Non-Covalent Interactions Atlas Benchmark Data Sets 2: Hydrogen Bonding in an Extended Chemical Space

Jan Řezáč*

Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 166 10 Prague, Czech Republic

E-mail: rezac@uochb.cas.cz

June 29, 2020

Abstract

The Non-Covalent Interactions Atlas (www.nciatlas.org) aims to provide a new generation of benchmark data sets for non-covalent interactions. The HB300SPX data set presented here extends the coverage of hydrogen bonds to phosphorus, sulfur and halogens up to iodine. It is again complemented by a set of dissociation curves, HB300SPX×10. The new data make it possible to analyze the transferability of the parametrization of e.g. dispersion corrections for DFT from simple organic molecules to a broader chemical space. The HB300SPX×10 has also been used for the extension of the parametrization of hydrogen-bonding corrections in the semiempirical PM6-D3H4X and DFTB3-D3H5 methods to additional elements.

1 Introduction

The Non-Covalent Interactions Atlas (NCIA, www.nciatlas.org) is a collection of benchmark data sets of interaction energies intended for the validation and parametrization of approximate computational methods. The first two data sets already published, HB375 and IHB100, cover hydrogen bonds in organic molecules (comprised of H, C, N and O atoms).[1] The next data set in the series presented here, HB300SPX, extends the coverage of hydrogen bonds to sulfur, phosphorus and halogens (F to I). It features 300 model complexes where these elements act as both hydrogen-bond donors and acceptors. For many of these combinations, no benchmark data were previously available. Analogously to the other NCIA data sets, ten-point dissociation curves have been constructed for each systems, forming the HB300SPX×10 data set.

The motivation to develop this data set was twofold: First and foremost, such data are necessary for the development of empirically parametrized computational methods ranging from forcefields to semiempirical quantum-mechanical (SQM) methods and machine learning (ML) approaches. The lack of data beyond the few most common organic elements often translates into the limited applicability of these methods to more general problems. Second, even methods that rely solely on global parameters (such as various dispersion corrections for DFT) may be affected by the limited diversity of systems used for their parametrization. New data covering a wider chemical space can be used to test their robustness and improve their development in the future.

The HB300SPX data set has been constructed using the same methods as the HB375 set to ensure their compatibility. The geometries have been optimized at the DFT-D3 level and verified to be true minima. The benchmark interaction energies have been computed using the composite CCSD(T)/CBS scheme with MP2/CBS extrapolated from aug-cc-pVTZ and aug-cc-pV5Z basis sets and a CCSD(T) correction calculated in heavy-aug-cc-pVTZ basis. This setup balances high accuracy and a cost still applicable to large data sets.[2–4] Bromine and iodine are treated slightly differently in order to account for effects specific to heavier

elements.[5] The complete protocol is outlined below; additional information can be found in the paper introducing the NCIA project.[1]

In the NCIA project, special emphasis is placed on publishing the data sets openly in a form that simplifies their reuse. The benchmark interaction energies as well as the results of all the other calculations used in this study are available in several formats in the Supporting Information of this paper as well as at the www.nciatlas.org website. The data set is annotated with metadata enabling the selection of specific systems or subsets defined here. The calculations and analysis of the results on the data set can readily be automated using the Cuby framework.[6,7]

Besides introducing the data set, this paper also focuses on its applications to benchmarking dispersion-corrected DFT methods and to the parametrization of hydrogen-bonding corrections for semiempirical methods. DFT-D and other related means of introducing London dispersion to DFT rely on damping functions that control the transition between the short-ranged correlation included in the DFT functional and the dispersion correction itself. The transferability of its parametrization (done solely or predominantly on simple organic molecules) to other elements is a good test of the robustness of the method, and hydrogen bonds make it possible to explore these effects at short range, where they are most important. The new data enabled the identification of an issue in the description of the interactions of iodine in the popular D3 and D4 dispersion corrections.[8,9]

What is more specific to hydrogen bonds is their description in semiempirical QM methods. The approximations involved at this level lead to a serious underestimation of their interaction energies. This issue is well documented and various corrections had been proposed.[10–12] Their parametrization was, however, limited to hydrogen bonds of nitrogen, oxygen and possibly sulfur because no reference data were available for other elements. With the new data, we revisit the hydrogen-bonding corrections in the PM6-D3H4X[12–14] and DFTB3-D3H5[15] methods and extend their parametrization to more elements. This is an important part of our efforts in the application of SQM methods in computer-aided drug

design, where the studied compounds often include heteroatoms covered by this data set and their interactions have to be described as accurately as possible.

The present data sets focused on hydrogen bonds will soon be complemented by compatible sets covering other types of interactions. Two larger data sets covering London dispersion and repulsive contacts in the same chemical space are now under development. Another data set of sigma-hole interactions will follow closely. Together, these data sets should provide accurate benchmarks for all important non-covalent interactions in this chemical space.

2 Methods

2.1 Data Set Construction

First, a preliminary set of complexes featuring possible H-bonds involving the studied elements was built automatically by connecting monomers, in which hydrogen-bond donor and acceptor sites were labeled. The monomers were chosen as the simplest molecules featuring different valence states and chemical environments of the atoms. The studied elements (P, S, F, Cl, Br and I) were used as both H-bond donors and acceptors, and additional H-bond donors with CH, NH and OH groups were added to provide complete coverage when combined with the HB375 set. The monomers present in the data set are: acetamide, acetic acid, acetone, ammonia, bromic acid, molecular bromine (Br_2), chloric acid, molecular chlorine (Cl_2), difluorophosphine, dimethyldisulfide, dimethylether, dimethylsulfide, dimethylsulfone, dimethylsulfoxide, disulfur dichloride, ethyne, molecular fluorine (F_2), hydrogen bromide, hydrogen chloride, hydrogen fluoride, hydrogen iodide, molecular iodine (I_2), methane, methanethiol, methanol, methylamine, methyl azide, methyl bromide, methyl chloride, methyl cyanate, methyl fluoride, methyl iodide, nitromethane, N-methylacetamide, phosphorus sesquisulfide (P_4S_3), phosphaalkene $(CH_3)_2C{=}PCH_3$, phosphaalkyne $CH_3C{\equiv}P$, phosphine, phosphoric acid, phosphorine, molecular phosphorus (P_4), phosphorus trichloride, phosphoryl chloride, pyridine, sulfane, sulfur dioxide, tetrafluoromethane, tetrahydrothio-

phene, thioacetone, trimethylamine, trimethylphosphate, trimethylphosphine, trimethylphosphite and water. The initial set of ca 700 dimers contained all combinations of the elements on both sides of the H-bonds.

After the optimization of the structures (described below), only geometries with hydrogen bonds were retained. The presence of the H-bond was defined as a contact between a hydrogen bound to element X and another atom Y shorter than the sum of their van der Waals (vdW) radii with the XHY angle larger than 120°. These criteria are weaker than in the HB375 data set, where contact shorter than 90% of the van der Waals distance was required, but the less electronegative elements studied here do not form so short H-bonds. If there are more such contacts, the one with the shortest distance relative to the sum of vdW radii is used to classify the system. Next, all (CH,NH,OH)–(N,O) H-bonds were removed as these are covered by the HB375 data set. Finally, since this selection contained too many H-bonds with an oxygen acceptor, the number of systems for each XH–O combination was limited to ten randomly selected species. This procedure resulted in a set of 300 H-bonded complexes, which form the HB300SPX data set. They are listed in the Supporting Information, Table S1.

The combinations of elements covered by the set and the numbers of systems representing each combination are summarized in Table 1. The coverage is not homogeneous because some of the heavier elements are not very likely to form hydrogen bonds, and few combinations are missing completely because the molecules prefer different interaction motifs. Namely, phosphorus is such a weak H-bond donor that there are no PH–S H-bonds, and iodine is a weak acceptor, as a result of which no NH–I or SH–I contacts have been found. In most cases, iodine clearly prefers the formation of halogen bonds. Some other systems may not be prototypical hydrogen bonds even if they match the geometric criteria; this is analyzed and discussed below.

Since the properties of the H-bond are the most affected by the acceptor atoms, the data set is organized into larger groups labeled XH–N, XH–O, XH–P, XH–S, XH–F, XH–Cl, XH–

Table 1: Counts of the combinations of elements forming hydrogen bonds in the HB300SPX data set. Hydrogen-donor groups XH in rows, acceptor atoms Y in columns.

	N	O	S	P	F	Cl	Br	I
CH	—	—	7	5	12	3	1	1
NH	—	—	7	8	5	1	1	
OH	—	—	7	7	9	9	5	7
SH	10	10	7	6	1	1	1	
PH	4	1		1	1	1	1	1
FH	7	10	7	6	5	4	2	3
ClH	6	10	7	7	3	5	2	3
BrH	4	10	6	7	2	4	2	2

Br and XH–I. Analogously to the other NCIA data sets, these groups define the numbering of the systems in the set (group number, dot, number of a system in the group) and are used in the analysis of the results. The information on the groups is summarized in Table 2. At this level of resolution, the sizes of the largest groups are well balanced, and even the least populated group with 17 entries provides a sufficient number of systems for most applications.

Table 2: The composition of the HB300SPX data set: Groups by hydrogen-bond acceptor element, their size, and the average interaction energy (kcal/mol) in each group.

#	Group	Size	$\langle \Delta E^{int} \rangle$
1	XH–N	34	-4.9
2	XH–O	51	-4.9
3	XH–P	52	-2.3
4	XH–S	54	-3.2
5	XH–F	41	-2.5
6	XH–Cl	32	-1.6
7	XH–Br	17	-1.8
8	XH–I	19	-1.5

Analogously to the previous NCIA data sets, smaller predefined subsets of the HB300SPX dataset were generated by clustering analysis.[1,16] The desired number of clusters (20, 50, 100 and 200) was formed, and the representative systems from each cluster comprise the respective subset. The data set is clustered by the errors of the methods used or tested in this paper (48 correlated, DFT and SQM methods); the clusters thus gather systems in which these

methods perform similarly. Therefore, the cluster representatives define diverse, information-rich subsets optimized for method testing. A complete-linkage clustering algorithm is used, in which the similarity of each two systems is evaluated as a Pearson correlation coefficient between their error vectors. A cluster representative is then selected as the system most similar to all the other members of the cluster.

It has been verified that the subsets generated by the clustering cover all the chemically-defined groups of the data set and that the representation of these groups in the subsets is proportional to the size of the subset. This has indicated that the similarity measure used in the clustering, which is based on the results of calculations, also reflects well the chemical diversity of the data set. These subsets are listed in the SI in table S2, and the assignment of the systems to them is also indicated in the data files provided.

2.2 Geometry Preparation

The geometries of the systems were prepared and optimized using a protocol consistent with the one used to build previous NCIA data sets.[1] The complexes are prepared automatically from monomer structures in which possible H-bond donor and acceptor atoms are labelled. Here, a larger number of systems (about 700) was built and optimized, out of which the final data set of 300 complexes was selected as described above. The preparation of the systems consists of multiple rounds of geometry optimizations and adjustments until true minima are obtained (which is verified by vibrational analysis). In the first rounds, small random adjustments are made in the intermolecular coordinates. If this is not sufficient to locate a reliable minimum, a short molecular dynamics is applied to anneal the geometry (cooling the system from 20 to 0 K over 500 fs). Unlike in the HB375 data set, where longer annealing was first applied to all the systems, here we use it only to resolve problematic cases; otherwise, we prefer local minima close to the initial structure as these are more likely to conserve the intended interaction motif. The final geometries are guaranteed to have the root mean square (RMS) of the gradient lower than 0.01 kcal/mol/Å and real vibrational

frequencies.

All the calculations in this protocol are performed at the B3LYP-D3(BJ)/def2-QZVP level.[17] The use of DFT is necessary for performance reasons; even at this level, the preparation of the geometries consumes as many resources as the final benchmark calculations. However, a large basis set and a very fine DFT grid are used to ensure the best possible results. A hybrid functional is needed to prevent artifacts due to delocalization error. In the previous paper, it was found that specifically B3LYP-D3(BJ) does not perform optimally around the equilibrium distances of hydrogen bonds (in terms of interaction energies), but other studies[18] as well as our own tests suggest that the geometries optimized at this level are close to high-level benchmarks. We thus keep the same setup for consistency with the other data sets built at this level.

To check the geometries further, the MP2/def2-QZVP gradient was calculated on all the structures and its intermolecular component was separated. The RMS of the Cartesian gradient averaged over the data set is 0.15 kcal/mol/Å, which indicates that DFT-D3 and MP2 geometries would be similar. The difference is, however, slightly larger than in the HB375 set, where the average RMS of the gradient is 0.09 kcal/mol/Å, which suggests that the present systems are more difficult to handle by DFT.

The set of ten-point dissociation curves, HB300SPX×10, is constructed from the equilibrium geometries using the protocol described in detail in the previous paper.[1] In short, the distance of the closest intermolecular contact (in this case the length of the H-bond) is scaled by a factor of 0.8, 0.85, 0.9, 0.95, 1.0, 1.05, 1.1, 1.25, 1.50 and 2.0, moving one of the molecules along the axis of the H-bond.

After the dissociation curves were computed at the benchmark level, the equilibrium distances interpolated from these curves were compared to the optimized geometries. In seven systems (3.50, 4.54, 5.18, 5.27, 6.14, 7.07 and 8.09), the error was larger than 5%, which was considered unacceptably large (as this is also the spacing of the points around the minimum in the dissociation curves). These systems were reoptimized at the MP2/def2-

QZVP level, which fixed this issue. In two of these systems (5.27 and 6.14), the geometry no longer matches our criteria for a H-bond, but the original XH–Y remains the closest contact between the molecules; therefore, they were left in their respective groups. Another attempt to reoptimize these structures using a double-hybrid DFT functional, DSD-BLYP-D3, did not improve the errors in the equilibrium distances well enough.

All the calculations in this protocol were carried out using Orca 4.2;[19] the RI approximation with the auxiliary basis sets corresponding to the AO basis was used.

2.3 Benchmark CCSD(T)/CBS Calculations

The benchmark interaction energies are computed using the same scheme as in the HB375 data set; all details can be found in the respective paper.[1] Important points at the dissociation curve (the equilibrium geometry and the closest point) are computed at the "gold level" with MP2/CBS extrapolated[20] from aug-cc-pVQZ and aug-cc-pV5Z basis sets, and the CCSD(T) correction is computed in the heavy-aug-cc-pVTZ basis.[21] It is now widely agreed that a true "gold-standard" CCSD(T)/CBS benchmark requires the CCSD(T) correction computed in the triple-zeta basis,[2,3,22] and this study shows that the basis sets used to obtain the MP2/CBS term should also be as large as possible. The remaining points were computed at the "silver level" with MP2/CBS extrapolated from aug-cc-pVTZ and aug-cc-pVQZ basis sets and CCSD(T) correction calculated in the aug-cc-pVDZ basis set. The silver-level calculations are then rescaled to the gold level using the procedure described earlier; the loss of accuracy in comparison with a full gold-level calculation has been found to be negligible (in the order of 0.001 kcal/mol).[1]

The only difference is in the treatment of bromine and iodine. Here, effective core potentials (ECPs) are used to account for the relativistic effects, and the corresponding variant of the basis set is used.[23] Moreover, it has been shown that the sub-valence d orbitals on these atoms significantly contribute to dispersion interactions.[5] We have thus excluded these orbitals from the frozen core, which calls for the use of basis sets from the aug-cc-pwCVXZ-PP

family.[24]

All the interaction energies were calculated using counterpoise correction.[25] Systems containing lighter atoms were calculated in Psi4 using density fitting with the default setup.[26] Systems containing Br and I were calculated in Turbomole 7.3 using RI approximation in the correlation energy calculation.[27] It has been verified that both of these programs with this setup produce practically identical results.

The results obtained along with the benchmark calculations were used to build interaction energies at a lower level, such as various flavors of MP2. Additional MP3 calculations in Turbomole were performed to obtain MP2.5 interaction energies.[28]

2.4 SAPT0 Interaction Energy Decomposition

A qualitative interaction energy decomposition has been performed using the approximate symmetry-adapted perturbation theory, SAPT0.[29,30] These calculations used the recommended jun-cc-pVDZ basis,[31] which offers favorable error compensation. This basis set is not available for iodine; in iodine-containing molecules, def2-TZVP and the accompanying ECP were used for the iodine atoms. The SAPT0 calculations were carried out in the Psi4 program.[26]

2.5 Density Functional Theory Calculations

The new data set is used for testing multiple DFT functionals with several different approaches to the treatment of non-covalent interactions. This analysis is more extensive than the tests published in the paper on the HB375 data set.[1] This data set was newly computed with additional functionals to allow comparison with the HB300SPX set (in both cases, the complete dissociation curves were computed).

The first set of DFT calculations uses the D3 dispersion correction.[8] The functionals BLYP, BP, PBE, B97D, TPSS, SCAN, B3LYP, BHLYP, PBE0 and TPSSH are used with Becke-Johnson (BJ) damping.[17] Additionally, BLYP, B3LYP, TPSS and TPSSH were also

used with the optimized power (OP) damping,[32] and B3LYP and M062X functionals with the original zero damping. The results of M062X without a dispersion correction were also evaluated as this functional recovers some dispersion effects by other means. The D4 correction,[9] a successor of D3 using charge-dependent C_6 coefficients, was applied to BLYP and B3LYP.

Next, a series of DFT functionals with a non-local van der Waals correction[33] was examined: BLYP-NL and B3LYP-NL apply this correction to standard DFT functionals,[34] while B97M-V and ωB97X-V were parametrized with this correction included.[35,36] Their variants ωB97X-D3 and ωB97X-D3BJ were added for comparison.[37,38] Finally, DSD-BLYP and DSD-BLYP-D3 represent recent broadly parametrized double-hybrid functionals.[39]

All the DFT calculations have been carried out in the def2-QZVP basis set.[40] Firstly, it is large enough to minimize the errors caused by the basis set size below the magnitude of the errors specific to DFT functionals and dispersion corrections, which is a prerequisite for making a valid comparison of these methods. Second, it is the basis set for which the D3 and D4 dispersion corrections were parametrized, so that their performance should be optimal in this basis. All the DFT calculations were run in Orca 4.2[19] using RI approximation. The D3 correction was added to the common functionals using the Cuby framework,[6,7] because it also implements the OP damping. The D4 correction was computed using the standalone program provided by the authors.[41]

2.6 Semiempirical QM Calculations.

The HB300SPX data set is used to test several semiempirical QM methods targeted at non-covalent interactions. PM6 method[13] with the D3H4 corrections[12] for London dispersion and hydrogen bonding and PM7[42] (which already includes similar corrections) represent the classical SQM methods. These calculations were performed in MOPAC 2016.[43]

Next, there are several variants of the third-order self-consistent charge density-functional tight binding method (DFTB3),[44] all using the 3OB parameter set.[45,46] This method and its

version with D3 dispersion correction (DFTB3-D3)[47] use the XH-damping (with the recommended exponent of 4.0) as a hydrogen-bonding correction. In DFTB3-D3H4, it is replaced by a standalone hydrogen-bonding correction[12] with updated parameters,[48] and DFTB3-D3H5 integrates a different hydrogen-bonding correction into the self-consistent charge procedure.[15] All DFTB3 calculations were carried out in the DFTB+ program, which now implements all of these corrections as well.[49,50]

Finally, the GFN2-xTB method[51] is a recent empirical tight-binding model covering a wide range of elements and designed with non-covalent interactions as one of the goals of its parametrization. These calculations were performed in the standalone code provided by the authors of the method.[52]

3 Data Availability

The HB300SPX×10 benchmark data set itself as well as other results presented in this paper are available in the Supporting Information and at the NCIA website www.nciatlas.org. At the website, it is also possible to browse all the systems. The data set of equilibrium geometries labeled here as HB300SPX is not published separately, as it is a subset of the HB300SPX×10 set.

The data set consists of geometries in .xyz format (with additional information on the monomers etc.), and a table of benchmark interaction energies and metadata. The metadata describe the classification of the systems and their assignment to predefined groups and subsets. The results of all the tested methods, and the interaction energy components used to construct the CCSD(T)/CBS benchmark are available in a separate table.

Additionally, all this information is also provided in the form of a machine-readable YAML data file used to automate the calculations in the Cuby framework. The data set will also be bundled with a future version of the framework. Cuby can be used to automate calculations on the data set and process the results. YAML is a structured data format[53]

that is human-readable and accessible from all common programming languages, so these data can easily be used outside of Cuby as well.

4 Results and Discussion

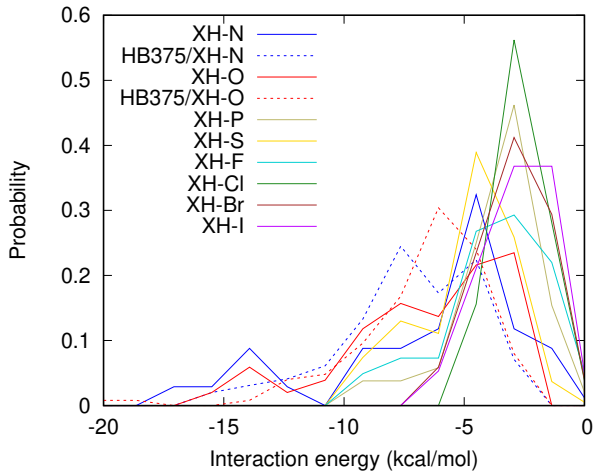
4.1 The Nature of the Interactions in the HB300SPX Data Set

All the complexes in the data set feature a geometric arrangement of a hydrogen bond. Nevertheless, it is useful to discuss also other properties of these interactions before using the data set. Here, the interaction energies in the data set and their decomposition are analyzed.

The average interaction energies in each group of the HB300SPX data set are listed in Table 2 and their distributions are plotted in Figure 1. Two subsets of the HB375 set corresponding to XH–N and XH–O groups (X being N or O) are added for comparison. Apparently, the hydrogen bonds in the HB300SPX set are weaker than the hydrogen bonds between the nitrogen and oxygen from the HB375 set. This also applies to the XH–N and XH–O groups, where the interactions are strongest because here the X can be a less electronegative element than just O and N in HB375. They are followed by the XH–S group, where both the electronegativity and polarizability of sulfur contribute to the strength of the interaction. The next groups are XH–F and XH–P, where only one of these contributions is strong. Finally, the weakest interactions are formed by the halogens from chlorine to iodine.

To verify whether these interactions can be called hydrogen bonds, the nature of the interaction has to be explored in more detail. A valuable insight can be provided by interaction energy decomposition. Here, SAPT0 is used to compute the ratio of first-order electrostatics, induction and dispersion to their sum (R^{ele}, R^{ind}, R^{disp}). In this way, we focus on the attractive terms, whereas the first-order exchange covering Pauli repulsion is neglected. The average values of these ratios are listed in Table 3 along with the same quantities computed for the XH–N and XH–O subsets of the HB375 data set (X being limited

Figure 1: Distributions of interaction energies in the groups of the HB300SPX data set and in the comparable subsets of HB375.



to N and O). The plots of these ratios for each system are displayed at the NCIA website. Overall, the low fraction of dispersion (below 26 %, with the exception of the XH–I group with 35 %) indicates that the studied interactions are proper hydrogen bonds mainly driven by electrostatics. Even in the case of iodine, this number would be higher if it was not a hydrogen bond. In the XH–N and XH–O groups, this character is slightly weaker than in the analogous subsets of HB375 because some of the hydrogen-bond donors used here are weaker. The only exception to the expected trends is the lower contribution of electrostatics in the XH–Cl group than in XH–Br despite the opposite trend in their electronegativity (the difference in which is, however, only small). This is a result of the selection of the systems – the XH–Cl complexes used here are weaker than the XH–Br ones (see Table 2).

4.2 Benchmark Calculations and Other Correlated Wavefunction Methods

First, this section discusses the benchmark interaction energies. At the closest and equilibrium geometries, both gold and silver levels are available. The root-mean-square difference between them is 0.072 kcal/mol in the equilibrium and 0.173 kcal/mol in the compressed geometries. This difference can be attributed mainly to the basis set used in the

CCSD(T) correction (aug-cc-pVDZ at the silver level, heavy-aug-cc-pVTZ at the gold level); the MP2/CBS interaction energies used in these two setups (extrapolation from aug-cc-pV[T,Q]Z and aug-cc-pV[Q,5]Z) differ only by 0.022 kcal/mol in the equilibrium and by 0.053 kcal/mol at the short distances. The difference between the gold-level setup and the true complete basis set limit should be smaller than the differences between these two setups.

It is also interesting to compare these results with the same differences evaluated in the HB375 data set (equilibrium geometries only). There, the MP2/CBS values at the silver and gold levels differ only by 0.003 kcal/mol, and practically all the difference (0.053 kcal/mol) comes from the CCSD(T) part of the calculation. This indicates that the calculations in a more extended chemical space are more demanding; it also justifies the use of the quintuple-zeta basis in the MP2/CBS term (which is costly and has hardly brought any improvement in the HB375 data set).

The components of the benchmark calculations can be used to construct several other correlated methods, which may be interesting because of their lower cost. At the MP2 level, we explore MP2 in finite basis sets and at the CBS limit (extrapolated from the same basis sets like in the benchmark), its spin-component scaled variants SCS-MP2 and SCS-MI-MP2,[54,55] and dispersion-corrected MP2D[56] (here, iodine-containing systems were omitted because the method lacks parameters for this element). Additional MP3/heavy-aug-

Table 3: The average ratios of SAPT0 electrostatics, induction and dispersion to their sum (in %) in the groups of the HB300SPX set and analogously selected subsets of HB375.

Set	Group		R^{ele}	R^{ind}	R^{disp}
HB300SPX	1	XH–N	54.3	20.0	25.6
	2	XH–O	57.5	18.9	23.7
	3	XH–P	48.2	27.9	23.9
	4	XH–S	50.8	25.8	23.4
	5	XH–F	59.6	22.2	18.1
	6	XH–Cl	45.1	29.8	25.1
	7	XH–Br	50.1	24.9	25.0
	8	XH–I	38.6	26.8	34.6
HB375	–	XH–N	59.4	19.2	21.4
	–	XH–O	61.3	17.1	21.6

cc-pVTZ calculations were performed in order to build MP2.5 interaction energies.[28] Finally, double-hybrid DFT functionals were added to this comparison because their computational complexity is comparable to MP2. The results are summarized in Figure 2 and provided in a tabular form in the Supporting Information, Table S3. Tables S4 and S5 therein also list the corresponding relative errors and other error measures evaluated in the HB300SPX data set.

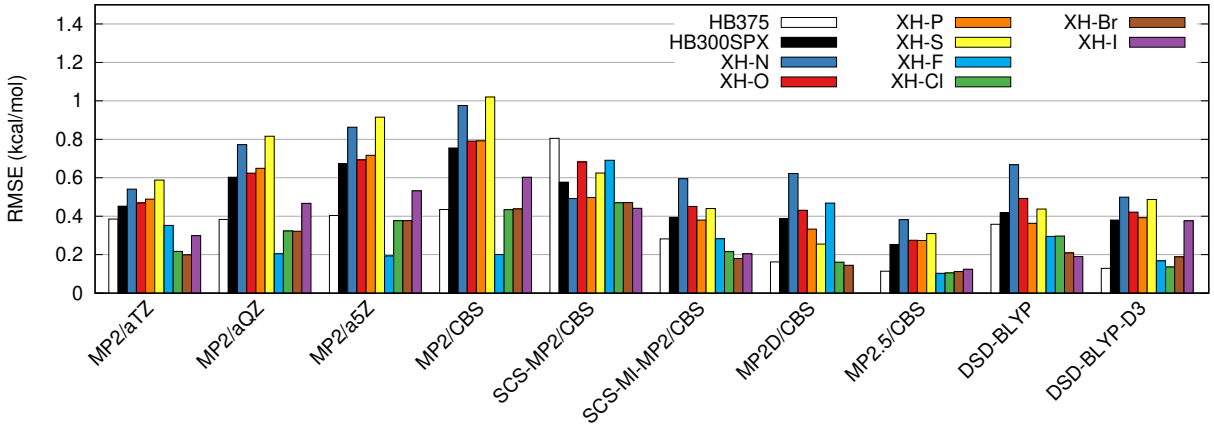
The MP2 calculations in different basis sets and extrapolated to the CBS limit illustrate well the compensation between the method overestimating interaction energies and the basis set size. This error compensation is strongest in the smallest aug-cc-pVTZ basis, and the error increases steadily in the series towards the CBS. This effect is not visible in the HB375 data set, where the MP2 error does not change much with the basis-set size. The common statement that MP2 works well for hydrogen bonds seems to be valid only for H-bonds of second-period elements. In heavier elements, the contribution of London dispersion increases and the limitations of MP2 become more significant.

SCS-MP2 is the only method here that works better in the HB300SPX than in the HB375 data set, but the error is rather large. The remaining MP2-based approaches, SCS-MI-MP2, MP2D and the DSD-BLYP functionals perform similarly, with errors (RMSEs) around 0.4 kcal/mol. This, however, applies only to this data set; in HB375, on the other hand, MP2D and DSD-BLYP-D3 work significantly better than the other two methods. Further improvement can be achieved only at the cost of more complex calculations, with MP2.5 having the RMSE of 0.25 kcal/mol.

4.3 DFT Calculations – General Trends

The results of the DFT methods subject to this study are summarized in Figure 3, where the RMSE in the HB300SPX set is compared to HB375. The actual values as well as RMSE in the individual groups of the data set are listed in Table 4. Analogous tables of relative errors and systematic errors (MSE) are available in the Supporting Information as Tables

Figure 2: The errors of correlated wavefunction calculations and double-hybrid DFT functionals in the HB300SPX data set (black, individual groups in color) and in HB375 (white).



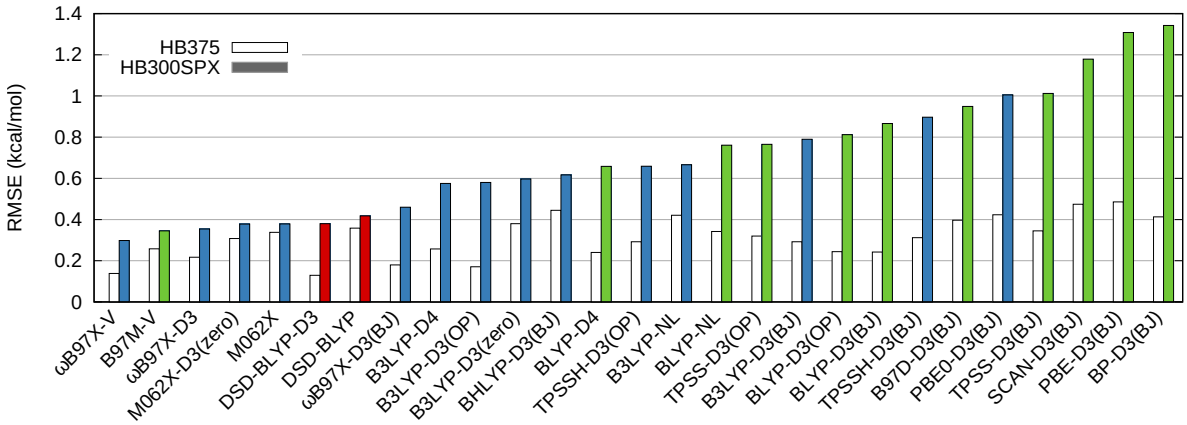
S6 and S7. The plots of all these data are also available as Figures S1, S2 and S3 in the SI.

The best results (with the RMSE of 0.30 kcal/mol) are achieved with the ωB97X-V functional, a range-separated hybrid GGA functional with non-local vdW dispersion correction.[36] Moreover, the errors in all the groups of the data set are very similar (no worse than 0.35 kcal/mol), which indicates that the method is very robust. These results only confirm its excellent performance in other tests. However, this accuracy comes at a price – the calculations are more demanding than common, simpler functionals. At the second and third place, there are two functionals closely related to the top one, B97M-V and ωB97X-D3. The former is a meta-GGA functional using the same treatment of dispersion and parametrized using a similar combinatorial procedure. At the cost of only a small loss of accuracy (which is more pronounced in the HB375 set), it brings much better efficiency because it does not use exact exchange. ωB97X-D3 is then a reparametrization of ωB97X-V with the dispersion term replaced by the D3 correction. It is interesting that this functional does not share the problems found in other D3-corrected methods discussed below. Next comes M06-2X with and without D3 correction. However, their accuracy is significantly worse in the HB375 set, where they are outperformed by many other methods.

Only then come the two double-hybrid functionals, DSD-BLYP-D3 and DSD-BLYP. The former performed exceptionally well in the HB375 set, and its error of 0.4 kcal/mol seems

to be rather large for a functional of this level. Some part of this error clearly comes from the D3 dispersion, which is discussed in detail in the following paragraphs. The common DFT functionals with *a posteriori* empirical dispersion corrections follow with errors starting at 0.6 kcal/mol, which is twice as much as the best functional tested. Here, the errors in the HB300SPX become larger although many of the methods still perform very well in HB375. This trend is so widespread that it is a strong argument for a hypothesis that the parametrization of the dispersion corrections is biased towards simple organic molecules, where it may mask the errors of the DFT itself.

Figure 3: The errors of DFT calculations in the HB300SPX data set (in color) and in HB375 (white). The methods are sorted by increasing RMSE in HB300SPX. The colors indicate pure GGA and meta-GGA functionals (green), hybrid and range-separated hybrid functionals (blue) and double-hybrids (red).



4.4 DFT calculations – issues in the dispersion correction

Next, we analyze the results of different dispersion corrections more closely. Here, B3LYP is used as an example because all the corrections discussed here are available for this functional. The D3 correction has been applied with three different damping functions (BJ, OP and zero),[8,17,32] and the D4[9] and NL[34] corrections are also available. The errors in the individual groups of the HB300SPX data sets are plotted in Figure 4, and the distance-dependence of the RMSE along the HB300SPX×10 curves is shown in Figure 5. The D3(BJ) variant yields

Table 4: The errors of DFT methods (RMSE, in kcal/mol) in the HB375 data set and in the HB300SPX set and its groups. The methods are sorted by increasing RMSE in HB300SPX.

Method	HB375	HB300SPX	XH–N	XH–O	XH–P	XH–S	XH–F	XH–Cl	XH–Br	XH–I
ωB97X-V	0.138	0.298	0.356	0.324	0.346	0.347	0.128	0.211	0.251	0.243
B97M-V	0.258	0.345	0.350	0.369	0.214	0.329	0.288	0.419	0.428	0.468
ωB97X-D3	0.217	0.354	0.470	0.359	0.416	0.320	0.311	0.248	0.304	0.285
M062X-D3(zero)	0.308	0.378	0.488	0.485	0.303	0.293	0.426	0.293	0.302	0.295
M062X	0.338	0.379	0.527	0.466	0.328	0.302	0.362	0.315	0.322	0.298
DSD-BLYP-D3	0.129	0.380	0.499	0.421	0.392	0.487	0.168	0.136	0.189	0.376
DSD-BLYP	0.358	0.418	0.668	0.493	0.363	0.437	0.295	0.297	0.209	0.190
ωB97X-D3(BJ)	0.180	0.460	0.370	0.366	0.632	0.563	0.381	0.247	0.293	0.505
B3LYP-D4	0.252	0.575	0.691	0.558	0.652	0.746	0.346	0.292	0.325	0.535
B3LYP-D3(OP)	0.171	0.580	0.641	0.508	0.689	0.811	0.340	0.274	0.290	0.524
B3LYP-D3(zero)	0.380	0.597	0.768	0.678	0.586	0.760	0.327	0.355	0.381	0.432
BHLYP-D3(BJ)	0.445	0.617	0.734	0.892	0.542	0.566	0.626	0.277	0.289	0.381
BLYP-D4	0.239	0.658	0.618	0.376	0.751	0.948	0.434	0.451	0.591	0.781
TPSSH-D3(OP)	0.292	0.659	0.983	0.484	0.806	0.844	0.408	0.299	0.287	0.341
B3LYP-NL	0.421	0.666	0.911	0.808	0.616	0.764	0.409	0.414	0.428	0.491
BLYP-NL	0.342	0.761	0.867	0.679	0.742	1.003	0.500	0.593	0.731	0.761
TPSS-D3(OP)	0.320	0.765	1.051	0.520	0.931	1.016	0.520	0.393	0.441	0.474
B3LYP-D3(BJ)	0.292	0.790	0.853	0.748	0.920	1.069	0.426	0.453	0.473	0.742
BLYP-D3(OP)	0.244	0.812	0.741	0.507	0.946	1.218	0.482	0.527	0.632	0.856
BLYP-D3(BJ)	0.242	0.866	0.723	0.520	1.051	1.287	0.508	0.574	0.661	0.973
TPSSH-D3(BJ)	0.312	0.897	1.125	0.695	1.117	1.200	0.501	0.400	0.528	0.762
B97D-D3(BJ)	0.397	0.949	0.964	0.486	1.185	1.411	0.722	0.473	0.545	0.809
PBE0-D3(BJ)	0.423	1.005	1.368	0.973	1.177	1.293	0.448	0.428	0.515	0.724
TPSS-D3(BJ)	0.345	1.012	1.207	0.736	1.233	1.370	0.589	0.535	0.733	0.950
SCAN-D3(BJ)	0.474	1.179	1.648	1.319	1.103	1.451	0.640	0.726	0.816	0.863
PBE-D3(BJ)	0.486	1.308	1.610	1.072	1.507	1.766	0.659	0.775	0.983	1.179
BP-D3(BJ)	0.413	1.342	1.386	1.045	1.557	1.946	0.788	0.705	0.899	1.334

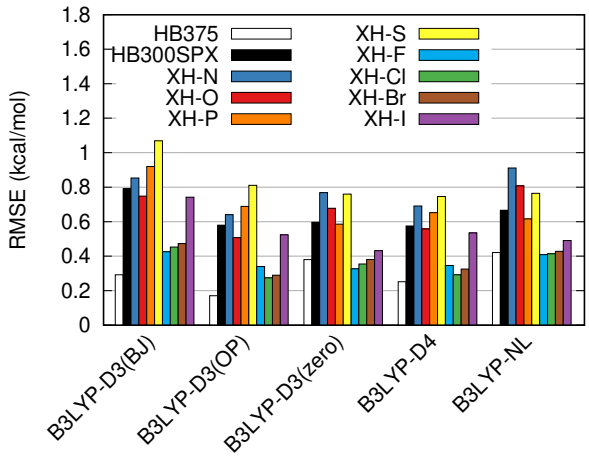
the worst results, which is an issue that has already been reported in the HB375 data set – this parametrization of the damping function exhibits large errors specifically around the equilibrium distance of hydrogen bonds. This issue is removed in D3(OP). The distance-dependence of the errors in D3(zero) is not smooth, which again implies that the damping function does not describe the transition between the DFT correlation and the dispersion correction properly. D4 uses the Becke-Johnson damping, but with different parameters. In HB300SPX, it works well, and the distance-dependence of the errors in HB300SPX×10 is nice and smooth. In B3LYP-NL, the error at short distances grows rather fast, but smoothly.

An important insight can be gained by comparing the errors in the XH–I group to the others. Since the iodine atom is large and its electron density is diffuse, it is difficult to define its radius. Nevertheless, this quantity is a crucial parameter in the damping of the dispersion correction. The D3(BJ), D3(OP) and D4 corrections derive this radius from the dispersion coefficient, while D3(zero) uses an independently calculated value. As the former three approaches yield a significantly larger error in this group, it is likely that this derivation of the damping radius is not reliable for heavier elements. Moreover, the same issue is visible in other methods with the BJ damping function, including the transition from DSD-BLYP to DSD-BLYP-D3, where the overall error decreases but the error in the XH–I group almost doubles. On the other hand, a comparison of the distance-dependence of the errors in the individual groups (available in Supporting Information, Figures S4 and S5) suggests that the good performance of B3LYP-D3(zero) in the case of iodine and bromine may be a result of an error compensation between the value of the damping radius and the shape of the damping function, because the curves are not monotonous. Only in B3LYP-NL, the error in the XH–I group is both small and decays smoothly with the intermolecular distance.

To test the hypothesis on the role of the damping radii rigorously, a modification of the BJ damping has been reparametrized with the radii used in zero damping (which will be labeled as BJ'$_{r0}$) and in the original setup (BJ'). The parametrization is identical to that used for fitting D3 to calculations in a smaller basis set;[57] it minimizes the RMSE weighted

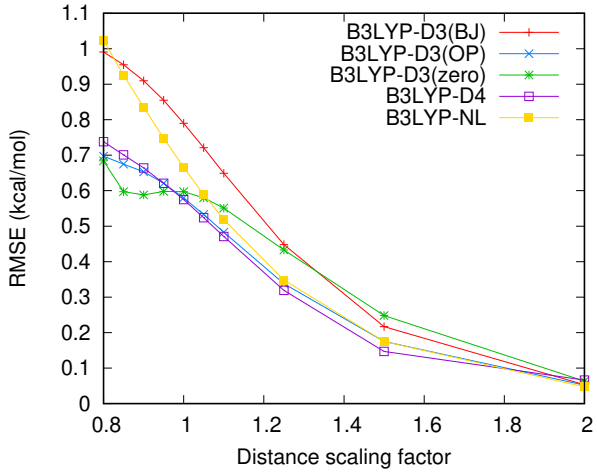
by the average interaction energy in S66x8, X40 and L7 data sets.[58–61] The resulting parameters are provided in the SI, Table S8. The B3LYP-D3(BJ') setup using the original radii performs slightly better than the default parametrization with the RMSE of 0.63 kcal/mol in HB300SPX, but the error in the XH–I group remains rather large, at 0.61 kcal/mol, which is twice as much as in the XH–Br group (0.33 kcal/mol). Using alternative radii in B3LYP-D3(BJ'$_{r0}$) has slightly increased the errors in the training set, but the performance in HB300SPX is better with the RMSE of 0.52 kcal/mol, and at least some improvement is visible in each group of the data set. However, the key difference lies in the XH–I group, where the RMSE drops to 0.33 kcal/mol, becoming closer to the XH–Br group with 0.24 kcal/mol. These results are a strong argument suggesting that the damping radii used in the BJ damping are becoming unreliable for iodine, and this likely applies also to other heavy elements as well.

Figure 4: The errors of DFT calculations with B3LYP functional and different dispersion corrections in HB375 and HB300SPX data sets and in the individual groups of HB300SPX.



Another outstanding feature in all but the few best DFT-D methods is the large error in the XH–S group, and this applies to some extent to XH–P systems as well. This issue is even more pronounced if we look at the errors relative to the strength of the interactions in each group (Table S6 and Figure S2 in the SI). Looking at signed errors, it is clear that the strength of these interactions is overestimated by these DFT methods. The distribution of the errors in

Figure 5: The errors of DFT calculations with B3LYP functional and different dispersion corrections in the HB300SPX×10 data set as a function of a intermolecular distance scaling factor.



the group is even, which means that this is a systematic trend rather than spurious behavior of some systems. Moreover, the interaction energy decomposition discussed above does not indicate any special character of the interactions in these groups. These interactions should not be difficult for DFT, so it is likely that this error stems from the dispersion correction. Unlike in the XH–I interaction, this issue is not specific only to some damping functions, so it must be some general effect. If we assume that the C_6 coefficients are reliable, the most likely explanation is that there is some imbalance in the atomic radii used in the damping function between the second and third period elements, and that the parametrization of DFT-D is biased towards the second period because these elements constitute the majority of the training set. This is, however, only a hypothesis that has to be revised as soon as more benchmark data are available also for other interactions than hydrogen bonds.

4.5 DFT Calculations – Dissociation Curves

The results of the analysis of DFT errors as a function of intermolecular distance scaling in HB300SPX×10 dissociation curves are presented in the Supporting Information; here, they will only be summarized briefly (Table S9 lists the overall RMSE as a function of the distance, and the errors in the individual groups are plotted in Figures S4 and S5). Overall,

the errors in the closest point are about 50% larger than in the equilibrium, which is not a bad result. In some combinations of a DFT functional and dispersion correction, the error drops to a value lower than in equilibrium in specific groups. This is a result of the interplay of the damping function in the dispersion correction, which has been already discussed.

At the other end of the dissociation curves, another trend can be observed. Although the errors there are consistently very small, they tend to be larger for non-hybrid DFT functionals, which indicates that the source of this error is overdelocalization of charge due to a self-interaction error.

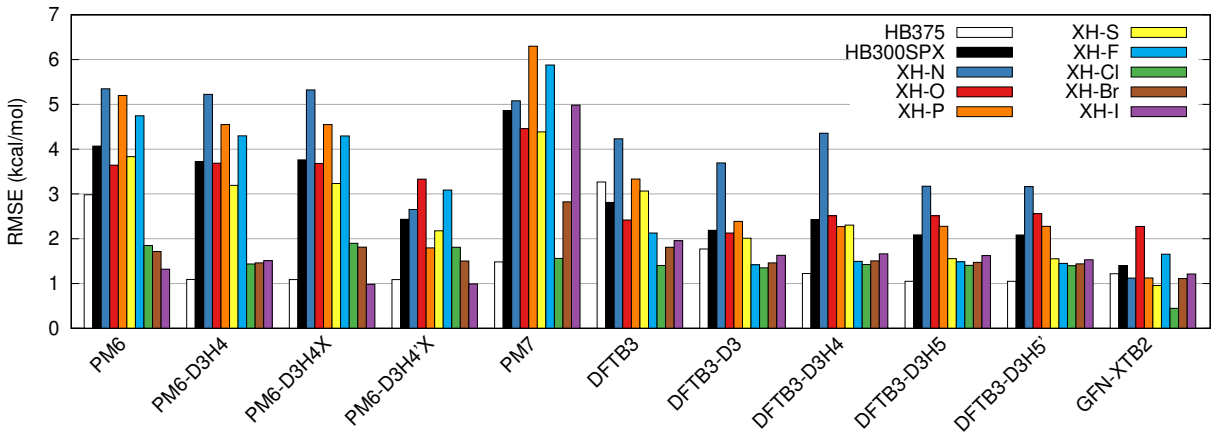
4.6 Semiempirical QM calculations

The present data set are more challenging for semiempirical QM methods. Their performance in such systems is limited not only by the approximations involved, but also by the lack of the reference data needed for their development. That is especially true for the recent NDDO-based methods such as PM6 and PM7, where large number of parameters was fitted to a limited set of mostly experimental data. The next approach tested here is DFTB3, which is less empirical because the core of the method is derived from DFT calculations. However, only little attention was paid to non-covalent interactions. In both cases, the issues in the description of non-covalent interactions were later addressed by additional corrections. For example in PM6-D3H4X, three corrections are applied – a general D3 correction for London dispersion, H4 correction for hydrogen bonding (which applies only to H-bonds of nitrogen and oxygen), and the X correction, which adds the extra repulsion needed to fix the description of halogen bonds. Similarly, DFTB3 is complemented by the D3 dispersion and a more advanced correction for H-bonds, H5. Finally, the GFN2-xTB method, an empirical tight-binding approach, is the only method in whose development non-covalent interactions were properly considered.

The results obtained with these methods, with and without the corrections discussed here, are summarized in the plot in Figure 6. More data (the RMSE and MSE by groups,

and by the distance scaling factor in the HB300SPX×10) are provided in the Supporting Information in Tables S10 – S13 and in Figure S6. These data also include the PM6-D3H4'X and DFTB3-D3H5' methods, which extend the parametrization of the corrections using the HB300SPX data set (their parametrization is described in the section below).

Figure 6: The errors of semiempirical QM calculations in the HB300SPX data set (black, individual groups in color) and in HB375 (white). The RMSE in kcal/mol.



The largest errors were obtained with PM7, although this method includes built-in corrections for non-covalent interactions and the training set comprised some non-covalent complexes. PM6, even without corrections, performs slightly better. The dispersion correction slightly improves the results and the H4 correction has no effect here – it only applies to hydrogen bonds between nitrogen or oxygen, which are not included in the set. The X correction improves the description of XH–I bonds but increases the error in XH–Cl and XH–Br groups. Only when the H4 correction is extended to the extra elements used here (in the PM6-D3H4'X method), the RMSE drops to 2.4 kcal/mol. The error is still rather large, but the improvement is reasonable in comparison to the uncorrected method. DFTB-D3 performs better as long as a dispersion correction is added, with the RMSE close to 2 kcal/mol. Hydrogen-bonding corrections play only a small role here, and no improvement is achieved when the parametrization of the H5 correction is extended to halogens in the DFTB3-D3H5' method. The only method to reach higher accuracy (the RMSE of 1.4 kcal/mol) is GFN2-xTB. It was developed with non-covalent interactions in mind (GFN stands for Geometries,

Frequencies and Non-covalent interactions), and so far it has proven itself to be a robust method applicable to a wide range of systems.

4.7 Extension of PM6-D3H4X and DFTB3-D3H5 to Additional Elements

The HB300SPX set was used to extend the parametrization of the H4 correction for hydrogen bonds in the PM6-D3H4X method.[12] The H4 correction is a product of multiple terms (Equation 1 in ref. 12), out of which only the radial part depends on adjustable parameters determining the strength of the correction. This term also uses several fixed parameters defining the shape of the potential, which were developed for nitrogen and oxygen only. These are applicable also to fluorine, which has similar size, but for the other elements discussed here, the form of the correction has to be modified. The radial part of the correction $f_{rad}(r_{DA})$ (Equation 2 in ref. 12), which is a function of the distance r_{DA} between the hydrogen-bond donor and acceptor (atoms X and Y in an XH–Y H-bond), is formulated as a smooth polynomial defined by a short-distance cutoff $r_{DA,0} = 1.5$Å, the position of the minimum $r_{DA,min} = 3.0$Å and a long-distance cutoff $r_{DA,cut} = 5.5$Å. To apply the same formula to the additional elements, the actual donor–acceptor distance r_{DA} is rescaled and enters the calculation as:

$$r_{DA,scaled} = r_{DA} * \frac{r_{DA,min}}{r_{XY,avg}}, \tag{1}$$

where $r_{XY,avg}$ is the average XY distance measured on the equilibrium geometries from the data set. This modification can be easily applied to the original implementation of the correction, and it is now available in the Cuby framework. The strength of the correction is determined by pairwise parameters c_{DA} for each combination of elements acting as a hydrogen donor and acceptor in the H-bond. Here, the original parameters are preserved for H-bonds of oxygen and nitrogen, and only new parameters are added for other elemental

combinations. Since not all the element pairs are covered well enough by the data set, the correction has only been extended to element combinations represented by at least five systems. Furthermore, the combinations in which the correction would be repulsive or very weak (c_{DA} after parametrization < 0.1) have been excluded. The resulting method is labeled here as PM6-D3H4'X.

The parameters were optimized one by one on the respective subset of HB300SPX by means of gradient optimization with the target of minimizing the RMSE. Note that only equilibrium geometries are used here – this prevents problems with cases where the error is even larger at shorter distances. The resulting parameters are listed in the SI, Tables S14 and S15. The error in the complete HB300SPX data set (i.e. at equilibrium distances) drops from 3.76 to 2.44 kcal/mol. What is more important is the change in the systematic error. The original PM6-D3H4X has the MSE of 2.35 kcal/mol in the HB300SPX set, while PM6-D3H4'X has a much smaller MSE of 0.14 kcal/mol. As it was not the target of the optimization, its reduction shows that the correction is successful in the removal of the systematic underestimation of the strength of hydrogen bonds in PM6. More details on the performance of PM6-D3H4'X in HB300SPX and HB300SPX×10 data sets can be found in Figure 6 and in the Supporting Information Tables S10–S13.

Since the whole data set was used for the parametrization, there is only a limited amount of data that can be used to validate it. There are ten hydrogen bonds in the X40 data set, and the RMSE in this subset of X40 drops from 4.48 to 1.41 kcal/mol when we pass from PM6-D3H4X to PM6-D3H4'X. The error in the whole data set becomes only 0.88 kcal/mol, which is a very good result – hydrogen bonds were the last issue there to be fixed, and PM6-D3H4X yielded the RMSE of 2.3 kcal/mol. This also indicates that the extension of the correction does not interfere with the description of halogen bonds, which comprise the major part of the X40 data set.

Despite these positive results, this extension of the correction is only a proof of concept rather than a method that can be used as a black box. It shows how far it is possible to

get with standalone corrections to existing methods, but it is clear that future methods should handle this problem by other means than by introducing pairwise parameters. Some of the parameters (especially for H-bonds of fluorine) have very large values, which suggests that the problem is bigger than what can be safely fixed with a standalone correction. The PM6-D3H4'X method should thus be used with care; it might lead to unexpected artifacts in other systems not covered by the limited validation performed here.

A similar reparametrization has been applied to the H5 correction in the DFTB3-D3H5 method. Here, the spatial component is expressed using van der Waals radii of the atoms, and no modifications to the form of the corrections are needed. There is only one parameter per element determining the strength of the correction to be added. The correction already covers nitrogen, oxygen and sulfur. For the remaining elements, new parameters have been derived using the respective groups of the HB300SPX data set. No improvement can be achieved for phosphorus; despite a positive MSE in this group, the RMSE grows when the correction is applied. For the remaining elements, the improvements are practically negligible as DFTB3-D3H5 does not exhibit any significant systematic errors there. The results (labeled as DFTB3-D3H5') are provided in the plots and tables here and in the Supporting Information. The values of the parameters used are listed in the SI, Table S16, but they are provided only for the record and there is no need to use them in practice.

5 Conclusions

The HB300SPX×10 data set extends the Non-Covalent Interactions Atlas to hydrogen bonds involving sulfur, phosphorus and halogens. It contains 300 systems covering all relevant combinations of these elements, for which a total of 3000 CCSD(T)/CBS data points are available. The main purpose of this data set is the development of future methods, but this paper also brings several interesting findings discovered using this data set.

This data set is a tough test for dispersion-corrected DFT calculations. It includes heav-

ier elements, which were often underrepresented in the parametrization of the corrections; hydrogen bonds make them interact at very short distances, where the transition between uncorrected DFT and the correction occurs. As a result, some DFT functionals that perform well in simple organic molecules yield much larger errors here. On the other hand, there are DFT methods that performed surprisingly well, e.g. the functionals from the ωB97 family. It has also been found that the atomic radii used in the D3 and D4 dispersion corrections with Becke-Johnson damping lead to large errors for iodine, which suggests that the procedure used to derive them may not be suitable for heavier elements.

Hydrogen bonds of less common elements are even more challenging for semiempirical QM methods. Not only because of the approximations involved, but also because such systems are underrepresented or even missing in the reference data on which these methods are parametrized. Even with all possible corrections applied, both PM6 and DFTB3 yield errors at least twice as large as in the HB375 data set of hydrogen bonds of simple organic molecules. It is difficult to correct these errors further; in PM6-D3H4, the improvement is achieved only at the cost of introducing a large number of parameters, and extending the DFTB3-D3H5 to halogens has not brought any improvement.

It is clear that this issue can be properly addressed only if non-covalent interactions are considered from the very beginning of the development of a method and suitable model systems are included in the reference data. A good example of this approach is the GFN2-xTB method, which has been found to perform well even in these challenging systems. Here lies the main importance of the HB300SPX×10 data set – it provides data that can be used to improve the next generation of semiempirical QM methods.

So far, the Non-Covalent Interactions Atlas covers only hydrogen bonds, which is not enough to develop a general-purpose method. However, most of the remaining data needed are now under development. The following data sets will cover London dispersion, repulsive contacts and σ-hole interactions in the same or slightly larger chemical space.

6 Acknowledgements

We acknowledge the support from the Czech Science Foundation, Grant No. 19-13905S, and from the European Regional Development Fund, OP RDE, Project: Chemical Biology for Drugging Undruggable Targets (Chem-BioDrug, No. CZ.02.1.01/0.0/0.0/16_019/0000729). This work is part of the Research Project RVO 61388963 of the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences. It has also been supported by the Ministry of Education, Youth and Sports from the Large Infrastructures for Research, Experimental Development and Innovations project IT4Innovations National Supercomputing Center, LM2015070.

Supporting Information Available

The Supporting Information contains: 1) additional tables and figures supporting the main text, including tables of the data used in the plots featured here, 2) geometries of all the systems and tables of benchmark interaction energies needed for the reproduction of the results presented here, and 3) a machine-readable definition of the data set containing the metadata describing the categorization of the systems.

References

(1) Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding. *J. Chem. Theory Comput.* **2020**, *16*, 2355–2368, Publisher: American Chemical Society.

(2) Řezáč, J.; Dubecký, M.; Jurečka, P.; Hobza, P. Extensions and applications of the A24 data set of accurate interaction energies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19268–19277.

(3) Kodrycka, M.; Patkowski, K. Platinum, gold, and silver standards of intermolecular interaction energy calculations. *J. Chem. Phys.* **2019**, *151*, 070901.

(4) Řezáč, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.

(5) Kesharwani, M. K.; Manna, D.; Sylvetsky, N.; Martin, J. M. L. The X40x10 Halogen Bonding Benchmark Revisited: Surprising Importance of (n–1)d Subvalence Correlation. *J. Phys. Chem. A* **2018**, *122*, 2184–2197.

(6) Řezáč, J. Cuby: An integrative framework for computational chemistry. *J. Comput. Chem.* **2016**, *37*, 1230–1237.

(7) Řezáč, J. Cuby 4, software framework for computational chemistry. 2015; `http://cuby4.molecular.cz/`.

(8) 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.

(9) 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.

(10) Řezáč, J.; Fanfrlík, J.; Salahub, D.; Hobza, P. Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749–1760.

(11) Korth, M.; Pitoňák, M.; Řezáč, J.; Hobza, P. A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods. *J. Chem. Theory Comput.* **2010**, *6*, 344–352.

(12) Řezáč, J.; Hobza, P. Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. *J. Chem. Theory Comput.* **2012**, *8*, 141–151.

(13) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J Mol Model* **2007**, *13*, 1173–1213.

(14) Řezáč, J.; Hobza, P. A halogen-bonding correction for the semiempirical PM6 method. *Chem. Phys. Lett.* **2011**, *506*, 286–289.

(15) Řezáč, J. Empirical Self-Consistent Correction for the Description of Hydrogen Bonds in DFTB3. *J. Chem. Theory Comput.* **2017**, *13*, 4804–4817.

(16) Gordon, A. D. A Review of Hierarchical Classification. *Journal of the Royal Statistical Society. Series A (General)* **1987**, *150*, 119–137.

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

(18) Witte, J.; Goldey, M.; Neaton, J. B.; Head-Gordon, M. Beyond Energies: Geometries of Nonbonded Molecular Complexes as Metrics for Assessing Electronic Structure Approaches. *J. Chem. Theory Comput.* **2015**, *11*, 1481–1492.

(19) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **2018**, *8*, e1327.

(20) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *The Journal of Chemical Physics* **1997**, *106*, 9639–9646.

(21) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975.

(22) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, δMP2CCSD(T): Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases. *The Journal of Chemical Physics* **2011**, *135*, 194102.

(23) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16-18 elements. *J. Chem. Phys.* **2003**, *119*, 11113–11123.

(24) 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*, 10548–10560, WOS:000179495000015 bibtex: cc-pwCVXZ.

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

(26) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. Psi4: an open-source ab initio electronic structure program. *WIREs Comput Mol Sci* **2012**, *2*, 556–565.

(27) TURBOMOLE v7.3. 2018; `http://www.turbomole.com`.

(28) Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. Scaled MP3 Non-Covalent Interaction Energies Agree Closely with Accurate CCSD(T) Benchmark Data. *ChemPhysChem* **2009**, *10*, 282–289.

(29) Hohenstein, E. G.; Sherrill, C. D. Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010**, *133*, 014101.

(30) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.

(31) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.

(32) Witte, J.; Mardirossian, N.; Neaton, J. B.; Head-Gordon, M. Assessing DFT-D3 Damping Functions Across Widely Used Density Functionals: Can We Do Better? *J. Chem. Theory Comput.* **2017**, *13*, 2043–2052.

(33) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103, Publisher: American Institute of Physics.

(34) 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, Publisher: American Chemical Society.

(35) 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, Publisher: American Institute of Physics.

(36) 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, Publisher: The Royal Society of Chemistry.

(37) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density

Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272, Publisher: American Chemical Society.

(38) 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, Publisher: American Chemical Society.

(39) Kozuch, S.; Gruzman, D.; 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.

(40) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(41) DFTD4 program. 2019; `https://github.com/dftd4/dftd4`, original-date: 2019-02-28T15:45:22Z.

(42) Stewart, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J Mol Model* **2013**, *19*, 1–32.

(43) Stewart, J. J. P. MOPAC 2016. 2016; `http://openmopac.net/`.

(44) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

(45) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.

(46) Kubillus, M.; Kubař, T.; Gaus, M.; Řezáč, J.; Elstner, M. Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. *J. Chem. Theory Comput.* **2015**, *11*, 332–342.

(47) Grimme, S. Towards First Principles Calculation of Electron Impact Mass Spectra of Molecules. *Angew. Chem. Int. Ed.* **2013**, *52*, 6306–6312.

(48) Miriyala, V. M.; Řezáč, J. Description of non-covalent interactions in SCC-DFTB methods. *J. Comput. Chem.* **2017**, *38*, 688–697.

(49) Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshaye, M. Y.; Dumitrică, T.; Dominguez, A.; Ehlert, S.; Elstner, M.; van der Heide, T.; Hermann, J.; Irle, S.; Kranz, J. J.; Köhler, C.; Kowalczyk, T.; Kubař, T.; Lee, I. S.; Lutsker, V.; Maurer, R. J.; Min, S. K.; Mitchell, I.; Negre, C.; Niehaus, T. A.; Niklasson, A. M. N.; Page, A. J.; Pecchia, A.; Penazzi, G.; Persson, M. P.; Řezáč, J.; Sánchez, C. G.; Sternberg, M.; Stöhr, M.; Stuckenberg, F.; Tkatchenko, A.; Yu, V. W.-z.; Frauenheim, T. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *J. Chem. Phys.* **2020**, *152*, 124101, Publisher: American Institute of Physics.

(50) DFTB+. `http://www.dftb-plus.info/`.

(51) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(52) XTB, Semiempirical Extended Tight-Binding Program Package. 2019; `https://github.com/grimme-lab/xtb`, original-date: 2019-09-30T12:40:09Z.

(53) The Official YAML Web Site. `http://yaml.org/`.

(54) 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.

(55) DiStasio, R.; Head-Gordon, M. Optimized spin-component scaled second-order Moller-Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007**, *105*, 1073–1083.

(56) Řezáč, J.; Greenwell, C.; Beran, G. J. O. Accurate Noncovalent Interactions via Dispersion-Corrected Second-Order Møller–Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2018**, *14*, 4711–4721.

(57) Hostaš, J.; Řezáč, J. Accurate DFT-D3 Calculations in a Small Basis Set. *J. Chem. Theory Comput.* **2017**,

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

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

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

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