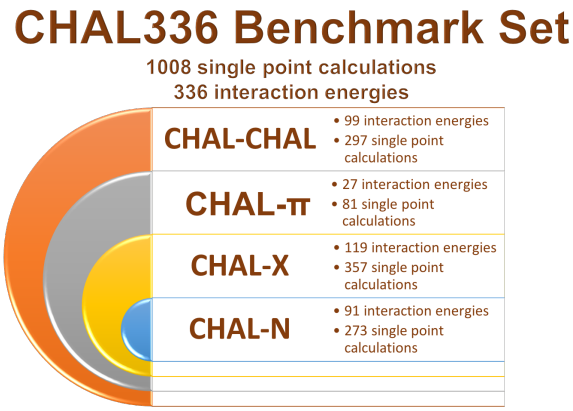

Figure 2: Overview of the CHAL336 benchmark set.

a thorough analysis of 72 of them, which also includes an assessment of London-dispersion corrections and the density error. We recommend those methods to DFT users that are the currently best for the computational treatment of CB interactions but also inform DFT developers on shortcomings of current methods; as such our findings also assist future method developments.

2 Overview of the CHAL336 benchmark set

Our aim is to design the first comprehensive benchmark set for intermolecular CB interactions. The strength of CB interactions increases from top to bottom in group 16 and based on selection criteria outlined in Section 4.1, we included S, Se and Te as CB donors. The resulting CHAL336 benchmark set covers 336 dimers of varying size. For each quantum-chemical method that a user would like to assess, 1008 single point calculations have to be conducted to obtain the 336 relative energies for the subsequent statistical analysis. As shown in Fig. 2, the 336 dimers are divided in four main categories. The first category is called the “CHAL-CHAL subset", and it includes problems where chalcogen-containing species act both as CB donors and Lewis bases. This subset comprises a total of 99 dimers. In the second category, 27 complexes are characterized by chalcogen-π bonding (CHAL-π subset), where interactions take place between the σ-hole of the chalcogen center and an

unsaturated moiety; this is not to be confused with π-hole interactions. The third category contains 119 complexes stabilized by chalcogen-halogen interactions (CHAL-X subset), while the fourth subset (CHAL-N) comprises 91 dimers with chalcogen-nitrogen interactions.

A detailed description of all 336 dimers follows in Section 5. All reference values in this study are zero-point vibrational energy exclusive, non-relativistic, and only involve valence electrons, as is standard in DFT benchmark studies. All structures can be obtained as a zip archive in the Supporting Information (SI).

3 General computational details

MOLPRO 2019.2[128] was used to obtain reference values for the Weizmann-n composite scheme mentioned in the next section. TURBOMOLE 7.3[129–131] was used for all geometry optimizations. ORCA4.2.0 and ORCA4.2.1[132,133] were employed for all remaining wave-function and DFT single-point calculations. DFT-D3[7,8] and DFT-D4[134,135] type dispersion corrections were obtained with the respective standalone programs by Grimme and co-workers.[136,137]

We applied Domain-Based Local Pair Natural Orbital Coupled-Cluster Theory with Singles Doubles and Perturbative Triple Excitations [DLPNO-CCSD(T)][138] in its latest linear-scaling implementation.[139] We employed the frozen-core approximation to all electron-correlation methods in ORCA, including double-hybrid DFT, to prevent basis-set superposition errors in the treatment of core-core electron correlation.[140] Core electrons of heavy elements (Te and beyond) were replaced with effective-core-potentials of the def2-ECP type in all cases.[141] Specific details on geometry optimizations, tested DFAs, atomic-orbital (AO) basis sets, technical setups for reference-energy and DFT calculations are discussed in the next sections.

Throughout this paper, we use the terms "interaction" or "binding energy" (ΔE) as

described below:

$$\Delta E = E_{12} - E_1 - E_2 \ , \tag{1}$$

where E_{12} is the total energy of the dimer and $E_{1/2}$ are the total energies of the individual monomers. While calculating ΔE, the monomers assume the same internal coordinates as in the bound dimer, as is common in similar studies.[9,11,20,142–144] If $\Delta E < 0$, a dimer is stable.

Any deviation during our method assessment is defined as the difference in interaction energies between the evaluated method and the reference value. Given that ΔE values are negative, a negative deviation indicates an overestimation of the binding energies and vice versa. Statistical values discussed herein are mean deviations (MDs), mean absolute deviations (MADs), root-mean-square deviations (RMSDs) and error ranges (ERs).

We advise to pay close attention to the resulting binding energies after a successful SCF procedure. Careful analysis revealed some unusual values for some DFAs that were against chemical intuition. Moreover, some of those values changed when different initial guesses were tested. In the majority of cases, we decided to use either an Extended Hückel initial guess or the results of a successfully converged DFT calculation that had led to a reasonable interaction energy.

4 First steps leading to the CHAL336 benchmark set

4.1 Geometry optimizations and selection criteria

All geometry pre-optimizations were performed with Turbomole's multigrid option "m4" and the final optimizations with grid "7" for the numerical integration of the exchange-correlation potentials. The latter is one of Turbomole's largest grid and minimizes the emergence of any artifacts that may impact the geometries of noncovalently bound structures. The self-consistent-field (SCF) calculations in each geometry cycle were carried out with a convergence criterion of 10^{-7} E_h. The geometry convergence criterion with respect to the change in total

energy between two cycles was set to 10^{-7} E_h. The Resolution-of-the-Identity approximation to the Coulomb integrals (RI-J)[131] was employed with appropriate auxiliary basis from the Turbomole basis-set library.[145] All model systems were pre-optimized at the PBEh-3c[146] hybrid DFT level. PBEh-3c is a low-cost DFT method based on a specifically designed small AO atomic-orbital double-ζ basis set with basis-set superposition error correction.[147] It provides a proper description of London- dispersion effects[7,8] and reasonably accurate geometries and properties. The resulting geometries were further optimized at the more accurate[8,148,149] dispersion-corrected PW6B95[150]-D3(BJ)[7,8]/def2-TZVPD[151] level of theory. We observed geometry convergence problems for some of systems containing Te or I at the PW6B95-D3(BJ)/def2-TZVPD level. Such systems were optimized with PBE0[152,153]-D3(BJ)/def2-TZVPD, instead. Some geometries of the initial model dimers or monomers were taken from Refs 51,58,73,86–88,154 and 155, but were re-optimized at the above mentioned levels of theory. Most other model systems had not been investigated before and are discussed here for the first time.

We started this study with nearly 1000 model dimers (3000 structures including monomers) because we allowed for different permutations of both chalcogen atoms and other functional groups. Out of those dimers, we short-listed 336 dimers based on the following criteria:

- We excluded dimers with an absolute binding energy < 0.5 kcal mol^{-1} based on the accurate reference level of theories defined in the following sections.

- We excluded dimers having oxygen as CB donor. While such cases have been shown to be possible,[156,157] the resulting interaction is quite weak, such that such interactions have often been ruled out or mentioned to be "rare".[98,158–161]

- Upon inspection post optimization, we excluded dimers that did not fulfill the directionality criteria of CB interactions, for instance wrong relative orientation of the monomers.

- In systems that displayed isomers, we only included those with the stronger CB inter-

actions.

- While some of the initial structures contained pnictogens other than N we decided to leave those out due to smaller CB interactions and little new insights that would not have warranted the increase in size of the benchmark set and computational resources.
- We excluded those complexes that displayed hydrogen-bonding along with CB interactions.

4.2 Establishing reference values

CCSD(T) at the complete basis set limit (CBS) is known as the "gold standard" method of chemical accuracy when the inclusion of higher-order excitations (true Triples and beyond) is not possible. However, the cost of a canonical CCSD(T) calculation scales as $\mathcal{O}(\mathrm{N}^7)$ (where N is the number of AOs or system size), which is why it is not always computationally feasible for chemically interesting systems.

While not affecting the formal scaling behavior, explicitly correlated approaches, such as the R12 and F12 variants,[162–166] allow using smaller AO basis sets with faster convergence to the CBS limit. Other alternatives to canonical CCSD(T) are various localized approaches.[138,139,167–173] In this work, we always strive to obtain the most reliable reference values for a given system size. Therefore, we employ a hierarchy of protocols based on explicitly or locally-correlated approaches as outlined next.

4.2.1 W1-F12 reference values

Reference binding energies of smaller systems could be calculated at the highly accurate W1-F12[174] level of theory. W1-F12 is a composite, thermochemical wave-function protocol that recovers the CCSD(T)/CBS limit through a series of finite-basis-set calculations that involve explicitly-correlated techniques. The accuracy of W1-F12 lies within the threshold of chemical accuracy, usually defined as 0.1 kcal mol^{-1} for NCI energies,[175] and it has been

previously shown how it delivers accurate reference energies for the assessment of lower-level methods.[20] Extrapolation exponents used in our W1-F12 calculations were based on the suggested values for systems containing second-row elements.[174] Moreover, we only considered the resulting energies for valence electrons, as is standard in DFT benchmarking. Additional calculations of core-electronic effects revealed that they were indeed negligible for our purposes. For instance, they only contributed -0.03 and -0.07 kcal mol^{-1} to the total binding energy of $C_2H_6O\cdots SC_2H_6$ and $C_2H_6S\cdots S(CH_3)CN$, respectively. All SCF calculations were done with a convergence threshold of 10^{-6} E_h and the resulting SCF energies were corrected with the complementary auxiliary basis (CABS) singles correction.[162,176–178] W1-F12 calculations were done with the RI approximations using OPTRI[179] auxiliary basis sets within the CABS approach.

Besides being computationally more demanding for larger systems, our W1-F12 treatments were also limited because the AO basis sets prescribed for this protocol are not defined for elements beyond Ar. Therefore, we often had to apply the methodology described in the following section.

4.2.2 DLPNO-CCSD(T)/CBS variants

An efficient and successful local-correlation approach is Neese and co-workers' DLPNO-CCSD(T) scheme.[138,139] DLPNO-CCSD(T)/CBS has been regarded as an ideal alternative to W1-F12 before and we follow in the footsteps of previous benchmark studies.[20,180] However, the remaining question is which basis-set types should be chosen, and in this section we conduct appropriate tests to determine this. These tests are inspired by recently established protocols for benchmarking enzymatically catalyzed reactions.[181] Herein, we follow those protocols with just some variations in basis sets.

In order to obtain the CBS limit, we used two-point extrapolation techniques. Hartree-Fock (HF)/CBS total energies ($E_{SCF}(CBS)$) were obtained as:[182]

$$E_{SCF}(CBS) = \frac{E_{SCF}(X)exp(-\alpha\sqrt{Y}) - E_{SCF}(Y)exp(-\alpha\sqrt{X})}{exp(-\alpha\sqrt{Y}) - exp(-\alpha\sqrt{X})} \ , \tag{2}$$

while the extrapolation of the correlation energies (E_{corr}) was performed as:[183]

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

where α and β are optimized, basis-set specific constants (see Table S1 in the SI) and X and Y are the basis set's cardinal numbers.

The electron-correlation steps in the DLPNO-CCSD(T) calculations were computed with the RI approximation using an appropriate auxiliary basis sets. All SCF calculations preceding the actual DLPNO-CCSD(T) steps were done with the default convergence criterion of 10^{-6} E_h.

Generally, extrapolation involving triple- and quadruple-ζ basis sets should be preferred over double- and triple-ζ sets.[181] Appropriate Dunning basis sets, namely aug-cc-pVTZ/aug-cc-pVQZ[184–186] (or aug-cc-pwCVTZ/aug-cc-pwCVQZ[187]) are the commonly recommended ones in this context. However, they have only been defined for elements up to Kr. Ahlrichs-type basis sets could be a possible alternative. Although they were originally developed for DFT methods, they have been successfully used for *ab-initio* calculations as well, with the advantage of being faster due to a smaller number of primitive Gaussian-type orbitals.[20,181] These basis sets are available for elements up to Rn.

Herein, we investigated the basis-set combinations aug-cc-pVTZ/aug-cc-pVQZ, aug-cc-pwCVTZ/aug-cc-pwCVQZ, def2-TZVPP/def2-QZVPP,[141] ma-def2-TZVPP/ma-def2-QZVPP,[188] and def2-TZVPPD/def2-QZVPPD,[151] where "aug", "ma" and "D" are identifiers for additional diffuse functions. All our tests were done with the "TightPNO" setup as suggested by Neese, Martin, and co-workers, which has been deemed particularly important for NCI energies and makes the differences to canonical CCSD(T) results negligible to minimal.[189] For a recent analysis of errors in DLPNO calculations, see Ref. 190.

During our first calculations, we noticed that the aug-cc-pVT/QZ basis sets behaved quite differently from the Ahlrichs ones for heavier elements, such as Se. For instance, the absolute difference between DLPNO-CCSD(T)/CBS(aug-cc-pVTZ/aug-cc-pVQZ) and DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP) for $C_2H_6O\cdots Se(CH_3)CN$ is 0.52 kcal mol^{-1}. For an element of the n^{th} principal shell, with n being the principal quantum number, the $(n-1)$d electrons are not frozen in ORCA's default "chemical-core". Contrary to Ahlrichs basis sets, the correlation consistent Dunning basis sets are minimal in the sub-valence region, which has been shown to negatively affect the description of halogen-bonding interactions.[110] This explains our observed large difference, as electron-correlation is underestimated for the $(n-1)$d electrons. Indeed, when using the "weighted" core-valence variant aug-cc-pwCVT/QZ, the results are more consistent and the absolute difference between DLPNO-CCSD(T)/CBS(aug-cc-pWCTZ/aug-cc-pwCQZ) and DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP) is only 0.08 kcal mol^{-1}. This automatically rules out the conventional correlation-consistent Dunning basis sets from further study.

At this stage, we arrive at three different strategies to reach the CBS limit with DLPNO-CCSD(T): aug-cc-pwCVTZ/aug-cc-pwCVQZ (strategy A), def2-TZVPPD/def2-QZVPPD (strategy B), and ma-def2-TZVPP/ma-def2-QZVPP (strategy C). We compare those against W1-F12 reference numbers for 15 systems and also provide statistical data in Table 1. When comparing strategy A against the W1-F12 numbers, we see very low statistical values, for instance MAD = 0.09 kcal mol^{-1} and ER = 0.19 kcal mol^{-1} (Table 1). Therefore, strategy A serves as our next best-possible benchmark to asses the other strategies whenever W1-F12 calculations are not feasible. Strategy C is the next best choice with an MAD of 0.14 kcal mol^{-1} compared to strategy B (MAD = 0.28 kcal mol^{-1}). The only difference between both basis-set types used in the two strategies are the diffuse functions. The def2-T/QZVPPD type basis sets underestimate the binding energies systematically with MD = 0.28 kcal mol^{-1}, while the ma-def2-TVZPP type seems to be better suited with MD = 0.12 kcal mol^{-1}. In passing, we note that using diffuse functions is advised. For instance, the MAD for DLPNO-

Table 1: DLPNO-CCSD(T)/CBS interaction energies and statistics with reference to W1-F12 energies (kcal mol^{-1}) calculated with the aug-cc-pwCVTZ/aug-cc-pwCVQZ (A), def2-TZVPPD/def2-QZVPPD (B) and ma-def2-TZVPP/ma-def2-QZVPP (C) AO basis sets.

No.	System	W1-F12	ΔA^a	ΔB^a	ΔC^a
1	$C_2H_6O\cdots SC_2H_6$	−3.39	0.14	0.23	0.16
2	$C_2H_6O\cdots S(CH_3)CN$	−4.98	0.16	0.29	0.16
3	$C_2H_6S\cdots S(CH_3)CN$	−4.72	0.16	0.29	0.14
4	$C_4H_4O\cdots SHCl$	−3.45	0.07	0.34	0.23
5	$C_4H_4O\cdots SHF$	−4.60	−0.01	0.34	0.20
6	$C_4H_4S\cdots SHCl$	−2.22	0.04	0.19	0.10
7	$C_4H_4S\cdots SHF$	−4.11	0.11	0.24	0.12
8	$F_2CS\cdots Cl^-$	−7.79	0.05	0.23	0.25
9	$F_2CS\cdots F^-$	−17.33	0.13	0.39	0.03
10	$H_2CS\cdots Cl^-$	−0.66	0.03	0.21	0.15
11	$H_2CS\cdots F^-$	−5.20	0.13	0.51	0.15
12	$OCS\cdots Cl^-$	−9.77	0.07	0.20	0.21
13	$OCS\cdots F^-$	−19.55	0.07	0.35	−0.08
14	$SCS\cdots Cl^-$	−10.58	0.04	0.07	0.06
15	$SCS\cdots F^-$	−21.25	0.18	0.32	−0.05
MD			0.09	0.28	0.12
MAD			0.09	0.28	0.14
RMSD			0.11	0.30	0.15
ER			0.19	0.44	0.33

^{a} ΔA/B/C= $\Delta E_{A/B/C}$ -ΔE_{W1-F12}.

CCSD(T)/CBS(def2-TZVPP/def2-QZVPP) (strategy D) is 0.56 kcal mol^{-1} (see Table S2).

W1-F12 and strategy-A calculations are only feasible for a limited number of elements, which means that Ahlrichs basis sets will have to be used for the majority of cases. The latter also have the advantage of being computationally friendlier. Therefore, we intend to discard strategy A from the development of our benchmark set in Section 5. In order to be more confident about the expected error when using Ahlrichs basis sets, we extend our analysis of strategies B and C and now consider 38 systems, with the same 15 W1-F12 references as before and an additional 23 reference values based on strategy A. The results in Table 2 confirm again that strategy C has a better MAD and MD than strategy B with MAD = 0.12 kcal mol^{-1} vs. 0.16 kcal mol^{-1} and MD = 0.02 kcal mol^{-1} vs. 0.11 kcal mol^{-1}, respectively. As such, the ma-def2-TZVPP/ma-def2-QZVPP basis-set pair was

chosen for the CBS extrapolation to obtain reference values in this study. Note that this recommendation aligns well with the previously made recommendations for enzymatically catalyzed reactions.[181]

Table 2: DLPNO-CCSD(T)/CBS interaction energies and statistics with reference to W1-F12 and DLPNO-CCSD(T)/CBS(aug-cc-pwCVTZ/aug-cc-pwCVQZ) energies (kcal mol^{-1}) calculated with the def2-TZVPPD/def2-QZVPPD (B) and ma-def2-TZVPP/ma-def2-QZVPP (C) AO basis sets.

No.	System	ref.	ΔB[a]	ΔC[a]
1	$C_2H_6O\cdots SC_2H_6$[b]	−3.39	0.23	0.16
2	$C_2H_6O\cdots S(CH_3)CN$[b]	−4.98	0.29	0.16
3	$C_2H_6O\cdots SeC_2H_6$[c]	−3.49	−0.04	−0.07
4	$C_2H_6O\cdots Se(CH_3)CN$[c]	−5.66	0.00	−0.08
5	$C_2H_6Se\cdots S(CH_3)CN$[c]	−4.52	0.11	−0.14
6	$C_2H_6Se\cdots SeC_2H_6$[c]	−3.78	−0.02	−0.11
7	$C_2H_6Se\cdots Se(CH_3)CN$[c]	−5.36	0.05	−0.13
8	$C_2H_6S\cdots S(CH_3)CN$[b]	−4.72	0.29	0.14
9	$C_2H_6S\cdots SeC_2H_6$[c]	−3.73	−0.03	−0.05
10	$C_2H_6S\cdots Se(CH_3)CN$[c]	−5.43	0.00	−0.12
11	$C_4H_4O\cdots SeHBr$[c]	−3.87	0.15	−0.03
12	$C_4H_4O\cdots SeHCl$[c]	−4.28	0.06	0.01
13	$C_4H_4O\cdots SeHF$[c]	−5.70	−0.04	−0.04
14	$C_4H_4O\cdots SHBr$[c]	−3.07	0.23	0.10
15	$C_4H_4O\cdots SHCl$[b]	−3.45	0.34	0.23
16	$C_4H_4O\cdots SHF$[b]	−4.60	0.34	0.20
17	$C_4H_4S\cdots SeHBr$[c]	−4.00	−0.12	−0.15
18	$C_4H_4S\cdots SeHCl$[c]	−4.28	−0.12	−0.12
19	$C_4H_4S\cdots SeHF$[c]	−5.46	−0.20	−0.22
20	$C_4H_4S\cdots SHBr$[c]	−2.01	0.15	0.01
21	$C_4H_4S\cdots SHCl$[b]	−2.22	0.19	0.10
22	$C_4H_4S\cdots SHF$[b]	−4.11	0.24	0.12
23	$F_2CS\cdots Cl^-$[b]	−7.79	0.23	0.25
24	$F_2CSe\cdots Cl^-$[c]	−11.04	−0.05	0.19
25	$F_2CSe\cdots F^-$[c]	−24.38	0.04	−0.08
26	$F_2CS\cdots F^-$[b]	−17.33	0.39	0.03
27	$H_2CS\cdots Cl^-$[b]	−0.66	0.21	0.15
28	$H_2CSe\cdots Cl^-$[c]	−3.28	0.12	0.25
29	$H_2CSe\cdots F^-$[c]	−10.57	0.08	−0.03
30	$H_2CS\cdots F^-$[b]	−5.20	0.51	0.15
31	$OCS\cdots Cl^-$[b]	−9.77	0.20	0.21
32	$OCSe\cdots Cl^-$[c]	−13.19	−0.09	0.10
33	$OCSe\cdots F^-$[c]	−27.74	0.05	−0.06
34	$OCS\cdots F^-$[b]	−19.55	0.35	−0.08
35	$SCS\cdots Cl^-$[b]	−10.58	0.07	0.06
36	$SCSe\cdots Cl^-$[c]	−13.05	−0.16	0.02
37	$SCSe\cdots F^-$[c]	−26.09	−0.08	−0.34
38	$SCS\cdots F^-$[b]	−21.25	0.32	−0.05
MD			0.11	0.02
MAD			0.16	0.12
RMSD			0.20	0.14
ER			0.71	0.59

[a] $\Delta B/C = \Delta E_{B/C} - \Delta E_{ref.}$.
[b] ref. values were calculated at the W1-F12 level.
[c] ref. values were calculated at the DLPNO-CCSD(T)/CBS(aug-cc-pwCVTZ/aug-cc-pwCVQZ) level (strategy A).

4.2.3 Composite scheme

In the previous section, we have determined a protocol for DLPNO-CCSD(T)/CBS reference energy calculations. For larger systems, where a conventional CBS extrapolation is very expensive, it has become common practice to determine reference energies with composite schemes that attempt to estimate the CBS limit for a high-level (HL) electron-correlation method.[9,181,191,192] The idea is to obtain a CBS total energy for a low-level (LL) method followed by adding the difference in correlation energies between HL and LL computed with a smaller basis set, which leads to the following expression for the estimated (est.) HL/CBS number:

$$E(est.HL/CBS) = E(LL, CBS) + E_{corr}(HL/small\,basis) - E_{corr}(LL/small\,basis) \quad (4)$$

Ref. 181 investigated various LL methods and basis-set combinations for enzymatically catalyzed reactions to obtain est. DLPNO-CCSD(T)/CBS values. Herein, we adopt their best-performing protocol, which we dub strategy E, with only a minor modification, namely that we employ diffuse functions for all calculations. In strategy E, we first obtained a spin-component scaled[193] DLPNO-SCS-MP2[194]/CBS(ma-def2-TZVPP/ma-def2-QZVPP) values that were subsequently corrected for the difference between DLPNO-CCSD(T) and DLPNO-SCS-MP2 with the ma-def2-TZVPP basis set. In all cases, TightPNO thresholds were applied.

The focus of our initial tests is to investigate the robustness of the above-mentioned scheme. The values that are ideally to be replicated are the DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP) ones. The resulting comparison for 48 dimers is shown in Table 3. Surprisingly, the differences between the properly extrapolated and the composite scheme seem to be relatively large with strategy E having an MAD of 0.17 kcal mol^{-1}. Upon closer inspection, we identified the fluoride-containing systems to be an anomaly. When excluding those, the MAD dropped to 0.08 kcal mol^{-1} with a nearly perfect MD of 0.03 kcal

mol^{-1}. These are acceptably low differences given the computational gain from applying strategy E. To our knowledge, it has not been reported before that the ma-def2 basis sets may not be suitable for fluoride-containing systems in such composite schemes.

We can conclude that in the present work, the best strategy to obtain reference values is the W1-F12 scheme. DLPNO-CCSD(T)/CBS numbers can be used as the next possible alternative when W1-F12 calculations are not feasible. Considering the range of elements that need to be covered and additional benefits from relying on fewer Gaussian-type orbitals, we opt for strategy C as our next best choice, which uses the ma-def2-TZVPP/ma-def2-QZVPP combination. Strategy E can be used to estimate DLPNO-CCSD(T)/CBS value if strategy C is too resource-demanding. In the following section, we apply the recommended strategies to all 336 systems and present the final version of the CHAL336 benchmark set.

5 The CHAL336 benchmark set

To the best of our knowledge, CHAL336 is the largest benchmark set for CB interactions and its reference values are the most accurate published for such interactions to date. Tables 4-7 show the names of all 336 dimers, their interaction energies, and the levels of theory at which those have been obtained. Across the whole set, the interaction energies range from -0.66 to -71.77 kcal mol^{-1} with an average value (ΔE_{av}) of -14.09 kcal mol^{-1}. As mentioned in Section 2, all 336 dimers are divided into four subsets each of which is discussed individually in the following.

5.1 The CHAL-CHAL subset

99 dimers represent typical chalcogen-chalcogen interactions; a total of 297 single-point calculations need to be carried out for this subset. The interaction energies in this subset range from -2.00 to -34.18 kcal mol^{-1}, with $\Delta E_{av} = -10.90$ kcal mol^{-1}. The reference interaction energies of 13 dimers were calculated at the W1-F12 level and the remaining 86 systems were

Table 3: Est. DLPNO-CCSD(T)/CBS (strategy E) interaction energies and statistics with reference to DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP) energies (kcal mol^{-1}).

No.	System	$\Delta E_{ref.}$	ΔE^a
1	S2···Br^-	−21.62	0.18
2	S2···Cl^-	−23.7	0.09
3	S2···F^-	−45.69	0.42
4	S2···I^-	−17.94	0.09
5	S3···Br^-	−12.99	0.16
6	S3···Cl^-	−13.96	−0.07
7	S3···F^-	−25.16	0.25
8	S3···I^-	−10.54	−0.04
9	S4···Br^-	−8.43	−0.01
10	S4···Cl^-	−9.21	−0.09
11	S4···F^-	−17.4	0.13
12	S4···I^-	−6.92	−0.03
13	S5···Br^-	−12.8	0.06
14	S5···Cl^-	−13.9	−0.02
15	S5···F^-	−25.77	0.18
16	S5···I^-	−10.44	−0.00
17	Se2···Br^-	−28.52	−0.02
18	Se2···Cl^-	−31.49	−0.16
19	Se2···F^-	−56.9	0.58
20	Se2···I^-	−24.03	−0.19
21	Se3···Br^-	−14.8	0.06
22	Se3···Cl^-	−16.13	−0.05
23	Se3···F^-	−30.35	0.40
24	Se3···I^-	−12.05	−0.03
25	Se4···Br^-	−10.06	0.13
26	Se4···Cl^-	−11.05	0.04
27	Se4···F^-	−22.32	0.37
28	Se4···I^-	−7.95	−0.02
29	Se5···Br^-	−15.49	−0.02
30	Se5···Cl^-	−17.16	−0.07
31	Se5···F^-	−33.19	0.31
32	Se5···I^-	−12.62	−0.05
33	Te2···Br^-	−40.1	0.20
34	Te2···Cl^-	−43.85	0.13
35	Te2···F^-	−71.77	0.70
36	Te2···I^-	−34.65	−0.03
37	Te3···Br^-	−21.81	0.17
38	Te3···Cl^-	−24.11	0.09
39	Te3···F^-	−45.76	0.70
40	Te3···I^-	−17.95	−0.01
41	Te4···Br^-	−15.74	0.16
42	Te4···Cl^-	−17.61	0.06
43	Te4···F^-	−36.84	0.53
44	Te4···I^-	−12.59	0.03
45	Te5···Br^-	−23.14	0.12
46	Te5···Cl^-	−25.76	0.12
47	Te5···F^-	−48.63	0.55
48	Te5···I^-	−19.04	0.02
MD			−0.13 (−0.03[b])
MAD			0.17 (0.08[b])
RMSD			0.25 (0.10[b])
ER			0.89 (0.39[b])

[a] $\Delta E = \Delta E_E - \Delta E_{ref.}$. This is not to be confused with ΔE in eq. 1, but this is the error for strategy E. [b] statistics on 38 systems after excluding fluoride-containing dimers.

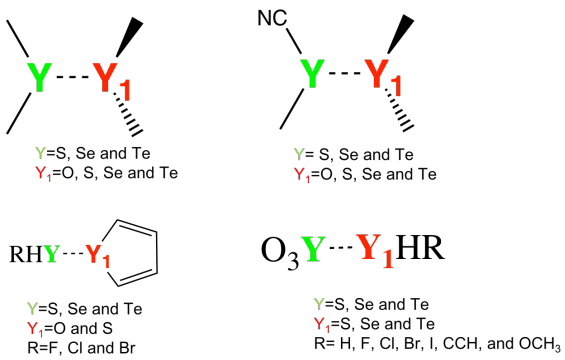

Figure 3: Lewis structures of the model systems included in CHAL-CHAL subset of the CHAL336 benchmark set.

obtained with strategy C [DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP)]; see Table 4 for details.

Example Lewis structures of the systems included in the CHAL-CHAL subset are shown in Fig. 3. A total of 8 dimers consist of dimethyl-substituted monomers of the form $C_2H_6Y_1\cdots YC_2H_6$ (where Y = S, Se and Te; Y_1 = O, S, Se and Te), as inspired by Refs 86–88. For those, ΔE ranges from -3.39 kcal mol^{-1} ($C_2H_6O\cdots SC_2H_6$) to -5.49 kcal mol^{-1} ($C_2H_6Te\cdots TeC_2H_6$). This is an expected result given that σ-holes are stronger for Te than for the lighter chalcogen atoms. The average interaction energy for these dimers is -4.36 kcal mol^{-1}. In 11 dimers, a cyano group replaces one of the methyl groups to increase the strength of the Lewis-acid center[86–88] (see Fig. 3). This results in stronger CB interactions that range from -4.66 kcal mol^{-1} ($C_2H_6Se\cdots S(CH_3)CN$) to -8.66 kcal mol^{-1} ($C_2H_6Te\cdots Te(CH_3)CN$) with $\Delta E_{av}-6.36$ kcal mol^{-1}.

18 complexes in this category involve either furan or thiophene as Lewis base and have the form furan$\cdots Y_1HR$ and thiophene$\cdots Y_1HR$, where Y_1 = S, Se and Te, and R = F, Cl and Br, as inspired by dimers shown in Ref. 58 The interaction energies for these dimers vary from -2.00 ($C_4H_4S\cdots SHBr$) to -7.97 kcal mol^{-1} ($C_4H_4S\cdots TeHF$) with $\Delta E_{av} = -4.86$ kcal mol^{-1}. For a given substituent R and Lewis base, dimers involving TeHR have the highest absolute interaction energies ($\Delta E_{av} = -6.67$ kcal mol^{-1}), followed by SeHR ($\Delta E_{av} = -4.69$ kcal mol^{-1}) and SHR ($\Delta E_{av} = -3.22$ kcal mol^{-1}). The CB interaction increases with more

electronegative substituents R for a given Lewis acid. For instance, the average interaction energies of furan···Y_1HF, furan···Y_1HCl, furan···Y_1HBr are −6.01, −4.63 and −4.15 kcal mol^{-1}, respectively.

The above discussed dimers are σ-hole complexes, but we also intend to cover π-hole bonded dimers in this category. For this purpose we included a total of 62 of the form YO_3···Y_1HR (where Y = S, Se and Te; Y_1 = S, Se and Te; R = H, F, Cl, Br, I, CCH, OCH_3 and NC), based on ideas published in Ref. 73. In such complexes, YO_3 generally acts as the CB donor through the π-hole and Y_1HR as the Lewis base. Surprisingly and contrary to our initial expectations, a number of those complexes turned out to be stabilized by a combination of both π-hole···lone-pair and σ-hole···lone-pair interactions. In such complexes, Y_1HR acts as the CB donor and the oxygens of YO_3 act as Lewis bases. This observation is based on the directionality of the CB interactions; Fig. S1 provides a visual comparison between a system stabilized exclusively by π-hole interactions (SeO_3···SeHBr) and one formed by both types of interactions (SeO_3···SeHF). The following systems display this form of mixed interaction: SeO_3···SeHF, SeO_3···SHF, SeO_3···TeHBr, SeO_3···TeHCl, SeO_3···TeHF, SeO_3···TeHNC, SO_3···SeHF, SO_3···SHF, TeO_3···SeHCl, TeO_3···SeHF, TeO_3···SeHNC, TeO_3···SHF, TeO_3···TeHBr, TeO_3···TeHF, TeO_3···TeHNC, and TeO_3···$TeOCH_3$.

The interaction energies of the YO_3···Y_1HR complexes range from −5.48 (SO_3···SHNC) to −34.18 kcal mol^{-1} (TeO_3···TeHF) with ΔE_{av} = −14.31 kcal mol^{-1}. Overall, those interactions are stronger than for the previously discussed systems. For a given substituent R, the absolute value of ΔE increases when moving from the lighter to the heavier CB donors, i.e. S···S-type dimers have the lowest absolute interaction energies (ΔE_{av} = −9.69 kcal mol^{-1}), followed by Se···Se-type (ΔE_{av} = −11.69 kcal mol^{-1}), and Te···Te-type dimers (ΔE_{av} = −25.15 kcal mol^{-1}). In YO_3···Y_1HR complexes, the nature of the substituent R has a significant influence on the interaction energies. For instance, the value of ΔE_{av} decreases from −13.97 to −17.58 kcal mol^{-1} for YO_3···Y_1H_2 and YO_3···Y_1OCH_3 dimers and increase to −9.85 kcal mol^{-1} for the YO_3···Y_1HNC complexes.

Table 4: Reference interaction energies (ΔE in kcal mol^{-1}) of the CHAL-CHAL subset.

Label	System	ΔE	Ref. strategy	Label	System	ΔE	Ref. strategy
CHAL-CHAL-1	$O_3Te\cdots TeHF$	−34.18	C[a]	**CHAL-CHAL-51**	$O_3S\cdots SeHBr$	−9.77	C
CHAL-CHAL-2	$O_3Te\cdots TeHOCH_3$	−30.72	C	**CHAL-CHAL-52**	$O_3Te\cdots SHNC$	−9.68	C
CHAL-CHAL-3	$O_3Te\cdots TeHCl$	−26.28	C	**CHAL-CHAL-53**	$O_3S\cdots SHCl$	−9.54	W1-F12
CHAL-CHAL-4	$O_3Se\cdots TeHF$	−25.69	C	**CHAL-CHAL-54**	$O_3S\cdots SeHCl$	−9.51	C
CHAL-CHAL-5	$O_3Te\cdots TeHBr$	−24.35	C	**CHAL-CHAL-55**	$O_3S\cdots SHBr$	−9.25	C
CHAL-CHAL-6	$O_3Te\cdots TeHNC$	−21.71	C	**CHAL-CHAL-56**	$O_3S\cdots SeHCCH$	−8.75	C
CHAL-CHAL-7	$O_3Te\cdots SeHF$	−20.14	C	**CHAL-CHAL-57**	$C_2H_6Te\cdots Te(CH_3)CN$	−8.66	C
CHAL-CHAL-8	$O_3Te\cdots TeH_2$	−19.76	C	**CHAL-CHAL-58**	$C_2H_6Se\cdots Te(CH_3)CN$	−8.46	C
CHAL-CHAL-9	$O_3Te\cdots SHOCH_3$	−19.27	C	**CHAL-CHAL-59**	$C_2H_6S\cdots Te(CH_3)CN$	−8.31	C
CHAL-CHAL-10	$O_3Te\cdots TeHCCH$	−19.02	C	**CHAL-CHAL-60**	$O_3S\cdots SHCCH$	−8.26	W1-F12
CHAL-CHAL-11	$O_3S\cdots TeHF$	−18.91	C	**CHAL-CHAL-61**	$O_3S\cdots TeHNC$	−8.16	C
CHAL-CHAL-12	$O_3Te\cdots SeHOCH_3$	−18.68	C	**CHAL-CHAL-62**	$C_2H_6O\cdots Te(CH_3)CN$	−7.97	C
CHAL-CHAL-13	$O_3Se\cdots TeHCl$	−18.37	C	**CHAL-CHAL-63**	$C_4H_4S\cdots TeHF$	−7.97	C
CHAL-CHAL-14	$O_3Te\cdots SeH_2$	−17.69	C	**CHAL-CHAL-64**	$C_4H_4O\cdots TeHF$	−7.68	C
CHAL-CHAL-15	$O_3Se\cdots TeHBr$	−16.89	C	**CHAL-CHAL-65**	$O_3Se\cdots SeHNC$	−7.00	C
CHAL-CHAL-16	$O_3Te\cdots SHF$	−16.85	C	**CHAL-CHAL-66**	$C_4H_4S\cdots TeHCl$	−6.57	C
CHAL-CHAL-17	$O_3Te\cdots SH_2$	−16.40	C	**CHAL-CHAL-67**	$O_3Se\cdots SHNC$	−6.23	C
CHAL-CHAL-18	$O_3Te\cdots SeHBr$	−16.11	C	**CHAL-CHAL-68**	$C_4H_4O\cdots TeHCl$	−6.16	C
CHAL-CHAL-19	$O_3Te\cdots SeHCCH$	−15.81	C	**CHAL-CHAL-69**	$C_4H_4S\cdots TeHBr$	−6.09	C
CHAL-CHAL-20	$O_3S\cdots TeHOCH_3$	−15.63	C	**CHAL-CHAL-70**	$C_2H_6O\cdots Se(CH_3)CN$	−5.74	C
CHAL-CHAL-21	$O_3Te\cdots SeHCl$	−15.62	C	**CHAL-CHAL-71**	$C_4H_4O\cdots SeHF$	−5.74	C
CHAL-CHAL-22	$O_3Te\cdots SHBr$	−15.25	C	**CHAL-CHAL-72**	$C_4H_4S\cdots SeHF$	−5.68	C
CHAL-CHAL-23	$O_3Se\cdots TeH_2$	−15.23	C	**CHAL-CHAL-73**	$O_3S\cdots SeHNC$	−5.62	C
CHAL-CHAL-24	$O_3Se\cdots SHOCH_3$	−14.83	C	**CHAL-CHAL-74**	$C_4H_4O\cdots TeHBr$	−5.57	C
CHAL-CHAL-25	$O_3Se\cdots TeHCCH$	−14.72	C	**CHAL-CHAL-75**	$C_2H_6S\cdots Se(CH_3)CN$	−5.55	C
CHAL-CHAL-26	$O_3Te\cdots SHCl$	−14.60	C	**CHAL-CHAL-76**	$C_2H_6Se\cdots Se(CH_3)CN$	−5.49	C
CHAL-CHAL-27	$O_3S\cdots SHOCH_3$	−14.56	W1-F12	**CHAL-CHAL-77**	$C_2H_6Te\cdots TeC_2H_6$	−5.49	C
CHAL-CHAL-28	$O_3Te\cdots SHCCH$	−14.41	C	**CHAL-CHAL-78**	$O_3S\cdots SHNC$	−5.48	W1-F12
CHAL-CHAL-29	$O_3Se\cdots SeHOCH_3$	−14.21	C	**CHAL-CHAL-79**	$C_2H_6Te\cdots Se(CH_3)CN$	−5.47	C
CHAL-CHAL-30	$O_3Se\cdots SeHF$	−13.80	C	**CHAL-CHAL-80**	$C_2H_6Se\cdots TeC_2H_6$	−5.12	C
CHAL-CHAL-31	$O_3Se\cdots TeHNC$	−13.69	C	**CHAL-CHAL-81**	$C_2H_6S\cdots TeC_2H_6$	−5.05	C
CHAL-CHAL-32	$O_3S\cdots TeHCl$	−12.81	C	**CHAL-CHAL-82**	$C_2H_6O\cdots S(CH_3)CN$	−4.98	W1-F12
CHAL-CHAL-33	$O_3S\cdots SeHOCH_3$	−12.72	C	**CHAL-CHAL-83**	$C_2H_6S\cdots S(CH_3)CN$	−4.72	W1-F12
CHAL-CHAL-34	$O_3Se\cdots SeH_2$	−12.63	C	**CHAL-CHAL-84**	$C_2H_6Se\cdots S(CH_3)CN$	−4.66	C
CHAL-CHAL-35	$O_3S\cdots TeHBr$	−12.61	C	**CHAL-CHAL-85**	$C_4H_4O\cdots SHF$	−4.60	W1-F12
CHAL-CHAL-36	$O_3S\cdots TeH_2$	−12.40	C	**CHAL-CHAL-86**	$C_2H_6O\cdots TeC_2H_6$	−4.58	C
CHAL-CHAL-37	$O_3S\cdots TeHCCH$	−11.83	C	**CHAL-CHAL-87**	$C_4H_4S\cdots SeHCl$	−4.40	C
CHAL-CHAL-38	$O_3Se\cdots SeHBr$	−11.69	C	**CHAL-CHAL-88**	$C_4H_4O\cdots SeHCl$	−4.27	C
CHAL-CHAL-39	$O_3Se\cdots SHF$	−11.56	C	**CHAL-CHAL-89**	$C_4H_4S\cdots SeHBr$	−4.15	C
CHAL-CHAL-40	$O_3Se\cdots SH_2$	−11.46	C	**CHAL-CHAL-90**	$C_4H_4S\cdots SHF$	−4.11	W1-F12
CHAL-CHAL-41	$O_3Se\cdots SeHCl$	−11.29	C	**CHAL-CHAL-91**	$C_4H_4O\cdots SeHBr$	−3.90	C
CHAL-CHAL-42	$O_3Se\cdots SeHCCH$	−11.19	C	**CHAL-CHAL-92**	$C_2H_6Se\cdots SeC_2H_6$	−3.89	C
CHAL-CHAL-43	$O_3Te\cdots SeHNC$	−11.05	C	**CHAL-CHAL-93**	$C_2H_6S\cdots SeC_2H_6$	−3.78	C
CHAL-CHAL-44	$O_3S\cdots SeHF$	−10.94	C	**CHAL-CHAL-94**	$C_2H_6O\cdots SeC_2H_6$	−3.56	C
CHAL-CHAL-45	$O_3S\cdots SHF$	−10.93	W1-F12	**CHAL-CHAL-95**	$C_4H_4O\cdots SHCl$	−3.45	W1-F12
CHAL-CHAL-46	$O_3Se\cdots SHBr$	−10.81	C	**CHAL-CHAL-96**	$C_2H_6O\cdots SC_2H_6$	−3.39	W1-F12
CHAL-CHAL-47	$O_3Se\cdots SHCl$	−10.47	C	**CHAL-CHAL-97**	$C_4H_4O\cdots SHBr$	−2.97	C
CHAL-CHAL-48	$O_3S\cdots SeH_2$	−10.35	C	**CHAL-CHAL-98**	$C_4H_4S\cdots SHCl$	−2.22	W1-F12
CHAL-CHAL-49	$O_3Se\cdots SHCCH$	−9.98	C	**CHAL-CHAL-99**	$C_4H_4S\cdots SHBr$	−2.00	C
CHAL-CHAL-50	$O_3S\cdots SH_2$	−9.78	W1-F12				

[a] DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP).

5.2 The CHAL-π subset

We included 27 systems representative of interactions between chalcogen atoms and π systems. Dimers of the form $R_2Y\cdots A$—with R = F, Cl and Br; Y = S, Se and Te; and A=acetylene (ac), ethylene (et) and 2-butyne (2-but)—were chosen for this purpose, as inspired by Ref. 154. The reference interaction energies for six systems were obtained with W1-F12 and the remaining according to strategy C; see Table 5 for details. ΔE values for this subset range from -2.19 kcal mol^{-1} ($Br_2S\cdots$ac) to -13.83 kcal mol^{-1} ($F_2Te\cdots$2but) and have an average value of -5.73 kcal mol^{-1}.

For a given π system A, the stability of the complexes increases with the electropositivity of the CB donor Y and the electronegativity of the covalently bound atom R. The ΔE_{av} values decrease from -3.37 ($R_2S\cdots$A) to -5.22 and -8.62 kcal mol^{-1} for the $R_2Se\cdots$A and $R_2Te\cdots$A dimers, respectively. The F, Cl and Br based complexes have ΔE_{av} values of -7.43, -5.00 and -4.78 kcal mol^{-1}, respectively. 2-but based complexes are the strongest (ΔE_{av}(2but)= -7.61 kcal mol^{-1}), followed by et-based (ΔE_{av}(et)= -5.43 kcal mol^{-1}) and ac-based ones (ΔE_{av}(ac)= -4.17 kcal mol^{-1}).

Table 5: Reference interaction energies (ΔE in kcal mol^{-1}) of the CHAL-π subset.

Label	System	ΔE	Ref. strategy	Label	System	ΔE	Ref. strategy
CHAL-π-1	$F_2Te\cdots$2but	-13.83	C[a]	**CHAL-π-15**	$Br_2Te\cdots$ac	-4.84	C
CHAL-π-2	$F_2Te\cdots$et	-12.78	C	**CHAL-π-16**	$Br_2S\cdots$2but	-4.55	C
CHAL-π-3	$Cl_2Te\cdots$2but	-9.71	C	**CHAL-π-17**	$Cl_2S\cdots$2but	-4.49	W1-F12
CHAL-π-4	$Br_2Te\cdots$2but	-9.27	C	**CHAL-π-18**	$Cl_2Se\cdots$et	-4.09	C
CHAL-π-5	$F_2Se\cdots$2but	-8.89	C	**CHAL-π-19**	$Br_2Se\cdots$et	-3.85	C
CHAL-π-6	$F_2Te\cdots$ac	-8.58	C	**CHAL-π-20**	$Cl_2Se\cdots$ac	-3.27	C
CHAL-π-7	$Cl_2Te\cdots$et	-6.91	C	**CHAL-π-21**	$F_2S\cdots$et	-3.25	W1-F12
CHAL-π-8	$Br_2Te\cdots$et	-6.35	C	**CHAL-π-22**	$Br_2Se\cdots$ac	-3.07	C
CHAL-π-9	$F_2Se\cdots$et	-6.31	C	**CHAL-π-23**	$F_2S\cdots$ac	-3.00	W1-F12
CHAL-π-10	$Cl_2Se\cdots$2but	-6.30	C	**CHAL-π-24**	$Br_2S\cdots$et	-2.66	C
CHAL-π-11	$Br_2Se\cdots$2but	-6.20	C	**CHAL-π-25**	$Cl_2S\cdots$et	-2.63	W1-F12
CHAL-π-12	$Cl_2Te\cdots$ac	-5.30	C	**CHAL-π-26**	$Cl_2S\cdots$ac	-2.29	W1-F12
CHAL-π-13	$F_2S\cdots$2but	-5.26	W1-F12	**CHAL-π-27**	$Br_2S\cdots$ac	-2.19	C
CHAL-π-14	$F_2Se\cdots$ac	-4.97	C				

[a] DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP).

5.3 The CHAL-X subset

To assess intermolecular chalcogen-halogen interactions, we compiled the CHAL-X subset, which consists of 119 dimers. Their interaction energies range from −0.66 to −71.77 kcal mol^{-1} with an average value of −22.35 kcal mol^{-1}. Reference energies for eight dimers are based on W1-F12, 84 on strategy C, and 27 on strategy E (est. DLPNO-CCSD(T)/CBS) due to their large size; see Table 6 for details. Contrary to the previous two subsets, CHAL-X consists of charged species.

CHAL-X contains 44 representative model systems of the form $H_2CY\cdots X^-$, $F_2CY\cdots X^-$, $O{=}C{=}Y\cdots X^-$, and $S{=}C{=}Y\cdots X^-$, where Y = S, Se and Te; X=F, Cl and Br. Their interaction energies range from −0.66 ($H_2CS\cdots Cl^-$) to −47.57 kcal mol^{-1} $O{=}C{=}Te\cdots F^-$ with an average value of −14.89 kcal mol^{-1}. $H_2CY\cdots X^-$-type complexes have the lowest absolute ΔE_{av} (−7.32 kcal mol^{-1}), followed by $F_2CY\cdots X^-$ (−16.22 kcal mol^{-1}), $S{=}C{=}Y\cdots X^-$ (−17.04 kcal mol^{-1}) and $O{=}C{=}Y\cdots X^-$ (−18.10 kcal mol^{-1}). There is a significant increase in stability when moving from S to Se and Te, with ΔE_{av} = −9.96, −11.93, and −22.33 kcal mol^{-1}, respectively.

In order to assess larger systems, some of which are based on monomers that have been used in experimental studies of similar CB interactions,[155] we included a total of 75 dimers formed between fluoride, chloride, bromide, and iodide anions and the following CB donors, for which we introduce the labels highlighted in bold and whose Se-based variants are shown in Fig. 4: 2-phenylbenzo[*d*]isothiazol-3(2*H*)-one (**S1**), isothiazol-3(2*H*)-one (**S2**), thiophen-3(2*H*)-one (**S3**), thiophene (**S4**), 1,2-selenazole (**S5**), 7*H*-benzo[4,5]isothiazolo[3,2-*b*]quinazolin-7-one (**S6**), 2-benzylbenzo[*d*]isothiazol-3(2*H*)-one (**S7**), 2-phenylbenzo[*d*][1,2]selenazol-3(2*H*)-one (**Se1**), 1,2-selenazol-3(2*H*)-one (**Se2**), selenophen-3(2*H*)-one (**Se3**), selenophene (**Se4**), 1,2-selenazole (**Se5**), 7*H*-benzo[4,5][1,2]selenazolo[3,2-*b*]quinazolin-7-one (**Se6**), 2-benzylbenzo[*d*][1,2]selenazol-3(2*H*)-one (**Se7**), 2-phenylbenzo[*d*][1,2]tellurazol-3(2*H*)-one (**Te1**), 1,2-tellurazol-3(2*H*)-one (**Te2**), tellurophen-3(2H)-one (**Te3**), tellurophene (**Te4**), 1,2-tellurazol-3(2*H*)-one (**Te5**), 7H-benzo[4,5][1,2]tellurazolo[3,2-*b*]quinazolin-7-one (**Te6**) and 2-benzylbenzo[*d*]

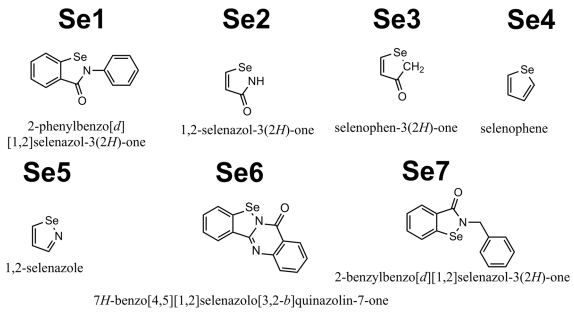

Figure 4: Lewis structures of **Se1**-**Se7** from the CHAL-X and CHAL-N subsets. S and Te equivalents are also used.

[1,2]tellurazol-3(2H)-one (**Te7**). **Se1**, **Se6**, **Se7** were taken from Ref. 155, while the other compounds were inspired by that study. The largest dimers in this subset are the ones formed with **S7**, **Se7**, and **Te7**, with 29 atoms in each case. The interaction energies of these 75 complexes range from −6.92 (S4···I$^-$) to −71.77 kcal mol^{-1} (Te2···F$^-$) with ΔE_{av} = −26.73 kcal mol^{-1}.

When assessing the CB-donor abilities of monomers **S1**-**S7**, **Se1**-**Se7**, and **Te1**-**Te7**, we see the expected trend that the Te-based complexes are more stable (ΔE_{av} = −35.59 kcal mol^{-1}), followed by Se-based (ΔE_{av} = −25.52 kcal mol^{-1}), and S-based ones (ΔE_{av} = −19.07 kcal mol^{-1}). The stability of Te-based complexes increase in the following order: **Te4** (−20.69 kcal mol^{-1}), **Te3** (−27.41 kcal mol^{-1}), **Te5** (−29.14 kcal mol^{-1}), **Te7** (−38.01 kcal mol^{-1}), **Te2** (−47.59 kcal mol^{-1}), **Te6** (−46.04 kcal mol^{-1}), and **Te1** (−46.12 kcal mol^{-1}). Interestingly, the order changes slightly for Se-based dimers: **Se4** (−12.85 kcal mol^{-1}), **Se3** (−18.33 kcal mol^{-1}), **Se5** (−19.61 kcal mol^{-1}), **Se2** (−35.24 kcal mol^{-1}), **Se7** (−29.39 kcal mol^{-1}), **Se1** (−33.89 kcal mol^{-1}), and **Se6** (−34.67 kcal mol^{-1}). When assessing the halide ions fluoride-based complexes have the the strongest CB interactions (ΔE_{av} = −38.31 kcal mol^{-1}), followed by Cl$^-$ (ΔE_{av} = −27.90 kcal mol^{-1}), Br$^-$ (ΔE_{av} = −25.22 kcal mol^{-1}) and I$^-$ (ΔE_{av} = −20.45 kcal mol^{-1}).

Table 6: Reference interaction energies (ΔE in kcal mol^{-1}) of the CHAL-X subset.

Label	System	ΔE	Ref. strategy	Label	System	ΔE	Ref. strategy
CHAL-X-1	Te2···F$^-$	−71.77	C[a]	**CHAL-X-61**	S1···I$^-$	−19.16	E
CHAL-X-2	Se2···F$^-$	−56.90	C	**CHAL-X-62**	Te5···I$^-$	−19.04	C
CHAL-X-3	Te6···Cl$^-$	−51.27	E[b]	**CHAL-X-63**	S6···I$^-$	−18.74	E
CHAL-X-4	Te1···Cl$^-$	−51.06	E	**CHAL-X-64**	SCTe···Br$^-$	−18.67	C
CHAL-X-5	Te5···F$^-$	−48.63	C	**CHAL-X-65**	F_2CTe···Br$^-$	−18.30	C
CHAL-X-6	OCTe···F$^-$	−47.57	C	**CHAL-X-66**	Te3···I$^-$	−17.95	C
CHAL-X-7	Te7···Cl$^-$	−46.84	E	**CHAL-X-67**	S2···I$^-$	−17.94	C
CHAL-X-8	Te1···Br$^-$	−46.82	E	**CHAL-X-68**	Te4···Cl$^-$	−17.61	C
CHAL-X-9	Te6···Br$^-$	−46.81	E	**CHAL-X-69**	S4···F$^-$	−17.40	C
CHAL-X-10	Te3···F$^-$	−45.76	C	**CHAL-X-70**	F_2CS···F$^-$	−17.33	W1-F12
CHAL-X-11	S2···F$^-$	−45.69	C	**CHAL-X-71**	Se5···Cl$^-$	−17.16	C
CHAL-X-12	Te2···Cl$^-$	−43.85	C	**CHAL-X-72**	S7···I$^-$	−16.71	E
CHAL-X-13	Te7···Br$^-$	−42.68	E	**CHAL-X-73**	Se3···Cl$^-$	−16.13	C
CHAL-X-14	F_2CTe···F$^-$	−42.61	C	**CHAL-X-74**	Te4···Br$^-$	−15.74	C
CHAL-X-15	SCTe···F$^-$	−41.06	C	**CHAL-X-75**	Se5···Br$^-$	−15.49	C
CHAL-X-16	Te1···I$^-$	−40.48	E	**CHAL-X-76**	SCTe···I$^-$	−15.15	C
CHAL-X-17	Te2···Br$^-$	−40.10	C	**CHAL-X-77**	Se3···Br$^-$	−14.80	C
CHAL-X-18	Te6···I$^-$	−40.04	E	**CHAL-X-78**	F_2CTe···I$^-$	−14.67	C
CHAL-X-19	Se6···Cl$^-$	−39.40	E	**CHAL-X-79**	S3···Cl$^-$	−13.96	C
CHAL-X-20	Se1···Cl$^-$	−38.20	E	**CHAL-X-80**	S5···Cl$^-$	−13.90	C
CHAL-X-21	Te4···F$^-$	−36.84	C	**CHAL-X-81**	OCSe···Cl$^-$	−13.09	C
CHAL-X-22	Se6···Br$^-$	−35.08	E	**CHAL-X-82**	SCSe···Cl$^-$	−13.03	C
CHAL-X-23	Te2···I$^-$	−34.65	C	**CHAL-X-83**	S3···Br$^-$	−12.99	C
CHAL-X-24	Se1···Br$^-$	−34.33	E	**CHAL-X-84**	S5···Br$^-$	−12.80	C
CHAL-X-25	Se7···Cl$^-$	−33.66	E	**CHAL-X-85**	Se5···I$^-$	−12.62	C
CHAL-X-26	Se5···F$^-$	−33.19	C	**CHAL-X-86**	Te4···I$^-$	−12.59	C
CHAL-X-27	Se2···Cl$^-$	−31.49	C	**CHAL-X-87**	Se3···I$^-$	−12.05	C
CHAL-X-28	Se3···F$^-$	−30.35	C	**CHAL-X-88**	OCSe···Br$^-$	−11.48	C
CHAL-X-29	Se7···Br$^-$	−30.09	E	**CHAL-X-89**	SCSe···Br$^-$	−11.28	C
CHAL-X-30	Se6···I$^-$	−29.54	E	**CHAL-X-90**	Se4···Cl$^-$	−11.05	C
CHAL-X-31	Se1···I$^-$	−29.14	E	**CHAL-X-91**	F_2CSe···Cl$^-$	−10.85	C
CHAL-X-32	Se2···Br$^-$	−28.52	C	**CHAL-X-92**	H_2CSe···F$^-$	−10.60	C
CHAL-X-33	OCSe···F$^-$	−27.80	C	**CHAL-X-93**	SCS···Cl$^-$	−10.58	W1-F12
CHAL-X-34	S6···Cl$^-$	−27.22	E	**CHAL-X-94**	S3···I$^-$	−10.54	C
CHAL-X-35	S1···Cl$^-$	−26.48	E	**CHAL-X-95**	S5···I$^-$	−10.44	C
CHAL-X-36	SCSe···F$^-$	−26.43	C	**CHAL-X-96**	Se4···Br$^-$	−10.06	C
CHAL-X-37	S5···F$^-$	−25.77	C	**CHAL-X-97**	H_2CTe···Cl$^-$	−10.05	C
CHAL-X-38	Te5···Cl$^-$	−25.76	C	**CHAL-X-98**	OCS···Cl$^-$	−9.77	W1-F12
CHAL-X-39	S3···F$^-$	−25.16	C	**CHAL-X-99**	SCS···Br$^-$	−9.41	C
CHAL-X-40	Te7···I$^-$	−24.50	E	**CHAL-X-100**	F_2CSe···Br$^-$	−9.39	C
CHAL-X-41	F_2CSe···F$^-$	−24.46	C	**CHAL-X-101**	S4···Cl$^-$	−9.21	C
CHAL-X-42	Se7···I$^-$	−24.43	E	**CHAL-X-102**	OCSe···I$^-$	−9.15	C
CHAL-X-43	H_2CTe···F$^-$	−24.42	C	**CHAL-X-103**	SCSe···I$^-$	−9.14	C
CHAL-X-44	Te3···Cl$^-$	−24.11	C	**CHAL-X-104**	H_2CTe···Br$^-$	−8.73	C
CHAL-X-45	Se2···I$^-$	−24.03	C	**CHAL-X-105**	OCS···Br$^-$	−8.61	C
CHAL-X-46	OCTe···Cl$^-$	−23.99	C	**CHAL-X-106**	S4···Br$^-$	−8.43	C
CHAL-X-47	S7···Cl$^-$	−23.79	E	**CHAL-X-107**	Se4···I$^-$	−7.95	C
CHAL-X-48	S2···Cl$^-$	−23.70	C	**CHAL-X-108**	F_2CS···Cl$^-$	−7.79	W1-F12
CHAL-X-49	S6···Br$^-$	−23.70	E	**CHAL-X-109**	SCS···I$^-$	−7.46	C
CHAL-X-50	S1···Br$^-$	−23.48	E	**CHAL-X-110**	F_2CSe···I$^-$	−7.20	C
CHAL-X-51	Te5···Br$^-$	−23.14	C	**CHAL-X-111**	S4···I$^-$	−6.92	C
CHAL-X-52	Se4···F$^-$	−22.32	C	**CHAL-X-112**	OCS···I$^-$	−6.85	C
CHAL-X-53	Te3···Br$^-$	−21.81	C	**CHAL-X-113**	H_2CTe···I$^-$	−6.62	C
CHAL-X-54	S2···Br$^-$	−21.62	C	**CHAL-X-114**	H_2CS···F$^-$	−5.20	W1-F12
CHAL-X-55	SCS···F$^-$	−21.25	W1-F12	**CHAL-X-115**	F_2CS···I$^-$	−5.08	C
CHAL-X-56	OCTe···Br$^-$	−21.21	C	**CHAL-X-116**	H_2CSe···Cl$^-$	−3.03	C
CHAL-X-57	S7···Br$^-$	−21.12	E	**CHAL-X-117**	H_2CSe···Br$^-$	−2.40	C
CHAL-X-58	SCTe···Cl$^-$	−21.04	C	**CHAL-X-118**	H_2CSe···I$^-$	−1.48	C
CHAL-X-59	F_2CTe···Cl$^-$	−20.79	C	**CHAL-X-119**	H_2CS···Cl$^-$	−0.66	W1-F12
CHAL-X-60	OCS···F$^-$	−19.55	W1-F12				

[a] DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP). [b] est. DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP).

5.4 The CHAL-N subset

The CHAL-N subset comprises 91 model systems representing chalcogen-nitrogen interactions whose strength range from −1.57 ($F_2CS\cdots NH_3$) to −33.67 kcal mol^{-1} ($TeO_3\cdots NH_3$) with ΔE_{av} = −9.23 kcal mol^{-1}. Reference energies for four dimers are based on the W1-F12 protocol, eight on strategy C, and 79 on strategy E; see Table 7 for details.

Similarly to CHAL-X, the CHAL-N subset contains a series of smaller model systems and a number of larger ones, with some of the latter being inspired by experiments.[155] The 12 smaller dimers contain ammonia as the Lewis base and F_2C=Y, O=C=Y, and S=C=Y (with Y = S, Se and Te) as CB donors. Their interaction energies range from −1.57 (F_2C=S$\cdots NH_3$) to −6.16 kcal mol^{-1} (O=C=Te$\cdots NH_3$) with ΔE_{av} = −3.20 kcal mol^{-1}. The binding strengths decrease when moving from Te to S with ΔE_{av}(Te-based) = −5.09 kcal mol^{-1}, ΔE_{av}(Se-based) = −2.67 kcal mol^{-1}, and ΔE_{av}(S-based) = −1.84 kcal mol^{-1}).

The 79 larger complexes contain the monomers **S1**-**S7**, **Se1**-**Se7**, and **Te1**-**Te7** from the CHAL-X subset paired with various nitrogenous bases (see Fig. 5), namely pyridine (**N1**), N,N-dimethylpyridin-4-amine (**N2**), quinuclidine (**N3**) and 1,4-diazabicyclo[2.2.2]octane (**N4**). With 49 atoms each, complexes between **N3** and **Se7**, and **Te7**, respectively, are the largest of this subset and the entire CHAL336 benchmark set. When assessing the CB donors, the average stabilities of the Te-based complexes increase in the following order: **Te4** (−6.83 kcal mol^{-1}), **Te3** (−8.03 kcal mol^{-1}), **Te5** (−9.61 kcal mol^{-1}), **Te7** (−15.18 kcal mol^{-1}), **Te2** (−16.72 kcal mol^{-1}), **Te6** (−17.12 kcal mol^{-1}), **Te1** (−17.38 kcal mol^{-1}). Similarly, the average stabilities of the Se-based compounds increase in the following order: **Se4** (−4.50 kcal mol^{-1}), **Se3**(−5.30 kcal mol^{-1}), **Se5**(−6.16 kcal mol^{-1}), **Se7**(−9.59 kcal mol^{-1}), **Se1** (−10.63 kcal mol^{-1}), **Se2**(−10.71 kcal mol^{-1}), and **Se6** (−11.10 kcal mol^{-1}). Sulfur-based complexes followed the same order as selenium. When assessing the influence of Lewis bases, **N3**-based dimers show the largest stability (ΔE_{av} = −10.10 kcal mol^{-1}), followed by **N4** (ΔE_{av} = −9.40 kcal mol^{-1}), **N2** (ΔE_{av} = −9.14 kcal mol^{-1}) and **N1** (ΔE_{av} = −7.82 kcal mol^{-1}).

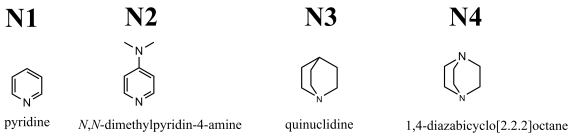

Figure 5: Lewis structure of the large nitrogenous bases (**N1**-**N4**) used in the CHAL-N subset.

Table 7: Reference interaction energies (ΔE in kcal mol^{-1}) of the CHAL-N subset.

Label	System	ΔE	Ref. strategy	Label	System	ΔE	Ref. strategy
CHAL-N-1	TeO_3···NH_3	−33.67	C[a]	**CHAL-N-47**	S2···N2	−7.63	E
CHAL-N-2	SO_3···NH_3	−30.85	W1-F12	**CHAL-N-48**	S6···N2	−7.55	E
CHAL-N-3	SeO_3···NH_3	−28.30	C	**CHAL-N-49**	Te4···N4	−7.41	E
CHAL-N-4	Te1···N2	−19.45	E[b]	**CHAL-N-50**	S1···N4	−7.33	E
CHAL-N-5	Te6···N2	−19.17	E	**CHAL-N-51**	S1···N2	−7.31	E
CHAL-N-6	Te2···N3	−18.55	E	**CHAL-N-52**	Se5···N3	−6.99	E
CHAL-N-7	Te1···N3	−17.95	E	**CHAL-N-53**	Se5···N4	−6.51	E
CHAL-N-8	Te6···N3	−17.81	E	**CHAL-N-54**	Te4···N2	−6.51	E
CHAL-N-9	Te2···N2	−17.22	E	**CHAL-N-55**	Te3···N1	−6.42	E
CHAL-N-10	Te2···N4	−17.16	E	**CHAL-N-56**	S2···N1	−6.41	E
CHAL-N-11	Te1···N4	−16.63	E	**CHAL-N-57**	S6···N1	−6.34	E
CHAL-N-12	Te7···N3	−16.35	E	**CHAL-N-58**	OCTe···NH_3	−6.16	C
CHAL-N-13	Te6···N4	−16.34	E	**CHAL-N-59**	S1···N1	−6.15	E
CHAL-N-14	Te1···N1	−15.49	E	**CHAL-N-60**	Se3···N3	−6.08	E
CHAL-N-15	Te7···N4	−15.24	E	**CHAL-N-61**	Se5···N2	−6.02	E
CHAL-N-16	Te6···N1	−15.16	E	**CHAL-N-62**	S7···N1	−5.99	E
CHAL-N-17	Te2···N1	−13.96	E	**CHAL-N-63**	Se3···N4	−5.71	E
CHAL-N-18	Te7···N1	−13.94	E	**CHAL-N-64**	Te4···N1	−5.46	E
CHAL-N-19	Se2···N3	−12.03	E	**CHAL-N-65**	S5···N3	−5.35	E
CHAL-N-20	Se6···N3	−11.95	E	**CHAL-N-66**	S5···N4	−5.14	E
CHAL-N-21	Se6···N2	−11.84	E	**CHAL-N-67**	Se5···N1	−5.13	E
CHAL-N-22	Se1···N2	−11.45	E	**CHAL-N-68**	Se3···N2	−5.07	E
CHAL-N-23	Se1···N3	−11.31	E	**CHAL-N-69**	Se4···N3	−5.05	E
CHAL-N-24	Se2···N4	−11.18	E	**CHAL-N-70**	S5···N2	−4.94	E
CHAL-N-25	Te5···N3	−11.14	E	**CHAL-N-71**	S3···N3	−4.93	E
CHAL-N-26	Se6···N4	−11.12	E	**CHAL-N-72**	Se4···N4	−4.78	E
CHAL-N-27	Se7···N3	−10.79	E	**CHAL-N-73**	S3···N4	−4.66	E
CHAL-N-28	Se2···N2	−10.70	E	**CHAL-N-74**	F_2CTe···NH_3	−4.64	C
CHAL-N-29	Se1···N4	−10.52	E	**CHAL-N-75**	S3···N2	−4.50	E
CHAL-N-30	Te5···N4	−10.37	E	**CHAL-N-76**	SCTe···NH_3	−4.47	C
CHAL-N-31	Se7···N4	−9.54	E	**CHAL-N-77**	Se3···N1	−4.33	E
CHAL-N-32	Se6···N1	−9.48	E	**CHAL-N-78**	Se4···N2	−4.32	E
CHAL-N-33	Te3···N3	−9.26	E	**CHAL-N-79**	S5···N1	−4.29	E
CHAL-N-34	Se1···N1	−9.23	E	**CHAL-N-80**	S4···N3	−4.17	E
CHAL-N-35	Te5···N2	−9.15	E	**CHAL-N-81**	S4···N4	−4.01	E
CHAL-N-36	Se2···N1	−8.93	E	**CHAL-N-82**	S3···N1	−3.90	E
CHAL-N-37	Te3···N4	−8.65	E	**CHAL-N-83**	S4···N2	−3.87	E
CHAL-N-38	Se7···N1	−8.43	E	**CHAL-N-84**	Se4···N1	−3.86	E
CHAL-N-39	S6···N3	−8.38	E	**CHAL-N-85**	S4···N1	−3.42	E
CHAL-N-40	S2···N3	−8.31	E	**CHAL-N-86**	OCSe···NH_3	−3.15	C
CHAL-N-41	S6···N4	−8.00	E	**CHAL-N-87**	SCSe···NH_3	−2.54	C
CHAL-N-42	Te4···N3	−7.93	E	**CHAL-N-88**	F_2CSe···NH_3	−2.33	C
CHAL-N-43	Te3···N2	−7.81	E	**CHAL-N-89**	OCS···NH_3	−2.11	W1-F12
CHAL-N-44	Te5···N1	−7.80	E	**CHAL-N-90**	SCS···NH_3	−1.83	W1-F12
CHAL-N-45	S2···N4	−7.77	E	**CHAL-N-91**	F_2CS···NH_3	−1.57	W1-F12
CHAL-N-46	S1···N3	−7.70	E				

[a] DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP). [b] est. DLPNO-CCSD(T)/CBS(ma-def2-TZVPP/ma-def2-QZVPP).

The above introduced CHAL-N complexes are formed by σ-hole interactions and we additionally include three dimers stabilized by π-hole ones, namely $TeO_3\cdots NH_3$, $SeO_3\cdots NH_3$, and $SO_3\cdots NH_3$ Interestingly, the Se-based dimer is the least stable of the three ($\Delta E = -28.30$ kcal mol^{-1}), followed by the S-based ($\Delta E = -30.85$ kcal mol^{-1}) and Te-based ones ($\Delta E = -33.67$ kcal mol^{-1}). The values also show that those complexes are more stable than the previously discussed σ-hole-bonded ones.

6 Exemplified test of DFT methods

6.1 Selected density functional approximations

Next, the CHAL336 benchmark set is used for an exemplified test of DFT methods, which constitutes the most comprehensive assessment of such methods for CB interactions. Rather than testing arbitrarily chosen methods, we selected density functional DFAs based on previous studies conducted with the GMTKN55 database. Functionals were chosen based on having been shown to be robust for the entire GMTKN55 benchmark database, for its NCI subcategories, or for some of its individual NCI subsets.[20,33,94,206,218,224] We also included methods that are very popular in the field—regardless of their established robustness or lack thereof—and methods that are relatively new. The included methods cover rungs 2-5 of "Jacob's Ladder".[227] We assess a total of 98 variations of dispersion-corrected and -uncorrected DFAs, and carry out a detailed analysis of 72 of them to identify the most reliable approaches for CB interactions: 22 DFAs belonging to the Generalized Gradient Approximation (GGA), 11 meta-GGA methods, 21 hybrids, and 18 double-hybrid DFAs. We apply different variants of dispersion corrections to the same underlying exchange-correlation approximations. Therefore, we assess 27 unique exchange-correction DFAs. The entire list of assessed DFAs and dispersion corrections are shown in Table 8. Note that other studies have shown that the double hybrid ωB97M(2)[228] is currently the best for GMTKN55,[224] however, we had to discard it from our study for technical reasons, as we observed severe convergence problems

Table 8: List of DFAs and their dispersion corrections used in the benchmark study.[a]

name	dispersion correction
	GGA
PBE[195]	-D3(BJ),[8,b] -D3(0),[7,c] -D4,[134,135,d] -NL[196,e]
revPBE[197]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
BLYP[198–200]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
BP86[198,201,202]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
OLYP[199,200,203]	-D3(BJ),[19] -D3(0),[19] -D4[134,135]
XLYP[199,200,204]	-D3(BJ),[20] -D3(0),[20] -D4[134,135]
OPBE[195,203]	-D3(BJ),[19] -D3(0),[19] -D4[134,135]
	meta-GGA
B97M[205]	-D3(BJ),[33] -D3(0),[33] -D4,[206] -V[205,e]
M06-L[207]	-D3(0)[19]
SCAN[208]	-D3(BJ),[209] -D3(0),[209] -D4[134,135]
TPSS[210]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
	hybrid
B3LYP[91,92]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
BHLYP[211]	-D3(BJ),[19] -D3(0),[19] -D4[134,135]
M06-2X[212]	-D3(0)[19]
M06[212]	-D3(0)[19]
PW6B95[150]	-D3(BJ),[8] -D3(0),[7] -D4[134,135]
MPW1B95[213]	-D3(BJ),[19] -D3(0),[19] -D4[134,135]
ωB97M[214]	-D3(BJ),[33] -D3(0),[33] -D4,[206] -V[214]
ωB97X[215]	-D3(BJ),[33] -D4,[206] -V[215]
	double hybrid
B2PLYP[216]	-D3(BJ),[19] -D3(0),[18] -D4[134,135]
B2NCPLYP[217]	-D3(BJ),[218] -D3(0)[218]
B2GPPLYP[219]	-D3(BJ),[19] -D3(0),[18] -D4[134,135]
DSD-BLYP[220]	-D3(BJ)[19] , -D3(0),[18] -D4[134,135]
DSD-PBEP86[221]	-D3(BJ),[221] -D4[134,135]
SOS0-PBE0-2[222]	-D3(BJ),[218] -D3(0),[218]
ωB97X-2[223,f]	-D3(BJ),[218] -D3(0)[218]
revDSD-PBEP86[224]	-D3(BJ)[224g]

[a] Articles that presented damping parameters for a particular dispersion correction are cited in each case. [b] DFT-D3 with Becke-Johnson damping.[7,8] [c] DFT-D3 with zero damping.[7]
[d] DFT-D4 in its EEQ[225] version including the default three-body[7] correction.[134,135] [e] VV10 correction.[226] [f] In its "TQZ" version[223]
[g] This functional has been parametrized specifically for DFT-D3(BJ).

for ωB97M(2) in QCHEM,[229] most likely due to the fact that the program only allows using the VV10[226] dispersion correction in its full-SCF implementation; for more details on different ways of applying such a correction see Refs 33,196,206 and 230. As double hybrids

make use of a nonlocal, MP2-type correction, we also include MP2 and SCS-MP2;[193] this is of particular interest, as MP2 has previously been used in the computational treatment of CB interactions.[51,58,73,84,86–88] We additionally assess those two methods with DFT-D3-type dispersion corrections, as it was shown how they can be beneficial.[218,231]

All calculations were carried out with an SCF criterion of $10^{-7}\mathrm{E}_h$ and ORCA's quadrature grid "3", followed by a non-iterative step with the larger grid "4". The nonlocal van-der-Waals (vdW) correlation kernel in ORCA was evaluated in the post-SCF manner with ORCA's "vdwgrid4" grid. This usually halves the computer time without the loss of accuracy.[33] The ma-def2-QZVPP AO basis set was used for all calculations, as it is close to the basis-set limit for DFT methods. All GGAs and meta-GGAs were treated with the RI-J approximation for Coulomb integrals and appropriate auxiliary basis sets.[145,232] The second-order perturbation part in double hybrids and MP2 were also computed with the RI approximation and appropriate auxiliary basis sets.[233]

6.2 The effect of including London-dispersion corrections

It has been evident from numerous studies that London-dispersion corrections play a significantly important role in DFT-based treatments of geometries, thermochemistry, kinetics, and NCIs.[8,19,20,35,148,149,206,234–239] Similarly to previous studies on NCIs,[20,142,143,240] it is most likely required to also apply dispersion corrections when dealing with CB interactions. Before we present our detailed analysis of mostly dispersion-corrected DFAs in Section 6.4, we exemplify the effect of the four therein applied London-dispersion corrections for select DFAs.

We begin our discussion with the BLYP, B3LYP, and B2PLYP functionals as BLYP-based representatives of rungs 2, 4 and 5 of Jacob's Ladder and show MDs and MADs of their dispersion-uncorrected and -corrected versions for CHAL336 in Fig. 6. For BLYP, the MAD decreases from 4.72 kcal mol^{-1} to 1.18 kcal mol^{-1} when using DFT-D3(BJ), while it decreases from 3.65 kcal mol^{-1} to 0.95 kcal mol^{-1} for B3LYP, and from 2.07 kcal mol^{-1}

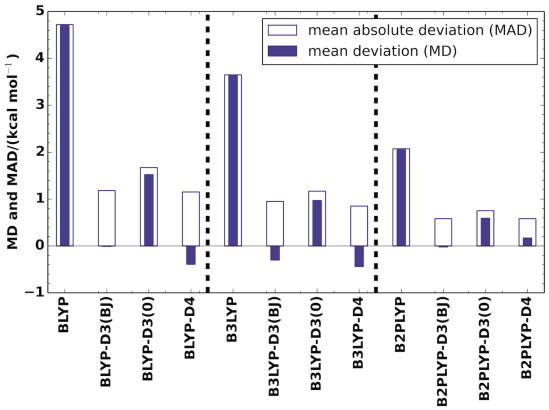

Figure 6: Mean deviations (MDs) and mean absolute deviations (MADs) for the CHAL336 benchmark set for dispersion-corrected and -uncorrected versions of BLYP, B3LYP and BPLYP. A negative MD indicates an overestimation of interaction energies. All MDs and MADs are based on ma-def2-QZVPP calculations.

to 0.58 kcal mol^{-1} for B2PLYP. The resulting dispersion contributions are way above the chemical accuracy of 0.1 kcal mol^{-1} and should consequently not be neglected. Adding a dispersion correction allows a more robust treatment of CB interactions with MDs being closer to zero, for instance, the MD of 4.70 kcal mol^{-1} for BLYP decreases to -0.01 kcal mol^{-1} for BLYP-D3(BJ). Similarly, the MDs decrease from 3.65 kcal mol^{-1} to -0.30 kcal mol^{-1} and from 2.07 kcal mol^{-1} to -0.02 kcal mol^{-1} for B3LYP-D3(BJ) and B2PLYP-D3(BJ), respectively. Adding the DFT-D4 dispersion correction leads to very similar MAD values as for DFT-D3(BJ) model, but with slightly more positive MDs, indicating a slightly underbinding tendency for this correction and functionals (Fig. 6). Moreover, the DFT-D3(0) correction shows even more positive MD values, which is expected due to the nature of its damping function, as discussed below.

While the findings shown in Fig. 6 align with what is expected, the benefit of using a dispersion correction seems to be DFA dependent. This in itself does not necessarily reflect negatively on the dispersion correction, but on the DFA itself. If a dispersion-corrected method performs worse than the underlying uncorrected functional, it means that the latter benefits from fortuitous error compensation, while the dispersion-corrected variant shows a "truer" picture of the DFA's performance; for a similar discussion of this issue in the context

of reaction barrier heights, see Ref. 20. We exemplify the perceived worsening of the statistics for PBE: while its MAD seems to improve from 1.81 kcal mol^{-1} to 1.49 kcal mol^{-1} when adding the DFT-D3(BJ) correction, its MD shows a strong decrease from 0.98 kcal mol^{-1} to -1.24 kcal mol^{-1}, which indicates strong overestimation. Severe overestimation for some DFT-D3(BJ)-corrected DFAs have also been observed for related halogen interactions.[20,109] DFT-D3(BJ) provides a non-zero dispersion energy even for two close-lying fragments, where DFT-D3(0) damps the dispersion energy down to zero in such cases; see Refs 8 or 241 for detailed discussions. This explains why PBE-D3(0) shows a smaller overbinding tendency with MD = -0.58 kcal mol^{-1}. DFT-D4 uses the same type of damping function as DFT-D3(BJ) and indeed PBE-D4's MD is very similar with a value of -1.64 kcal mol^{-1}. Note that this behavior is not solely a problem of Grimme-type, additive dispersion corrections. When using the non-local (NL), van-der-Waals VV10 term, the resulting PBE-NL method shows an even larger overbinding tendency with MD = -1.85 kcal mol^{-1}. This means that a method-inherent error in PBE itself must be the cause for the observed behavior, which is also the potential reason for other DFAs that show similar behavior. We discuss the most-likely explanation for this behavior in the following section.

6.3 A brief analysis of density-driven errors

It is well-established that dispersion corrections typically reduce binding-energy errors in weakly bound complexes.[8,19,20,35,148,149,206,234–240] However, on some occasions it can happen that uncorrected DFAs also seem to be attractive and that their resulting dispersion-corrected versions overestimate the binding energies, as we discussed in the previous section for CB interactions treated with various dispersion-corrected PBE variants. In 2018, Sim, Burke, and co-workers noticed that many DFAs have an unusually large density error for CB interactions,[119] and we suggest that this is the most-likely reason for the observed behavior in Section 6.2. The density error is one of the two components of the notorious self-interaction-error (SIE) of DFT, the other one being the functional error.[119–124] The functional error

arises from the fact that an approximate density functional is used. This means that even for an exact electron density, the resulting energy differs from the one expected for the (unknown) exact functional. The density error arises from the fact that an inaccurate DFA is used during the SCF procedure, thus, contaminating the resulting orbitals and densities. Density-corrected DFT (DC-DFT)[119–124] consists of a non-iterative evaluation of a DFA with an exact and SIE-free density; this would reduce the remaining SIE solely to the functional error. For one-electron systems, the true, exact density is the HF density; for a recently published comprehensive analysis of the one-electron SIE and its components in 74 DFAs, see Ref. 125. Due to the electron-correlation problem, HF cannot deliver the exact density for a many-electron system. However, for practical reasons and due to the fact that HF is SIE free, it has been suggested to use HF densities as a reliable proxy for exact densities in DC-DFT.[119,120,126,242]

In this section, we exemplify the impact of the DC-DFT procedure for ten systems, some of which show large and other small binding strengths; see SI for all systems and numbers (Tables S5-S12). Fig. 7 shows the difference between conventional and DC-DFT with and without the DFT-D3(BJ) dispersion correction for ten functionals for the $H_2Se\cdots I^-$ dimer, which shows one of strongest overestimating tendencies of the ten systems when DFT-D3(BJ) is added. The trends are the same for almost all DFAs, namely that upon using the DC-DFT scheme, MDs become significantly more positive which means that the binding strength of the dimer decreases. Therefore, adding a dispersion correction to DC-DFT results gives significantly smaller deviations. For instance, the average deviation for the ten tested functionals worsens from -0.33 kcal mol^{-1} to -1.65 kcal mol^{-1} when DFT-D3(BJ) is added to the conventional, SCF-based DFT result. On the other hand, DC-DFT's average deviation of 1.13 kcal mol^{-1} decreases to -0.19 kcal mol^{-1} when DFT-D3(BJ) is added.

Table 9 shows MADs and MDs for set of ten dimers. Results are shown for pure DFAs as well as their DFT-D3(BJ), -D4 and -D3(0) corrected versions. The trends are the same for almost all DFAs, namely that upon using the DC-DFT scheme, MDs become more positive

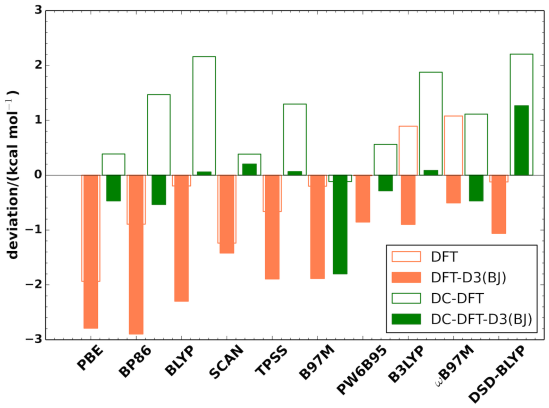

Figure 7: The difference between conventional and DC-DFT with and without the DFT-D3(BJ) correction for a select number of DFAs. The analysis is done for the $H_2Se\cdots I^-$ dimer. The y-axis represents the deviation from the reference for a given DFA (ΔE_{DFA} -$\Delta E_{ref.}$).

Table 9: MADs and MDs (in parenthesis) over ten select systems with and without dispersion corrections for conventional DFT and DC-DFT treatments.

	no dispersion		DFT-D3(0)		DFT-D3(BJ)		DFT-D4	
	DFT	DC-DFT	DFT	DC-DFT	DFT	DC-DFT	DFT	DC-DFT
PBE	0.97(0.33)	1.74(1.74)	0.54(−0.51)	0.93(0.90)	1.00(−1.00)	0.51(0.42)	1.21(−1.21)	0.41(0.20)
BP86	1.82(1.58)	3.05(3.05)	0.81(−0.05)	1.44(1.42)	1.21(−1.09)	0.54(0.38)	0.87(-0.74)	0.80(0.73)
BLYP	3.22(3.18)	4.51(4.51)	1.96(1.46)	2.78(2.78)	1.29(0.27)	1.59(1.59)	1.33(0.19)	1.56(1.52)
SCAN	0.65(−0.36)	0.72(0.01)	0.69(−0.58)	0.62(−0.21)	0.82(−0.72)	0.68(−0.35)	0.79(−0.68)	0.66(−0.31)
TPSS	1.19(1.04)	2.20(2.20)	0.38(−0.14)	1.02(1.02)	0.83(−0.78)	0.39(0.38)	1.02(−1.01)	0.23(0.15)
B97M	1.32(1.28)	2.63(2.32)	0.59(0.36)	2.24(1.40)	2.85(−2.84)	1.80(−1.80)	0.49(−0.48)	1.98(0.55)
PW6B95	0.61(0.61)	1.16(1.16)	0.26(−0.09)	0.47(0.47)	0.56(−0.48)	0.29(0.07)	2.20(−2.20)	1.65(−1.65)
B3LYP	2.42(2.42)	3.14(3.14)	1.07(0.97)	1.69(1.69)	0.41(−0.04)	0.71(0.68)	0.46(−0.01)	0.73(0.71)
ωB97M	2.09(2.09)	2.37(2.37)	1.03(0.94)	1.27(1.22)	0.48(−0.12)	0.61(0.16)	0.98(0.98)	1.27(1.27)
DSD-BLYP	0.68(0.66)	2.11(2.11)	0.69(0.13)	1.59(1.59)	0.97(-0.97)	0.93(0.49)	0.60(0.01)	1.47(1.47)

by a significant value. This can already be seen for the dispersion-uncorrected DFAs, where increases in MDs range from 0.28 kcal mol^{-1} (ωB97M) to 1.47 kcal mol^{-1} (BP86). Similar increases are also seen for MADs. The same decrease in overbinding tendency is observed for the dispersion-corrected results (Table 9). Given that reducing the density error decreases the binding strength of a given dimer, we now understand the previously discussed case where adding a dispersion correction led to a severe overestimation of the interaction energies for some methods. The main source for this problem is a density error that leads to a seemingly better result for a pure DFA, and there is no significant issue with the actual dispersion correction. Indeed, when returning to PBE-D3(BJ), its DC-DFT version has a better MD (0.42 kcal mol^{-1}) than its conventional SCF version (−1.00 kcal mol^{-1}; Table 9).

We advocate that the CHAL336 set is used by method developers in future for more detailed assessments of the density error in CB interactions. As DC-DFT applications are not straightforward to perform and as our main goal is to provide an overview of the current state of the field, we restrict our subsequent analysis of dispersion-corrected, SCF-based DFAs fully aware that the density error plays a role in determining the final ranking. The resulting recommended methods can then be safely applied by users until newer methods are developed that suffer less from density-driven errors for CB or similar interactions.

6.4 Discussion of CHAL336 and its categories

In the following, we analyze how dispersion-corrected DFAs perform for the separate categories of the CHAL336 benchmark set before we concluding with recommendations based on the entire set. The discussion is supported by Fig. 8, which shows average MADs for each assessed rung of Jacob's Ladder for each subset and the complete set. In the same spirit, Table 10 lists the best three DFAs for each rung and subset. Detailed results, including MDs, MADs, RMSDs, and ERs for each functional are shown in the SI (Tables S14-S117).

6.4.1 The CHAL-CHAL subset

For the description of chalcogen-chalcogen interactions, GGAs show the largest spread in MADs, varying from 1.05 kcal mol^{-1} [revPBE-D3(0)] to 6.81 kcal mol^{-1} [XLYP-D3(BJ)] (Fig. 8). The top-3 GGAs according to MADs are revPBE-D3(0), BLYP-D3(BJ), BP86-D3(0) and OLYP-D3(0) with MAD (MD) values of 1.05 (0.31) kcal mol^{-1}, 1.31 (-0.84) kcal mol^{-1}, 1.31 (-0.96) kcal mol^{-1}, and 1.34 (0.91) kcal mol^{-1}, respectively (Table 10). The worst-performing GGAs for this category are XLYP-D3(BJ) (MD = -6.81 kcal mol^{-1}), XLYP-D4 (MD = -5.09 kcal mol^{-1}) and OPBE-D3(BJ) (MD = -3.70 kcal mol^{-1}); they strongly overestimate CB interactions and cannot be recommended (Table S13). Note that XLYP-D3(BJ) and OPBE-D3(BJ) were chosen for this study, as they showed good performance for the intramolecular NCI test sets of GMTKN55 and halogen-bonding interactions,

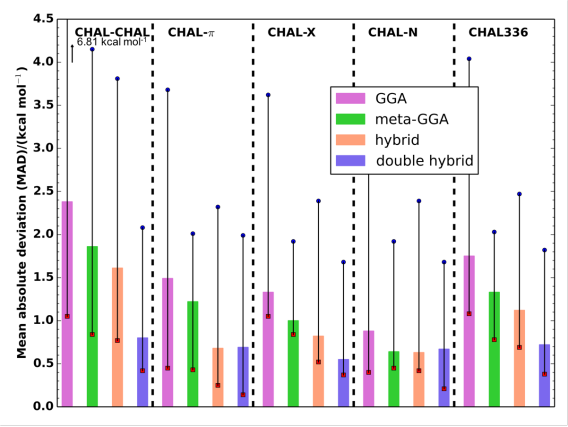

Figure 8: The bars represent the mean absolute deviations (MADs) averaged for GGAs, meta-GGAs, hybrids and double hybrids. The vertical lines show the range of MADs for each functional class. All values are based on ma-def2-QZVPP calculations. Results are shown for the four subsets and the entire CHAL336 set.

respectively; clearly, those findings do not translate to the CHAL-CHAL subset.

Among meta-GGAs, B97M-D3(0) gives the best result (MD = 0.71 kcal mol^{-1} and MAD = 0.84 kcal mol^{-1}), whereas its DFT-D3(BJ) variant yields a high MAD of 4.15 kcal mol^{-1} (MD= −4.15 kcal mol^{-1}) and ranks in the list of worst-performing meta-GGAs (Table S13). The B97M-V and B97M-D4 variants feature as the second- and third-best rung-3 DFAs for this category. This demonstrates the aforementioned impact of the different dispersion corrections. The worst-performing meta-GGAs for this subset are B97M-D3(BJ), SCAN-D3(BJ), SCAN-D4 and SCAN-D3(0) with MAD (MD) values of 4.15 (−4.15) kcal mol^{-1}, 2.61 (−2.52) kcal mol^{-1}, 2.54 (−2.42) kcal mol^{-1}, and 2.41 (−2.25) kcal mol^{-1}, respectively (Table S13).

The MADs for the 21 assessed hybrids range from 0.77 kcal mol^{-1} to 3.81 kcal mol^{-1}. While ωB97M-V, ωB97M-D3(BJ) and ωB97X-V are the best hybrid DFAs for the CHAL-CHAL subset, PW6B95-D4, ωB97X-D4 and ωB97X-D3(BJ) strongly overestimate CB interactions (Table S13).

The average MADs in Fig. 8 are significantly higher for (meta-)GGAs and hybrids than for double hybrids, which means Jacob's Ladder is on average reproduced. For instance, the average MAD values are 2.38 kcal mol^{-1} for GGAs, 1.86 kcal mol^{-1} for meta-GGAs,

Table 10: Top-three functionals according to MADs in each of the four assessed functional classes for the four subsets of CHAL336 and for the entire benchmark set. MDs are reported in parenthesis. All values are in kcal mol^{-1}.

	GGA	meta-GGA	hybrid	double hybrid
CHAL-CHAL	revPBE-D3(0) (1.05, 0.31) BLYP-D3(BJ) (1.31, −0.84) BP86-D3(0) (1.31, −0.96) OLYP-D3(0) (1.34, 0.91)	B97M-D3(0) (0.84, 0.71) B97M-V (0.92, −0.61) B97M-D4 (0.95, −0.80)	ωB97M-V (0.77, −0.28) ωB97M-D3(BJ) (0.84, 0.42) ωB97X-V (0.85, −0.37)	SOS0-PBE0-2-D3(BJ) (0.42, −0.07) ωB97X-2-D3(0) (0.49, −0.11) revDSD-PBEP86-D3(BJ) (0.50,−0.12)
CHAL-π	OLYP-D3(0) (0.45, 0.13) revPBE-D3(0) (0.53, −0.41) BLYP-D3(0) (0.74, 0.28)	B97M-V (0.43, −0.34) M06-L-D3(0) (0.44, −0.04) B97M-D3(0) (0.57, 0.36)	ωB97M-V (0.25, -0.05) ωB97X-V (0.35, 0.33) ωB97M-D3(BJ) (0.36, 0.29)	SOS0-PBE0-2-D3(BJ) (0.14, −0.09) B2PLYP-D4 (0.22, −0.11) B2NCPLYP-D3(0) (0.24, −0.23)
CHAL-X	BP86-D3(0) (1.05, −0.16) revPBE-D3(BJ) (1.18, 0.50) PBE-D3(0) (1.26, −0.78)	B97M-D4 (0.84, −0.06) TPSS-D3(0) (0.85, −0.21) B97M-V (0.94, −0.17)	PW6B95-D3(BJ) (0.52, −0.08) MPW1B95-D3(0) (0.54, 0.08) BHLYP-D4 (0.60, 0.35)	revDSD-PBEP86-D3(BJ) (0.37, 0.08) B2NCPLYP-D3(BJ) (0.38, 0.06) B2NCPLYP-D3(0) (0.38, 0.12) ωB97X-2-D3(0) (0.51, −0.06)
CHAL-N	OLYP-D3(BJ) (0.40, 0.16) revPBE-D3(BJ) (0.50, 0.18) PBE-D3(BJ) (0.54, −0.02)	TPSS-D3(BJ) (0.45, −0.11) SCAN-D4 (0.50, −0.22) SCAN-D3(0) (0.51, −0.12)	ωB97M-V (0.42, 0.23) MPW1B95-D3(BJ) (0.46, 0.39) ωB97X-D3(BJ) (0.46, −0.45) B3LYP-D3(BJ) (0.47, −0.20)	SOS0-PBE0-2-D3(BJ) (0.21, −0.07) revDSD-PBEP86-D3(BJ) (0.26, −0.09) B2NCPLYP-D3(0) (0.30, −0.10)
ALL	BP86-D3(0) (1.08, −0.53) PBE-D3(0) (1.14, −0.58) BLYP-D4 (1.15, −0.39)	B97M-D4 (0.78, −0.28) B97M-V (0.78, −0.14) TPSS-D3(0) (0.84, −0.21)	ωB97M-V (0.69, 0.26) PW6B95-D3(0) (0.72, 0.14) PW6B95-D3(BJ) (0.78, −0.21)	SOS0-PBE0-2-D3(BJ) (0.38, 0.12) revDSD-PBEP86-D3(BJ) (0.39, −0.07) B2NCPLYP-D3(BJ) (0.40, 0.04) B2NCPLYP-D3(0) (0.40, 0.11)

1.61 kcal mol^{-1} for hybrids, and 0.80 kcal mol^{-1} for double hybrids. The same trend can also be observed for RMSDs in Fig. S2 in the SI. Therefore, rung-5 DFAs should be used for chalcogen-chalcogen interactions to obtain reliable results. Among those, SOS0-PBE0-2-D3(BJ) comes in first place (MAD = 0.42 kcal mol^{-1}), closely followed by ωB97X-2-D3(0) (MAD = 0.49 kcal mol^{-1}), and revDSD-PBEP86-D3(BJ) (0.50 kcal mol^{-1}). Despite the latter's very good outcome, we note that the related and older DSD-BLYP-D3(BJ) and DSD-PBEP86-D3(BJ) double hybrids strongly overestimate chalcogen-chalcogen interactions with MAD (MD) values of 2.08 (−2.08) kcal mol^{-1} and 1.21 (−1.21) kcal mol^{-1}, respectively.

6.4.2 CHAL-π subset

For the CHAL-π subset, the average MAD values are significantly lower for double(-hybrids) than for the meta(-GGAs) (1.49 kcal mol^{-1} for GGAs, 1.22 kcal mol^{-1} for meta-GGAs, 0.68 kcal mol^{-1} for hybrids and 0.69 kcal mol^{-1} for double hybrids). We see a similar picture for average RMSDs (see Fig. S2).

Again OLYP-D3(0) (MAD = 0.45 kcal mol^{-1}; MD = 0.13 kcal mol^{-1}), revPBE-D3(0) (MAD = 0.53 kcal mol^{-1}; MD = −0.41 kcal mol^{-1}) and BLYP-D3(0) (MAD = 0.74 kcal mol^{-1}; MD = 0.28 kcal mol^{-1}) rank as the best performing GGAs, while XLYP-D3(BJ)

(MAD = 3.68 kcal mol^{-1}; MD = −3.68 kcal mol^{-1}), XLYP-D4 (MAD = 2.98 kcal mol^{-1}; MD = −2.98 kcal mol^{-1}) and OPBE-D3(BJ) (MAD = 2.49 kcal mol^{-1}; MD = −2.49 kcal mol^{-1}) are the worst performers. Among the assessed meta-GGAs, we recommend B97M-V (MAD = 0.43 kcal mol^{-1}; MD = −0.34 kcal mol^{-1}), M06-L-D3(0) (MAD = 0.44 kcal mol^{-1}; MD = −0.04 kcal mol^{-1}) and B97M-D3(0) (MAD = 0.57 kcal mol^{-1}; MD = 0.36 kcal mol^{-1}), whereas B97M-D3(BJ) (MAD = 2.01 kcal mol^{-1}; MD = −2.01 kcal mol^{-1}), TPSS-D4 and SCAN-D3(BJ) (both with MAD = 1.79 kcal mol^{-1} and MD = −1.79 kcal mol^{-1}) overestimate the chalcogen-π interactions.

Despite hybrids covering the larger MAD range than meta-GGAs, the best methods for this subset are ωB97M-V, ωB97X-V and ωB97M-D3(BJ) with MAD (MD) values of 0.25 (−0.05) kcal mol^{-1}, 0.35 (0.33) kcal mol^{-1} and 0.36 kcal mol^{-1} (0.29) kcal mol^{-1}, respectively.

Moving on to the double hybrids, we see that SOS0-PBE0-2-D3(BJ) ranks as the best double hybrid (MAD = 0.14 kcal mol^{-1}; MD = −0.09 kcal mol^{-1}). With MADs (MDs) of 0.22 (−0.11) kcal mol^{-1} and 0.24 (−0.23) kcal mol^{-1}, B2PLYP-D4 and B2NCPLYP-D3(0) complete the top three of double hybrids. The worst performing double hybrids for chalcogen-π interactions are DSD-BLYP-D3(BJ) (MAD = 1.99 kcal mol^{-1}; MD = −1.99 kcal mol^{-1}), DSD-PBEP86-D3(BJ) (MAD = 1.46 kcal mol^{-1}; MD = −1.46 kcal mol^{-1}) and DSD-BLYP-D3(0) (MAD = 1.30 kcal mol^{-1}; MD = −1.30 kcal mol^{-1}) with MADs even larger than for a large number of rung 2-4 DFAs.

6.4.3 The CHAL-X subset

For the CHAL-X subset, the average MAD values are significantly lower for double hybrids than the other rungs; for instance, average MADs are 1.95 kcal mol^{-1} for GGAs, 1.42 kcal mol^{-1} for meta-GGAs, 1.18 kcal mol^{-1} for hybrids and 0.70 kcal mol^{-1} for double hybrids. We see similar trends for average RMSDs (Fig. S2).

According to Table 10, BP86-D3(0), revPBE-D3(BJ) and PBE-D3(0) yield the best MADs for this category with MAD (MD) values of 1.05 (−0.16) kcal mol^{-1}, 1.18 (0.50)

kcal mol^{-1}, and 1.26 (-0.78) kcal mol^{-1}, respectively. Interestingly, the DFT-D3(BJ) and -D4 versions of XLYP overestimate the interactions in this section: their MADs (MDs) are 3.56 (-2.78) kcal mol^{-1} and 3.62 (-3.17) kcal mol^{-1}, respectively. Contrary to that, XLYP-D3(0) underestimates the interactions (MAD = 2.53 kcal mol^{-1}; MD = 2.38 kcal mol^{-1}). B97M-D4, TPSS-D3(0), and B97M-V are the best meta-GGAs for the CHAL-X subset with MAD (MD) values of 0.84 (-0.06) kcal mol^{-1}, 0.85 (-0.21) kcal mol^{-1}, and 0.94 (-0.17) kcal mol^{-1}, respectively. We observe a large gap when moving to the fourth-best meta-GGA TPSS-D3(BJ) (MAD = 1.22 kcal mol^{-1}; MD = -0.98 kcal mol^{-1}).

The best hybrids for this subset are PW6B95-D3(BJ), MPW1B95-D3(0) and BHLYP-D4 with MAD (MD) values of 0.52 (-0.08) kcal mol^{-1}, 0.54 (0.08) kcal mol^{-1}, and 0.60 (0.35) kcal mol^{-1}, respectively. Interestingly, PW6B95-D4 strongly overestimates the interaction energies (MD = -2.39 kcal mol^{-1}) and is the worst of the tested hybrids. The second- and third-worst performing hybrids—ωB97M-D3(0) and ωB97M-D3(BJ)—underestimate chalcogen-halogen interactions with MDs of 2.28 and 2.10 kcal mol^{-1}, respectively.

revDSD-PBEP86-D3(BJ) ranks as the best DFA for chalcogen-halogen interactions with MAD = 0.37 kcal mol^{-1} and a nearly perfect MD of 0.08 kcal mol^{-1}. This is closely followed by the two DFT-D3 variants of B2NCPLYP. While both B2NCPLYP-D3(0) and B2NCPLYP-D3(BJ) have an MAD value of 0.38 kcal mol^{-1}, the latter shows slightly better MDs (0.06 vs 0.12 kcal mol^{-1}, respectively). Again, DSD-BLYP-D3(BJ), DSD-PBEP86-D3(BJ) and DSD-BLYP-D4 overestimate the interaction energies with MDs of -1.58, -1.05, and -0.87 kcal mol^{-1}, respectively. Moreover, they are also outperformed by a large number of lower-rung DFAs.

6.4.4 The CHAL-N subset

Interestingly, the average MADs of the tested four rungs of Jacob's Ladder are smaller than 1 kcal mol^{-1} with rungs 3-5 being very close to one another: the average MADs are 0.88 (GGAs), 0.64 (meta-GGAs), 0.63 (hybrids) and 0.67 kcal mol^{-1} (double-hybrids). According

to Table 10, OLYP-D3(BJ), revPBE-D3(BJ) and PBE-D3(BJ) are the best GGAs with MAD (MD) values of 0.40 (0.16) kcal mol^{-1}, 0.50 (0.18) kcal mol^{-1}, and 0.54 ($-$0.02) kcal mol^{-1}, respectively. Among the worst-performing GGAs, we find XLYP-D4, XLYP-D3(BJ) and OLYP-D3(0) with MAD (MD) values of 1.80 ($-$1.46) kcal mol^{-1}, 1.77 ($-$1.47) kcal mol^{-1}, and 1.44 (1.44) kcal mol^{-1}, respectively. In fact, XLYP-D4 and XLYP-D3(BJ) were listed as the least accurate methods for the previous subsets as well; clearly showing that these methods are not ideal for CB interactions. Among the meta-GGAs, TPSS-D3(BJ) ranks in first position (MAD = 0.45 kcal mol^{-1}; MD = $-$0.11 kcal mol^{-1}). It is interesting to see that SCAN-D4 (MAD = 0.50 kcal mol^{-1}; MD = $-$0.22 kcal mol^{-1}) and SCAN-D3(0) (MAD = 0.51 kcal mol^{-1}; MD = $-$0.12 kcal mol^{-1}) rank as the next best meta-GGAs, while they were among the worst-performing DFAs for the CHAL-X subset. This indicates a lack of robustness for them. Among the assessed hybrids, we recommend ωB97M-V (MAD = 0.42 kcal mol^{-1}; MD = 0.23 kcal mol^{-1}), MPW1B95-D3(BJ) (MAD = 0.46 kcal mol^{-1}; MD = 0.39 kcal mol^{-1}) and ωB97X-D3(BJ) (MAD = 0.46 kcal mol^{-1}; MD = $-$0.45 kcal mol^{-1}), whereas PW6B95-D4 (MAD = 1.16 kcal mol^{-1}; MD = $-$1.16 kcal mol^{-1}), ωB97X-D4 (MAD = 1.05 kcal mol^{-1}; MD = $-$1.05 kcal mol^{-1}) and ωB97M-D4 (MAD = 0.80 kcal mol^{-1}; MD = 0.72 kcal mol^{-1}) rank as the worst-performing hybrids (Table S13).

Despite the double-hybrid MADs covering a larger range than rungs 2-4 (Fig. 8), the best methods for this subset are SOS0-PBE0-2-D3(BJ), revDSD-PBEP86-D3(BJ), B2NCPLYP-D3(0), and B2NCPLYP-D3(BJ) with MADs of 0.21, 0.26, 0.30 and 0.32 kcal mol^{-1}, respectively. DSD-BLYP-D3(BJ), DSD-BLYP-D4, and DSD-BLYP-D3(0) are the worst double hybrids in this section with MADs larger than for the worst meta-GGAs and hybrids due to large overestimation of interaction energies, as can be seen from their MDs ($-$1.63 kcal mol^{-1}, $-$1.24 kcal mol^{-1} and $-$1.22 kcal mol^{-1}, respectively).

6.4.5 Final analysis of the entire CHAL336 database

So far, we have discussed the different categories of the CHAL336 benchmark set separately and gave individual recommendations and warnings for them. However, we also saw that the best methods for one category may not be the best for another. While this may not be a problem if a user is interested in a very specialized problem that falls into one of the four categories, answering a more general question around CB interactions warrants the application of a reasonably robust and accurate DFA that works well across the entire benchmark set. We end the discussion of our results with such an analysis.

The Jacob's Ladder scheme is well reproduced for the entire CHAL336 set and that of the worst methods for a particular rung have larger MADs than the worst of any lower rungs. The average MAD for GGAs is 1.75 kcal mol^{-1}, which is closely followed by meta-GGAs (1.33 kcal mol^{-1}) and hybrids (1.12 kcal mol^{-1}). Double hybrids have the lowest average MAD with 0.72 kcal mol^{-1}. Again, the same trend can be observed for average RMSDs (see Fig. S2)

Fig. 9 shows the final ranking of all tested dispersion-corrected DFAs according to their MADs. The best method of the entire study is surprisingly the non-empirical double hybrid SOS0-PBE0-2-D3(BJ) (MAD = 0.38 kcal mol^{-1} and MD = 0.12 kcal mol^{-1}; see Table 10). While this particular DFA has been found to be the best non-empirical double-hybrid and also the best method for the calculation of reaction energies and barrier heights in enzymatically catalyzed reactions,[181,218] a broad study on GMTKN55 showed it to be surpassed by six other semi-empirical double hybrids.[218] Our hypothesis is that the latter may suffer from a larger density errors, which is why SOS0-PBE0-2-D3(BJ) appears to perform so well for CHAL336. That being said, SOS0-PBE0-2-D3(BJ) is closely followed by the two semi-empirical double hybrids revDSD-PBEP86-D3(BJ) and B2NCPLYP-D3(BJ), with MADs (MDs) of 0.39 (-0.07) kcal mol^{-1} and 0.40 (0.04) kcal mol^{-1}, respectively (Table 10). The first of those two is a recently developed method and one of the best for the GMTKN55 database,[224] while the latter has been developed for the treatment of NCIs. The DFT-D3(0)

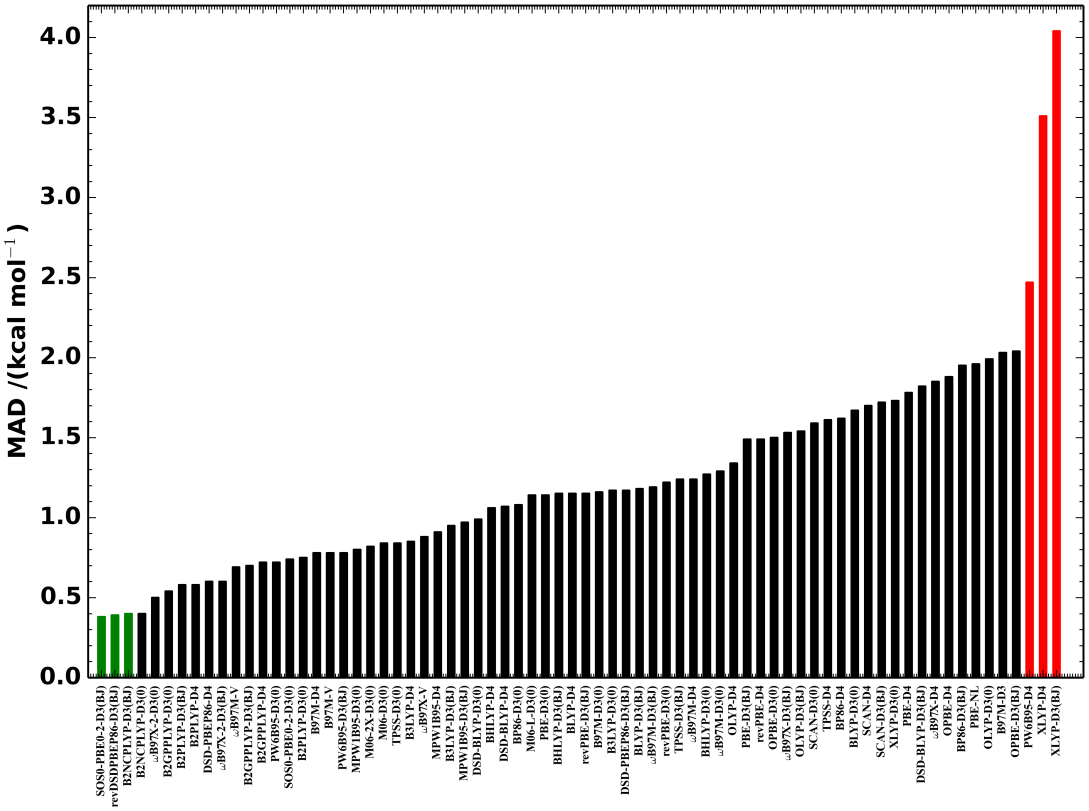

Figure 9: Overview of the MADs (kcal mol^{-1}) of all 72 tested dispersion-corrected DFAs for the CHAL336 benchmark set. The best three DFAs are highlighted in green, the worst three in red.

version of B2NCPLYP competes with its DFT-D3(BJ) variant and the two methods only differ in their MDs with the latter having a slightly better value (MD = 0.11 vs. 0.04 kcal mol^{-1}). We also note that a total of seven additional semi-empirical double hybrids follow before the first hybrid DFA appears in the overall ranking, showing once again the general superiority of double hybrids.

As mentioned in the introduction, MP2 has been a popular method to treat CB interactions in the past and while it is computationally similar to double hybrids, the results in Table 11 clearly show that it does not outperform the latter in either of the four subsets nor the entire benchmark set. Its MAD of 1.39 kcal mol^{-1} is even worse than seven of our assessed dispersion-corrected GGAs, namely BP86-D3(0), BLYP-D4, PBE-D3(0), revPBE-D3(BJ), BLYP-D3(BJ), revPBE-D3(0) and OLYP-D4 (Fig. 9). Allowing for a more balanced treat-

Table 11: MADs (in kcal mol^{-1}) for MP2 and SCS-MP2 approaches for CHAL336 and its four subsets.

	CHAL-CHAL	CHAL-π	CHAL-X	CHAL-N	ALL
MP2	0.88	2.65	1.44	1.51	1.39
MP2-D3(BJ)	0.77	2.29	1.27	0.78	1.07
MP2-D3(0)	0.79	2.29	1.29	1.05	1.16
SCS-MP2	1.80	0.39	1.07	0.49	1.07
SCS-MP2-D3(BJ)	0.45	0.83	0.73	0.86	0.69
SCS-MP2-D3(0)	0.56	0.65	0.59	0.77	0.64

ment of same and opposite spin electron correlation in SCS-MP2 reduces the MAD to 1.07 kcal mol^{-1}, which can be further reduced to 0.69 kcal mol^{-1} when combined with the DFT-D3(BJ) dispersion correction (Table 11). This makes SCS-MP2-D3(BJ) competitive with some hybrid DFAs, but it is still worse than many double hybrids (Fig. 9). Consequently, we do not see any reason why MP2-based methods should be used to treat CB interactions.

The best rung-4 methods for CB interactions are the range-separated hybrid ωB97M-V, and the global hybrids PW6B95-D3(0) and PW6B95-D3(BJ) with MADs (MDs) of 0.69 (0.26) kcal mol^{-1}, 0.72 (0.14) kcal mol^{-1}, and 0.78 ($-$0.21) kcal mol^{-1}, respectively (Table 10). These results align well with the general recommendations for the GMTKN55 database and confirm the overwhelming robustness of those functionals.[20,206] Even when corrected with DFT-D3(BJ), the popular B3LYP functional only ranks in 27th position of all tested methods with an MAD of 0.95 kcal mol^{-1}, and we do not recommend its usage.

The recommended hybrid functionals are good alternatives to double hybrids if the latter cannot be applied due to system size or other computational bottlenecks. Generally, it is not advised to use lower-rung methods in NCIs,[19,20] which is why we do not discuss them any further. The interested user can obtain the relevant information from Fig. 9, Table 10, and Table S13.

Finally, we would like to draw a comparison with the related halogen interactions. As reported in Refs 109 and 119, a number of dispersion-corrected approaches showed systematic overestimation of halogen-bonding interactions as well. It is worthwhile to point out that the best performers for the HAL59 set in GMTN55 were found to be OLYP-D3(BJ)

(GGA), B97M-D4 (meta-GGA), ωB97M-V (hybrid), as well as B2NCPLYP-D3(BJ) (double hybrid).[20,206,218] Our work therefore shows similarities with halogen-interactions for some methods, but also unique differences for others, which further justifies the development of CHAL336.

7 Summary and conclusions

We presented the—to our knowledge—largest benchmark set for chalcogen-bonding (CB) interactions (CHAL336) and used it in one of the most comprehensive DFT benchmark studies for such interactions. CHAL336 consists of 336 dimers and for each tested method, the user has to conduct 1008 single point calculations. The 336 dimers range from being small (4 atoms) to being relatively large (49 atoms) and are divided into four main categories: chalcogen-chalcogen, chalcogen-π, chalcogen-halogen, and chalcogen-nitrogen interactions. Most systems are stabilized by σ-hole interactions, while about 19 % are formed completely or partially by π-hole interactions.

We first obtained reliable geometries at an adequate level of theory that properly addresses London-dispersion interactions before we proceeded with a detailed analysis on how to obtain reference values with the best possible accuracy-cost ratio. When feasible, we used the accurate W1-F12 composite scheme to obtain reference values for CHAL336. We opted to use the efficient DLPNO-CCSD(T) method with TightPNO settings as an alternative to W1-F12. In this case, we obtained CBS-limit results from two-point extrapolations conducted with triple- and quadruple-ζ basis sets. After careful consideration of different basis-set combinations we used the minimally augmented Ahlrichs-type basis sets ma-def2-TZVPP and ma-def2-QZVPP. For larger systems for which that approach was not feasible, we used a composite scheme that estimated DLPNO-CCSD(T)/CBS numbers from lower-level DLPNO-SCS-MP2/CBS results that were combined with a triple-ζ-based correction for any shortcomings in the description of correlation energies. The latter strategy turned out

to be a reliable alternative to properly extrapolated DLPNO-CCSD(T)/CBS results with the exception of chalcogen-fluoride interactions.

The subsequent DFT benchmark study provided useful insights for both method developers and users. Unsurprisingly, London-dispersion corrections provided a significant contribution to the total binding energy, and as such, CB interactions should not be calculated without them. However, for some density functional approximations (DFAs), a dispersion-corrected result turned out to be worse than for the pure DFA due to significant overestimation of interaction energies. Inspired by earlier work in Ref. 119, we showed such behavior does not reflect negatively on the dispersion correction itself but indicated that the underlying DFA suffers from a large density-driven error. Further analysis of density-corrected DFT (DC-DFT) results corroborated that hypothesis and showed that DC-DFT results paired with dispersion corrections had the expected improved behavior. Therefore, CHAL336 may be useful for the development and assessment of new DFAs with smaller density-driven errors.

As DC-DFT is not a technique for routine applications, we investigated conventional DFAs to provide method users with an overview of how current methodologies perform for CB interactions. For this purpose, we assessed 98 variations of dispersion-corrected and -uncorrected DFAs and performed a detailed analysis of 72 of them to identify robust and reliable approaches. These 72 approaches cover rungs 2-5 of Jacob's Ladder and were 22 GGAs, 11 meta-GGAs, 21 hybrids and 18 double hybrids.

We first discussed those DFAs that performed well for the four different categories of the CHAL336 benchmark set before concluding with a comprehensive analysis of the entire set. We reproduced the Jacob's Ladder picture for the entire set and double-hybrid DFAs were on average more accurate than the lower rungs. They also turned out to be by far more accurate than the equally expensive MP2 and SCS-MP2 approaches, even when those had been combined with London-dispersion corrections. While MP2 has been used frequently to calculate CB interactions in the past, we see no reason why such a strategy should be

continued in the future.

The best method for the entire study was the non-empirical double hybrid SOS0-PBE0-2-D3(BJ). It was closely followed by the semi-empirical double hybrids revDSD-PBEP86-D3(BJ) and B2NCPLYP-D3(BJ), which also showed more robust performance in other contexts.[218,224] In general, double hybrids should be the method of choice and one should refer to lower rung methods only if double hybrid calculations are computationally not feasible. In such a scenario, we can recommend the range-separated hybrid ωB97M-V and the global hybrids PW6B95-D3(0) and PW6B95-D3(BJ). We do not advise to use the popular B3LYP functional nor any other lower-rung DFAs for CB interactions. However for completeness reasons we mention in passing that the best meta-GGAs were B97M-D4, B97M-V, and TPSS-D3(0), while the best GGAs turned out to be BP86-D3(0), PBE-D3(0), and BLYP-D4.

The recommended methods, particularly the double hybrids, can be used with any standard quantum-chemical code even if they have not been implemented via a keyword; see original references for details. As such, there is no technical roadblock from adopting our recommendations. In fact, we hope to inspire a change in computational protocols surrounding CB interactions that leads away from the commonly used methods, to the more robust and accurate ones recommended herein. Users and developers alike can set up their own calculations easily with the information provided in this manuscript and the SI. As such, we hope that the CHAL336 benchmark set will make a valuable contribution to future developments and applied studies.

Acknowledgement

We thank Prof. Jan M. L. Martin for an insightful discussion about Dunning basis sets. This work was supported by the Australian Research Council within the Discovery Project scheme (DP180101413). N. M. acknowledges a 'Melbourne International Engagement Award

(MIEA)' offered through the Melbourne India Postgraduate Program and a 'Melbourne Research Scholarship'. T. F. acknowledges the Australian Commonwealth Government for a Research Training Program Scholarship. L. G. is grateful for generous allocations of computational resources from the National Computational Infrastructure (NCI) Facility within the National Computational Merit Allocation Scheme (project fk5), and Research Platform Services (ResPlat) at The University of Melbourne (project punim0094). This research was also supported by the sustaining and strengthening merit-based access to the NCI LIEF Grant (LE190100021) facilitated by The University of Melbourne.

8 Notes

The authors declare no competing financial interest.

Supporting Information Available

The SI consists of four separate files. The first part (pdf file) contains additional DLPNO-CCSD(T) statistics, details on the DC-DFT analysis, RMSD-based analyses of DFT results, information on the worst-performing DFT methods, and statistical results for all assessed methods. The second part (zip file) contains all dimer and monomer geometries including information on any charged species (see included README file). The third part (zip file) contains information on the individual energy components to determine reference values according to strategies C and E. The fourth part (zip file) contains all DFT and MP2 data for each system (see included README file).

References

(1) Hobza, P.; Havlas, Z. Blue-shifting hydrogen bonds. *Chem. Rev.* **2000**, *100*, 4253–4264.

(2) Esrafili, M. D.; Elmi, F.; Hadipour, N. L. Density functional theory investigation of

hydrogen bonding effects on the oxygen, nitrogen and hydrogen electric field gradient and chemical shielding tensors of anhydrous chitosan crystalline structure. *J. Phys. Chem. A* **2007**, *111*, 963–970.

(3) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

(4) Ángyán, J.; Dobson, J.; Jansen, G.; Gould, T. *London dispersion forces in molecules, solids and nano-structures: an introduction to physical models and computational methods*; Royal Society of Chemistry, 2020.

(5) Stone, A. *The theory of intermolecular forces, Second Edition*; oUP oxford, 2013.

(6) El Kerdawy, A.; Murray, J. S.; Politzer, P.; Bleiziffer, P.; Hesselmann, A.; Görling, A.; Clark, T. Directional noncovalent interactions: repulsion and dispersion. *J. Chem. Theory Comput.* **2013**, *9*, 2264–2275.

(7) 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.

(8) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(9) Jurečka, P.; Šponer, 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.

(10) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, $\delta_{MP2}^{CCSD(T)}$: 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.

(11) Rezác, 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.

(12) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C.; Goddard III, W. A. Evaluation of B3LYP, X3LYP, and M06-class density functionals for predicting the binding energies of neutral, protonated, and deprotonated water clusters. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.

(13) Manna, D.; Kesharwani, M. K.; Sylvetsky, N.; Martin, J. M. Conventional and explicitly correlated ab initio benchmark study on water clusters: revision of the BEGDB and WATER27 data sets. *J. Chem. Theory Comput.* **2017**, *13*, 3136–3152.

(14) Setiawan, D.; Kraka, E.; Cremer, D. Strength of the pnictogen bond in complexes involving group Va elements N, P, and As. *J. Phys. Chem. A* **2015**, *119*, 1642–1656.

(15) Kozuch, S.; Martin, J. M. L. Halogen bonds: benchmarks and theoretical analysis. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.

(16) Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark calculations of noncovalent interactions of halogenated molecules. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.

(17) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate description of intermolecular interactions involving ions using symmetry-adapted perturbation theory. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.

(18) Goerigk, L.; Grimme, S. Efficient and accurate double-hybrid-meta-GGA density functionals — Evaluation with the extended GMTKN30 database for general main group

thermochemistry, kinetics and noncovalent interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

(19) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.

(20) 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.

(21) Schwabe, T.; Grimme, S. Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.

(22) Goerigk, L.; Grimme, S. A general database for main group thermochemistry, kinetics, and noncovalent interactions — Assessment of common and reparameterized (meta-)GGA density functionals. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.

(23) Grimme, S.; Steinmetz, M.; Korth, M. How to compute isomerization energies of organic molecules with quantum chemical methods. *J. Org. Chem.* **2007**, *72*, 2118–2126.

(24) Grimme, S. Seemingly simple stereoelectronic effects in alkane isomers and the implications for Kohn–Sham density functional theory. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460–4464.

(25) Gruzman, D.; Karton, A.; Martin, J. M. Performance of Ab initio and density functional methods for conformational equilibria of C_nH_{2n+2} alkane isomers (n= 4- 8). *J. Phys. Chem. A* **2009**, *113*, 11974–11983.

(26) Kesharwani, M. K.; Karton, A.; Martin, J. M. L. Benchmark ab initio conformational energies for the proteinogenic amino acids through explicitly correlated methods. Assessment of density functional methods. *J. Chem. Theory Comput.* **2016**, *12*, 444–454.

(27) Řeha, D.; Valdes, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S. Structure and IR spectrum of phenylalanyl–glycyl–glycine tripetide in the gas-phase: IR/UV experiments, Ab initio quantum chemical calculations, and molecular dynamic simulations. *Chem. - Eur. J.* **2005**, *11*, 6803–6817.

(28) Goerigk, L.; Karton, A.; Martin, J. M. L.; Radom, L. Accurate quantum chemical energies for tetrapeptide conformations: why MP2 data with an insufficient basis set should be handled with caution. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7028–7031.

(29) Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. The melatonin conformer space: benchmark and assessment of wave function and DFT methods for a paradigmatic biological and pharmacological molecule. *J. Phys. Chem. A* **2013**, *117*, 2269–2277.

(30) Csonka, G. I.; French, A. D.; Johnson, G. P.; Stortz, C. A. Evaluation of density functionals and basis sets for carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 679–692.

(31) Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Šponer, J. Quantum chemical benchmark study on 46 RNA backbone families using a dinucleotide unit. *J. Chem. Theory Comput.* **2015**, *11*, 4972–4991.

(32) Kozuch, S.; Bachrach, S. M.; Martin, J. M. Conformational equilibria in butane-1,4-diol: a benchmark of a prototypical system with strong intramolecular H-bonds. *J. Phys. Chem. A* **2014**, *118*, 293–303.

(33) Najibi, A.; Goerigk, L. The nonlocal kernel in van der Waals density functionals as an additive correction: an extensive analysis with special emphasis on the B97M-V and ωB97M-V approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738.

(34) Goerigk, L. How do DFT-DCP, DFT-NL, and DFT-D3 compare for the description of London-dispersion effects in conformers and general thermochemistry? *J. Chem. Theory Comput.* **2014**, *10*, 968–980.

(35) Grimme, S.; Schreiner, P. R. Steric crowding can stabilize a labile molecule: solving the hexaphenylethane riddle. *Angew. Chem., Int. Ed.* **2011**, *50*, 12639–12642.

(36) Wagner, J. P.; Schreiner, P. R. London dispersion in molecular chemistry—reconsidering steric effects. *Angew. Chem., Int. Ed.* **2015**, *54*, 12274–12296.

(37) Schreiner, P. R. Dispersion interactions. *Beilstein J. Org. Chem.* **2018**, *14*, 3076–3077.

(38) Song, L.; Schoening, J.; W:olper, C.; Schulz, S.; Schreiner, P. R. Role of London Dispersion Interactions in Ga-Substituted Dipnictenes. *Organometallics* **2019**, *38*, 1640–1647.

(39) Quesada-Moreno, M. M.; Pinacho, P.; Perez, C.; Sekutor, M.; Schreiner, P. R.; Schnell, M. Dispersion and hydrogen bond interactions in large complexes: the diadamanthyl ether case. 74th International Symposium on Molecular Spectroscopy. 2019.

(40) Fokin, A. A.; Bahonsky, V. V.; Koso, T. V.; Hoc, N. T.; Serafin, M.; Zhuk, T. S.; Rodionov, V. M.; Schreiner, P. R. Noncovalent interactions in crowded olefinic radical cations. *J. Org. Pharm. Chem.* **2020**, *18*, 05–13.

(41) Riley, K. E.; Murray, J. S.; Politzer, P.; Concha, M. C.; Hobza, P. Br··· O complexes as probes of factors affecting halogen bonding: interactions of bromobenzenes and bromopyrimidines with acetone. *J. Chem. Theory Comput.* **2009**, *5*, 155–163.

(42) Del Bene, J. E.; Alkorta, I.; Sánchez-Sanz, G.; Elguero, J. Structures, energies, bonding, and NMR properties of pnictogen complexes H_2XP: NXH_2 (X=H, CH_3, NH_2, OH, F, Cl). *J. Phys. Chem. A* **2011**, *115*, 13724–13731.

(43) Del Bene, J. E.; Alkorta, I.; Sánchez-Sanz, G.; Elguero, J. Interplay of F–H··· F hydrogen bonds and P···N pnictogen bonds. *J. Phys. Chem. A* **2012**, *116*, 9205–9213.

(44) Novák, M.; Foroutan-Nejad, C.; Marek, R. Asymmetric bifurcated halogen bonds. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6440–6450.

(45) Alkorta, I.; Rozas, I.; Elguero, J. Molecular complexes between silicon derivatives and electron-rich groups. *J. Phys. Chem. A* **2001**, *105*, 743–749.

(46) Bauzá, A.; Mooibroek, T. J.; Frontera, A. Tetrel-bonding interaction: rediscovered supramolecular force? *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 12317–12321.

(47) Grabowski, S. J. Tetrel bond–σ-hole bond as a preliminary stage of the SN 2 reaction. *Phys. Chem. Chem. Phys.* **2014**, *16*, 1824–1834.

(48) Scheiner, S. Weak H-bonds. Comparisons of CH···O to NH···O in proteins and PH···N to direct P···N interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13860–13872.

(49) Scheiner, S. Effects of multiple substitution upon the P···N noncovalent interaction. *Chem. Phys.* **2011**, *387*, 79–84.

(50) Alkorta, I.; Elguero, J.; Del Bene, J. E. pnictogen bonded complexes of PO_2X (X= F, Cl) with nitrogen bases. *J. Phys. Chem. A* **2013**, *117*, 10497–10503.

(51) Wang, W.; Ji, B.; Zhang, Y. Chalcogen bond: a sister noncovalent bond to halogen bond. *J. Phys. Chem. A* **2009**, *113*, 8132–8135.

(52) Scheiner, S. Detailed comparison of the pnictogen bond with chalcogen, halogen, and hydrogen bonds. *Int. J. Quantum Chem.* **2013**, *113*, 1609–1620.

(53) Adhikari, U.; Scheiner, S. Effects of charge and substituent on the S···N chalcogen bond. *J. Phys. Chem. A* **2014**, *118*, 3183–3192.

(54) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16789–16794.

(55) Metrangolo, P.; Carcenac, Y.; Lahtinen, M.; Pilati, T.; Rissanen, K.; Vij, A.; Resnati, G. Nonporous organic solids capable of dynamically resolving mixtures of diiodoperfluoroalkanes. *Science* **2009**, *323*, 1461–1464.

(56) Jabłoński, M. Energetic and geometrical evidence of nonbonding character of some intramolecular halogen···oxygen and other Y···Y interactions. *J. Phys. Chem. A* **2012**, *116*, 3753–3764.

(57) Brammer, L. Halogen bonding, chalcogen bonding, pnictogen bonding, tetrel bonding: origins, current status and discussion. *Faraday Discuss.* **2017**, *203*, 485–507.

(58) Esrafili, M. D.; Saeidi, N.; Baneshi, M. M. Chalcogen–chalcogen interactions in furan-YHX and thiophene-YHX complexes (X= F, Cl, Br; Y= S, Se): an ab initio study. *Bull. Chem. Soc. Jpn.* **2015**, *88*, 1683–1692.

(59) Burling, F. T.; Goldstein, B. M. Computational studies of nonbonded sulfur-oxygen and selenium-oxygen interactions in the thiazole and selenazole nucleosides. *J. Am. Chem. Soc.* **1992**, *114*, 2313–2320.

(60) Taylor, J. C.; Markham, G. D. The bifunctional active site of S-adenosylmethionine synthetase Roles of the active site aspartates. *J. Biol. Chem.* **1999**, *274*, 32909–32914.

(61) Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. Statistical characterization of non-bonded S··· O interactions in proteins. *Chem. Lett.* **2001**, *30*, 132–133.

(62) Iwaoka, M.; Takemoto, S.; Tomoda, S. Statistical and theoretical investigations on the directionality of nonbonded S···O interactions. Implications for molecular design and protein engineering. *J. Am. Chem. Soc.* **2002**, *124*, 10613–10620.

(63) Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. Weak nonbonded S··· X (X= O, N, and S) interactions in proteins. Statistical and theoretical studies. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 1611–1625.

(64) Coonan, M. H.; Craven, I. E.; Hesling, M. R.; Ritchie, G. L.; Spackman, M. A. Anisotropic molecular polarizabilities, dipole moments, and quadrupole moments of (CH2) 2X,(CH3) 2X, and C4H4X (X= O, S, Se). Comparison of experimental results and ab initio calculations. *J. Phys. Chem.* **1992**, *96*, 7301–7307.

(65) Iwaoka, M.; Tomoda, S. Nature of the intramolecular Se···N nonbonded interaction of 2-selenobenzylamine derivatives. An experimental evaluation by ^{1}H, ^{77}Se, and ^{15N} NMR spectroscopy. *J. Am. Chem. Soc.* **1996**, *118*, 8077–8084.

(66) Minkin, V. I.; Minyaev, R. M. Cyclic aromatic systems with hypervalent centers. *Chem. Rev.* **2001**, *101*, 1247–1266.

(67) Sanz, P.; Yáñez, M.; Mó, O. The role of chalcogen–chalcogen interactions in the intrinsic basicity and acidity of β-chalcogenovinyl (thio) aldehydes HC(=X)-CH=CH-CYH (X= O, S; Y= Se, Te). *Chem. - Eur. J.* **2002**, *8*, 3999–4007.

(68) Hayashi, S.; Nakanishi, W. Noncovalent Z···Z (Z= O, S, Se, and Te) interactions: how do they operate to control fine structures of 1, 8-dichalcogene-substituted naphthalenes? *Bull. Chem. Soc. Jpn.* **2008**, *81*, 1605–1615.

(69) Sánchez-Sanz, G.; Alkorta, I.; Elguero, J. Theoretical study of the HXYH dimers (X, Y= O, S, Se). Hydrogen bonding and chalcogen–chalcogen interactions. *Mol. Phys.* **2011**, *109*, 2543–2552.

(70) Sánchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J. Intermolecular weak interactions in HTeXH dimers (X= O, S, Se, Te): hydrogen bonds, chalcogen–chalcogen contacts and chiral discrimination. *ChemPhysChem* **2012**, *13*, 496–503.

(71) Esrafili, M. D.; Mohammadian-Sabet, F. An ab initio study on chalcogen–chalcogen bond interactions in cyclic $(SHX)_3$ complexes (X= F, Cl, CN, NC, CCH, OH, OCH_3, NH_2). *Chem. Phys. Lett.* **2015**, *628*, 71–75.

(72) Fellowes, T.; Harris, B. L.; White, J. M. Experimental evidence of chalcogen bonding at oxygen. *Chem. Commun.* **2020**, *56*, 3313–3316.

(73) Esrafili, M. D.; Mohammadian-Sabet, F. Homonuclear chalcogen–chalcogen bond interactions in complexes pairing YO_3 and YHX molecules (Y= S, Se; X= H, Cl, Br, CCH, NC, OH, OCH_3): influence of substitution and cooperativity. *Int. J. Quantum Chem.* **2016**, *116*, 529–536.

(74) Bondybey, V.; English, J. Infrared spectra of SO_3 polymers and complexes in rare gas matrices. *J. Mol. Spectrosc.* **1985**, *109*, 221–228.

(75) Larson, L. J.; Kuno, M.; Tao, F.-M. Hydrolysis of sulfur trioxide to form sulfuric acid in small water clusters. *J. Chem. Phys.* **2000**, *112*, 8830–8838.

(76) Loerting, T.; Liedl, K. R. Water-mediated proton transfer: a mechanistic investigation on the example of the hydration of sulfur oxides. *J. Phys. Chem. A* **2001**, *105*, 5137–5145.

(77) Mo, Y.; Gao, J. Polarization and charge-transfer effects in Lewis acid-base complexes. *J. Phys. Chem. A* **2001**, *105*, 6530–6536.

(78) Merino, G.; Bakhmutov, V. I.; Vela, A. Do cooperative proton-hydride interactions explain the gas-solid structural difference of BH_3NH_3? *J. Phys. Chem. A* **2002**, *106*, 8491–8494.

(79) Pszona, M.; Haupa, K.; Bil, A.; Mierzwicki, K.; Szewczuk, Z.; Mielke, Z. Clustering of sulfamic acid: ESI MS and theoretical study. *J. Mass Spectrom.* **2015**, *50*, 127–135.

(80) Azofra, L. M.; Alkorta, I.; Scheiner, S. Strongly bound noncovalent $(SO_3)_n$:H_2CO complexes (n=1, 2). *Phys. Chem. Chem. Phys.* **2014**, *16*, 18974–18981.

(81) Azofra, L. M.; Alkorta, I.; Scheiner, S. In *9th Congress on electronic structure: principles and applications (ESPA 2014): a conference selection from Theoretical Chemistry Accounts*; Ruiz-Lopez, M. F., Olivares del Valle, F. J., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2016; pp 159–166.

(82) Azofra, L. M.; Alkorta, I.; Scheiner, S. Noncovalent interactions in dimers and trimers of SO_3 and CO. *Theor. Chem. Acc.* **2014**, *133*, 1586.

(83) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(84) Sánchez-Sanz, G.; Alkorta, I.; Elguero, J. A theoretical study of the conformation of 2,2'-bifuran, 2,2'-bithiophene, 2,2'-bitellurophene and mixed derivatives: chalcogen–chalcogen interactions or dipole–dipole effects? *Comput. Theor. Chem.* **2011**, *974*, 37–42.

(85) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(86) Gleiter, R.; Haberhauer, G.; Werz, D. B.; Rominger, F.; Bleiholder, C. From noncovalent chalcogen–chalcogen interactions to supramolecular aggregates: experiments and calculations. *Chem. Rev.* **2018**, *118*, 2010–2041.

(87) Bleiholder, C.; Gleiter, R.; Werz, D. B.; Köppel, H. Theoretical investigations on heteronuclear chalcogen-chalcogen interactions: on the nature of weak bonds between chalcogen centers. *Inorg. Chem.* **2007**, *46*, 2249–2260.

(88) Bleiholder, C.; Werz, D. B.; Köppel, H.; Gleiter, R. Theoretical investigations on chalcogen-chalcogen interactions: what makes these nonbonded interactions bonding? *J. Am. Chem. Soc.* **2006**, *128*, 2666–2674.

(89) Jensen, F. *Introduction to Computational Chemistry*, 3rd ed.; John Wiley & Sons: Chichester, 2017.

(90) Wappett, D. A.; Goerigk, L. A guide to benchmarking enzymatically catalysed reactions: the importance of accurate reference energies and the chemical environment. *ChemRxiv* **2020**, ChemRxiv preprint, https://doi.org/10.26434/chemrxiv.13371098.v1 (accessed on 18 December 2020).

(91) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(92) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(93) Kruse, H.; Goerigk, L.; Grimme, S. *J. Org. Chem.* **2012**, *77*, 10824–10834.

(94) Goerigk, L.; Mehta, N. A trip to the density functional theory zoo: warnings and recommendations for the user. *Aust. J. Chem.* **2019**, *72*, 563–573.

(95) Benz, S.; López-Andarias, J.; Mareda, J.; Sakai, N.; Matile, S. Catalysis with chalcogen bonds. *Angew. Chem., Int. Ed.* **2017**, *56*, 812–815.

(96) Benz, S.; Macchione, M.; Verolet, Q.; Mareda, J.; Sakai, N.; Matile, S. Anion transport with chalcogen bonds. *J. Am. Chem. Soc.* **2016**, *138*, 9093–9096.

(97) Mahmudov, K. T.; Kopylovich, M. N.; da Silva, M. F. C. G.; Pombeiro, A. J. Chalcogen bonding in synthesis, catalysis and design of materials. *Dalton Trans.* **2017**, *46*, 10121–10138.

(98) Beno, B. R.; Yeung, K.-S.; Bartberger, M. D.; Pennington, L. D.; Meanwell, N. A. A survey of the role of noncovalent sulfur interactions in drug design. *J. Med. Chem.* **2015**, *58*, 4383–4438.

(99) Orian, L.; Toppo, S. Organochalcogen peroxidase mimetics as potential drugs: a long story of a promise still unfulfilled. *Free Radical Biol. Med.* **2014**, *66*, 65–74.

(100) P Wolters, L.; Orian, L. Peroxidase activity of organic selenides: mechanistic insights from quantum chemistry. *Curr. Org. Chem.* **2016**, *20*, 189–197.

(101) Tiezza, M. D.; Ribaudo, G.; Orian, L. Organodiselenides: organic catalysis and drug design learning from glutathione peroxidase. *Curr. Org. Chem.* **2019**, *23*, 1381–1402.

(102) Fellowes, T.; Van Koeverden, M. P.; White, J. M. Thermal conversion of a pyridine solvate to a de-solvate facilitated by rearrangement of chalcogen bonds. The solvate and non-solvate structures of N-(2-nitro-4-(3-oxobenzo[d][1,2]selenazol-2(3H)-yl)phenyl)picolinamide. *CrystEngComm* **2020**, *22*, 4023–4029.

(103) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(104) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *J. Chem. Phys.* **1997**, *106*, 1063–1079.

(105) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(106) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Assessment of Gaussian-3 and density-

functional theories on the G3/05 test set of experimental energies. *J. Chem. Phys.* **2005**, *123*, 124107.

(107) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15: a Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.* **2016**, *7*, 5032–5051.

(108) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(109) Kozuch, S.; Martin, J. M. Halogen bonds: benchmarks and theoretical analysis. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.

(110) Kesharwani, M. K.; Manna, D.; Sylvetsky, N.; Martin, J. M. The X40× 10 halogen bonding benchmark revisited: surprising importance of (n–1)d subvalence correlation. *J. Phys. Chem. A* **2018**, *122*, 2184–2197.

(111) Clark, T.; Hennemann, M.; Murray, J. Politzer P (2007) halogen bonding: the σ-hole: proceedings of "modeling interactions in biomolecules II", Prague, September 5th–9th. *J. Mol. Model.* **2005**, *13*, 291–296.

(112) Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding: an electrostatically-driven highly directional noncovalent interaction. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7748–7757.

(113) Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding and other σ-hole interactions: a perspective. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189.

(114) Clark, T.; Murray, J. S.; Politzer, P. Role of polarization in halogen bonds. *Aust. J. Chem.* **2014**, *67*, 451–456.

(115) Clark, T. Halogen bonds and σ-holes. *Faraday Discuss.* **2017**, *203*, 9–27.

(116) Clark, T.; Heßelmann, A. The coulombic σ-hole model describes bonding in CX 3 I? Y- complexes completely. *Phys. Chem. Chem. Phys.* **2018**, *20*, 22849–22855.

(117) Clark, T. Interaction of Radicals with σ-Holes. *J. Phys. Chem. A* **2019**, *123*, 3326–3333.

(118) Bauzá, A.; Alkorta, I.; Frontera, A.; Elguero, J. On the reliability of pure and hybrid DFT methods for the evaluation of halogen, chalcogen, and pnicogen bonds involving anionic and neutral electron donors. *J. Chem. Theory Comput.* **2013**, *9*, 5201–5210.

(119) Kim, Y.; Song, S.; Sim, E.; Burke, K. Halogen and Chalcogen Binding Dominated by Density-Driven Errors. *J. Phys. Chem. Lett.* **2019**, *10*, 295–301.

(120) Sim, E.; Song, S.; Burke, K. Quantifying density errors in DFT. *J. Phys. Chem. Lett.* **2018**, *9*, 6385–6392.

(121) Wasserman, A.; Nafziger, J.; Jiang, K.; Kim, M.-C.; Sim, E.; Burke, K. The importance of being inconsistent. *Annu. Rev. Phys. Chem.* **2017**, *68*, 555–581.

(122) Lee, D.; Furche, F.; Burke, K. Accuracy of electron affinities of atoms in approximate density functional theory. *J. Phys. Chem. Lett.* **2010**, *1*, 2124–2129.

(123) Kim, M.-C.; Sim, E.; Burke, K. Ions in solution: density corrected density functional theory (DC-DFT). *J. Chem. Phys.* **2014**, *140*, 18A528.

(124) Vuckovic, S.; Song, S.; Kozlowski, J.; Sim, E.; Burke, K. Density functional analysis: the theory of density-corrected DFT. *J. Chem. Theory Comput.* **2019**, *15*, 6636–6646.

(125) Lonsdale, D. R.; Goerigk, L. The one-electron self-interaction error in 74 density functional approximations: a case study on hydrogenic mono- and dinuclear systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 15805–15830.

(126) Santra, G.; Martin, J. M. L. What types of chemical problems benefit from density-corrected DFT? A probe using an extensive and chemically diverse test suite. *arXiv* **2020**, preprint ID: arxiv:2010.01519 (accessed on 18 December 2020).

(127) Karton, A.; Goerigk, L. Accurate reaction barrier heights of pericyclic reactions: surprisingly large deviations for the CBS-QB3 composite method and their consequences in DFT benchmark studies. *J. Comput. Chem.* **2015**, *36*, 622–632.

(128) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242–253.

(129) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 91–100.

(130) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: the program system Turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165–169.

(131) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **1995**, *240*, 283 – 290.

(132) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.

(133) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1327.

(134) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.

(135) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.;

Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.* **2019**, *150*, 154122.

(136) DFT-D3 V3.1, S.Grimme, University of Bonn, 2014.

(137) DFT-D4 V2.0, E. Caldeweyher, S. Ehlert and S. Grimme, University of Bonn, 2019.

(138) 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.

(139) 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.

(140) Schwabe, T. Systematic study of the basis set superposition error in core–electron correlation effects. *J. Phys. Chem. A* **2013**, *117*, 2879–2883.

(141) 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*.

(142) 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.

(143) Sure, R.; Grimme, S. Comprehensive benchmark of association (free) energies of realistic host–guest complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801.

(144) Mehta, N.; Abrahams, B. F.; Goerigk, L. Clam-like cyclotricatechylene-based capsules: identifying the roles of protonation state and guests as well as the drivers for stability and (anti-)cooperativity. *Chem. Asian J.* **2020**, *15*, 1301–1314.

(145) 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.

(146) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **2015**, *143*, 054107.

(147) Kruse, H.; Grimme, S. A geometrical correction for the inter-and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **2012**, *136*, 04B613.

(148) Goerigk, L.; Collyer, C. A.; Reimers, J. R. Recommending Hartree–Fock theory with London-dispersion and basis-set-superposition corrections for the optimization or quantum refinement of protein structures. *J. Phys. Chem. B* **2014**, *118*, 14612–14626.

(149) Goerigk, L.; Reimers, J. R. Efficient methods for the quantum chemical treatment of protein structures: the effects of london-dispersion and basis-set incompleteness on peptide and water-cluster geometries. *J. Chem. Theory Comput.* **2013**, *9*, 3240–3251.

(150) Zhao, Y.; Truhlar, D. G. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.

(151) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*.

(152) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(153) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110*, 5029–5036.

(154) Bortoli, M.; Ahmad, S. M.; Hamlin, T. A.; Bickelhaupt, F. M.; Orian, L. Nature and strength of chalcogen–π bonds. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27592–27599.

(155) Fellowes, T.; White, J. M. New insights into chalcogen bonding provided by co-crystal structures of benzisoselenazolinone derivatives and nitrogen bases. *CrystEngComm* **2019**, *21*, 1539–1542.

(156) Varadwaj, P. R.; Varadwaj, A.; Marques, H. M.; MacDougall, P. J. The chalcogen bond: can it be formed by oxygen? *Phys. Chem. Chem. Phys.* **2019**, *21*, 19969–19986.

(157) Varadwaj, P. R. Does oxygen feature chalcogen bonding? *Molecules* **2019**, *24*, 3166.

(158) Bauzá, A.; Mooibroek, T. J.; Frontera, A. The Bright Future of Unconventional σ/π-Hole Interactions. *ChemPhysChem* **2015**, *16*, 2496–2517.

(159) Mukherjee, A. J.; Zade, S. S.; Singh, H. B.; Sunoj, R. B. Organoselenium chemistry: role of intramolecular interactions. *Chem. Rev.* **2010**, *110*, 4357–4416.

(160) Huang, H.; Yang, L.; Facchetti, A.; Marks, T. J. Organic and polymeric semiconductors enhanced by noncovalent conformational locks. *Chem. Rev.* **2017**, *117*, 10291–10318.

(161) Benz, S.; Mareda, J.; Besnard, C.; Sakai, N.; Matile, S. Catalysis with chalcogen bonds: neutral benzodiselenazole scaffolds with high-precision selenium donors of variable strength. *Chem. Rev.* **2017**, *8*, 8164–8169.

(162) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(163) Fliegl, H.; Klopper, W.; Hättig, C. Coupled-cluster theory with simplified linear-r 12 corrections: the CCSD(R12) model. *J. Chem. Phys.* **2005**, *122*, 084107.

(164) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(165) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ quality coupled-cluster correlation energies with triple-ζ basis sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921–1930.

(166) Werner, H.-J.; Knizia, G.; Manby, F. R. Explicitly correlated coupled cluster methods with pair-specific geminals. *Mol. Phys.* **2011**, *109*, 407–417.

(167) Hampel, C.; Werner, H.-J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(168) Schütz, M. Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T). *J. Chem. Phys.* **2000**, *113*, 9986–10001.

(169) Schütz, M.; Werner, H.-J. Local perturbative triples correction (T) with linear cost scaling. *Chem. Phys. Lett.* **2000**, *318*, 370–378.

(170) Schütz, M.; Werner, H.-J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *J. Chem. Phys.* **2001**, *114*, 661–681.

(171) Adler, T. B.; Werner, H.-J. Local explicitly correlated coupled-cluster methods: efficient removal of the basis set incompleteness and domain errors. *J. Chem. Phys.* **2009**, *130*, 241101.

(172) Adler, T. B.; Werner, H.-J. An explicitly correlated local coupled cluster method for calculations of large molecules close to the basis set limit. *J. Chem. Phys.* **2011**, *135*, 144117.

(173) Ma, Q.; Werner, H.-J. Explicitly correlated local coupled-cluster methods using pair natural orbitals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1371.

(174) Karton, A.; Martin, J. M. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, *136*, 124114.

(175) Karton, A. A computational chemist's guide to accurate thermochemistry for organic molecules. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2016**, *6*, 292–310.

(176) Knizia, G.; Werner, H.-J. Explicitly correlated RMP2 for high-spin open-shell reference states. *J. Chem. Phys.* **2008**, *128*, 154103.

(177) Noga, J.; Kedžuch, S.; Šimunek, J. Second order explicitly correlated R12 theory revisited: a second quantization framework for treatment of the operators' partitionings. *J. Chem. Phys.* **2007**, *127*, 034106.

(178) Valeev, E. F. Improving on the resolution of the identity in linear R12 ab initio theories. *Chem. Phys. Lett.* **2004**, *395*, 190–195.

(179) Yousaf, K. E.; Peterson, K. A. Optimized auxiliary basis sets for explicitly correlated methods. *J. Chem. Phys.* **2008**, *129*, 184108.

(180) Goerigk, L.; Sharma, R. The INV24 test set: how well do quantum-chemical methods describe inversion and racemization barriers? *Can. J. Chem.* **2016**, *94*, 1133–1143.

(181) Wappett, D. A.; Goerigk, L. Toward a quantum-chemical benchmark set for enzymatically catalyzed reactions: important steps and insights. *J. Phys. Chem. A* **2019**, *123*, 7057–7074.

(182) Karton, A.; Martin, J. M. Comment on:"Estimating the Hartree–Fock limit from finite basis set calculations"[Jensen F (2005) Theor Chem Acc 113: 267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.

(183) 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. *Chemical Physics Letters* **1998**, *286*, 243–252.

(184) Dunning, Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(185) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(186) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*.

(187) Peterson, K. A.; Dunning, T. H. Accurate correlation consistent basis sets for molecular core-valence correlation effects: the second row atoms Al-Ar, and the first row atoms B-Ne revisited. *J. Chem. Phys.* **2002**, *117*.

(188) Zheng, J.; Xu, X.; Truhlar, D. G. Minimally augmented Karlsruhe basis sets. *Theor. Chem. Acc.* **2011**, *128*, 295–305.

(189) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M.; Neese, F. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

(190) 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* **2019**, *124*, 90–100.

(191) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. The magnitude of the CH/π interaction between benzene and some model hydrocarbons. *J. Am. Chem. Soc.* **2000**, *122*, 3746–3753.

(192) 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.

(193) 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.

(194) 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.

(195) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(196) 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.

(197) Zhang, Y.; Yang, W. Comment on "Generalized gradient approximation made simple". *Phys. Rev. Lett.* **1998**, *80*, 890.

(198) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(199) 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.

(200) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157*, 200–206.

(201) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.

(202) Perdew, J. P. Erratum: density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *34*, 7406.

(203) Handy, N. C.; Cohen, A. J. Left-right correlation energy. *Mol. Phys.* **2001**, *99*, 403–412.

(204) Xu, X.; Goddard, W. A. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proc. Natl. Acad. Sci. U. S. A.,* **2004**, *101*, 2673–2677.

(205) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: the search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.

(206) Najibi, A.; Goerigk, L. DFT-D4 counterparts of leading meta-generalized-gradient approximation and hybrid density functionals for energetics and geometries. *J. Comput. Chem.* **2020**, *41*, 2562–2572.

(207) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(208) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(209) Brandenburg, J.; Bates, J.; Sun, J.; Perdew, J. Benchmark tests of a strongly constrained semilocal functional with a long-range dispersion correction. *Phys. Rev. B* **2016**, *94*, 115144.

(210) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(211) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(212) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(213) Zhao, Y.; Truhlar, D. G. Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: the MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.

(214) Mardirossian, N.; Head-Gordon, M. ω B97M-V: a combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(215) Mardirossian, N.; Head-Gordon, M. ωB97X-V: a 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(216) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(217) Yu, F. Double-hybrid density functionals free of dispersion and counterpoise corrections for non-covalent interactions. *J. Phys. Chem. A* **2014**, *118*, 3175–3182.

(218) Mehta, N.; Casanova-Páez, M.; Goerigk, L. Semi-empirical or non-empirical double-hybrid density functionals: which are more robust? *Phys. Chem. Chem. Phys.* **2018**, *20*, 23175–23194.

(219) Karton, A.; Tarnopolsky, A.; Lamére, J.-F.; Schatz, G. C.; Martin, J. M. Highly accurate first-principles benchmark data sets for the parametrization and validation of density functional and other approximate methods. Derivation of a robust, generally applicable, double-hybrid functional for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.

(220) Kozuch, S.; D.Gruzman,; Martin, J. M. L. DSD-BLYP: a General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction. *J. Phys. Chem. C* **2010**, *114*, 20801–20808.

(221) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.

(222) Alipour, M. Seeking for spin-opposite-scaled double-hybrid models free of fitted parameters. *J. Phys. Chem. A* **2016**, *120*, 3726–3730.

(223) Chai, J.-D.; Head-Gordon, M. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.* **2009**, *131*, 174105.

(224) Santra, G.; Sylvetsky, N.; Martin, J. M. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129–5143.

(225) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network. *Phys. Rev. B* **2015**, *92*, 045131.

(226) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: the simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

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

(228) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: defining and testing the ω B97M (2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.

(229) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; III, H. L. W.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; Jr., R. A. D.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; III, W. A. G.; Gordon, M. S.;

Hehre, W. J.; Klamt, A.; III, H. F. S.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Voorhis, T. V.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(230) Hujo, W.; Grimme, S. Comparison of the performance of dispersion-corrected density functional theory for weak hydrogen bonds. *Phys. Chem. Chem. Phys.* **2011**, *13*, 13942–13950.

(231) L. Goerigk, Density Functional Theory approximations: development and evaluation for electronic ground and excited states, PhD thesis, Westfälische Wilhelms-Universität Münster, 2011.

(232) 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.

(233) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(234) Grimme, S.; Huenerbein, R.; Ehrlich, S. On the Importance of the Dispersion Energy for the Thermodynamic Stability of Molecules. *ChemPhysChem* **2011**, *12*, 1258–1261.

(235) Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.

(236) Hujo, W.; Grimme, S. Performance of non-local and atom-pairwise dispersion corrections to DFT for structural parameters of molecules with noncovalent interactions. *J. Chem. Theory Comput.* **2013**, *9*, 308–315.

(237) Rösel, S.; Quanz, H.; Logemann, C.; Becker, J.; Mossou, E.; Canadillas-Delgado, L.; Caldeweyher, E.; Grimme, S.; Schreiner, P. R. London dispersion enables the shortest intermolecular hydrocarbon H··· H contact. *J. Am. Chem. Soc.* **2017**, *139*, 7428–7431.

(238) Fokin, A. A.; Zhuk, T. S.; Blomeyer, S.; Pérez, C.; Chernish, L. V.; Pashenko, A. E.; Antony, J.; Vishnevskiy, Y. V.; Berger, R. J. F.; Grimme, S.; Logemann, C.; Schnell, M.; Mitzel, N. W.; Schreiner, P. R. Intramolecular London Dispersion Interaction Effects on Gas-Phase and Solid-State Structures of Diamondoid Dimers. *J. Am. Chem. Soc.* **2017**, *139*, 16696–16707.

(239) Kraus, P.; Frank, I. Density functional theory for microwave spectroscopy of noncovalent complexes: a benchmark study. *J. Phys. Chem. A* **2018**, *122*, 4894–4901.

(240) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking density functional methods against the S66 and S66x8 datasets for non-covalent interactions. *ChemPhysChem* **2011**, *12*, 3421–3433.

(241) Goerigk, L. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; Otero de la Roza, A., DiLabio, G. A., Eds.; Elsevier: Amsterdam, 2017; pp 195 – 219.

(242) Otero-de-la Roza, A.; LeBlanc, L. M.; Johnson, E. R. Dispersion XDM with hybrid functionals: delocalization error and halogen bonding in molecular crystals. *J. Chem. Theory Comput.* **2019**, *15*, 4933–4944.

Graphical TOC Entry