Extension and Evaluation of the D4 London Dispersion Model for Periodic Systems

Eike Caldeweyher, Jan-Michael Mewes, Sebastian Ehlert, and Stefan Grimme[a†]

London-dispersion effects are of great relevance to many aspects of materials science and for various condensed matter problems. In this work we present an adaptation and implementation of the DFT-D4 model [Caldeweyher *et al., J. Chem. Phys.*, 2019, **150**, 154122] for periodic systems. The main new ingredient are better computed reference polarizabilities for high coordination numbers (including alkaline metals, earth alkaline metals, and *d*-metals of group 3-5), which are consistently derived from periodic electrostatically embedded cluster calculations. Some technical extensions have been added concerning the coordination number, the partial charges, and the dispersion energy expression. To demonstrate the performance of the improved scheme, several test cases are considered, for which we compare D4 results to those of its predecessor D3(BJ) as well as to several other dispersion corrected methods. The largest improvements are observed for solid state polarizabilities of 16 inorganic salts, where the new D4 model achieves an unprecedented accuracy, surpassing its predecessor as well as other, computationally much more demanding approaches. For cell volumes and lattice energies of two sets of chemically diverse molecular crystals, the accuracy gain is less pronounced compared to the already excellently performing D3(BJ) method. For the challenging adsorption energies of small organic molecules on metallic as well as on ionic surfaces, DFT-D4 provides high accuracy similar to MBD/HI or uncorrected DFT/SCAN approaches. These results suggest the standard application of the proposed periodic D4 model as a physically improved yet computationally efficient dispersion correction for standard DFT calculations as well as low-cost approaches like semi-empirical or even force-field models.

1 Introduction

The efficiency of modern Kohn–Sham density functional theory [1,2] (DFT) enables its routine application to very large molecules with thousands of atoms, as well as the systematic screening of huge numbers of smaller compounds. [3] However, DFT as approximate mean-field electronic structure method can fail dramatically for certain types of interactions. The two most prominent and relevant problems are (i) the absence or severe underestimation of long-range electronic-correlation effects, which give rise to London dispersion interactions, [4] and (ii) the so-called self-interaction error (SIE). [5,6] The SIE problem will not be further discussed in this work, however, the focus of the present study is the extension of a correction scheme which attempts to solve the first mentioned London dispersion problem.

An accurate account of dispersion interactions is indispensable for any reasonable description of chemically relevant molecules and even more so for condensed-phase systems. Accordingly, it has shown, that dispersion-corrections systematically improve the accuracy of density functionals on all rungs of Jacob's ladder. [7] Hence, the development of dispersion correction schemes is a very active field of research which has spawned several different approaches. The unifying idea of all these approaches is to reintroduce the in principle well understood physics of London dispersion to the DFT framework. The existing approaches may be organized in two categories. On the one hand, there are self-consistent schemes in which dispersion is directly included in the functional, *e. g.*, *via* a response function and/or non-local dispersion kernels, and on the other with additive post-SCF-type corrections which model dispersion based on atomic polarizabilities. The latter may be subdivided in corrections which explicitly take into account the electron density, and semi-classical approaches that are electron density independent.

Most approaches incorporate the electron density into their theoretical apparatus. These include the exchange-hole dipole moment (XDM) approach of Becke and Johnson, [8,9] as well as the Tkatchenko–Scheffler (TS) model, [10] and its many-body dispersion (MBD) successor. [11,12] All those methods employ a Hirshfeld-

[a] *Mulliken Center for Theoretical Chemistry, Mulliken Center for Theoretical Chemistry, Bonn, Germany.*
† E-mail: grimme@thch.uni-bonn.de

type atomic partitioning of the electron density, either to rescale the polarizabilities of the neutral atoms (TS, MBD), or to density-weight the approximated expectation value of the squared total multipole operator for each atom within the molecule (XDM). For the TS related approaches, Bučko and co-workers[13,14] showed that an iterative Hirshfeld partitioning, in which the fractionally charged atomic reference state is determined self-consistently, can drastically improve the description of dispersion interactions.

Approaches of the first category, which self-consistently include dispersion in the density functional circumvent the somewhat arbitrary atomic partitioning of the electron density. However, this comes at the cost of a reduced flexibility — coupling to any method is hardly possible — and increased computational costs. Famous approaches of this class are the van der Waals density functionals[15–20] (vdW-DF), which are based on the fundamental adiabatic connection theory. A simplified construction scheme for the non-local correlation part has been introduced by Vydrov and Van Voorhis (rVV10).[21,22]

The perhaps most widespread correction schemes and main topic of this work are the semi-classical DFT-D methods of the last category.[4,23] Testing representatives from each of these categories, including TS(TS/HI), MBD(MBD/HI,MBD/FI), opt-vdw-DF2, rVV10, and DFT-D, we find that they are all equally able to accurately account for dispersion interactions in various scenarios within a DFT framework. However, the associated computational costs of these methods can differ drastically. The overarching principle for the cost of any such correction should be that the computational efficiency of the underlying DFT method must be retained. While this is in general the case for all of these methods in combination with hybrid functionals employing large Gaussian basis sets, the correction can become the bottleneck in low-cost GGA DFT calculations, as will be shown and discussed later in this article. For even lower-cost atomistic models such as semi-empirical molecular orbital or force-field methods, only the fastest approaches remain viable. An important point concerning the computational efficiency of additive dispersion-correction schemes is their account of the chemical environment. While it appears as a logical step to derive and/or scale the atomic polarizabilities based on the electron density, this also introduces a computationally demanding step, which typically slows down the calculation.

One attempt to overcome this problem forms the basis of the prominent DFT-D3 model,[24,25] in which the atomic polarizabilities are interpolated between atomic references based on fractional coordination numbers (CNs). This avoids the need to deal with the complex electron density, while the chemical environment is implicitly accounted-for by the CNs. The recently introduced successor DFT-D4 further improves this model by making the atomic reference polarizabilities charge-dependent.[26,27] However, to retain the computational efficiency, the charge-information is not derived from the electron density, but provided by a classical and thus very efficient electronegativity equilibration (EEQ) model calculation. While this charge-scaling improves the description for most molecular applications, in particular for transition metals,[28] there are still issues with certain metal-ions in highly polar and periodic environments, such as Na^+ in crystalline NaCl. By analyzing these failures, we noted that the charge-scaling is not the root cause of the problem, but rather the absence of suitable reference systems in the model. In the solid state, the coordination number of alkaline, alkaline-earth and early transition metals can approach values far beyond those for which references polarizabilities are available from molecular treatments, causing the otherwise accurate interpolation to become an unreliable extrapolation.

The present work addresses this problem by widening the scope of the references. For this, polarizabilities for new, highly coordinated systems are added based on pseudo-periodic model calculations with the periodic electrostatic embedded cluster method (PEECM). This includes new polarizabilities for alkaline, alkaline-earth, and early transition metal elements. Herein theoretically observed solid state polarizabilities are evaluated against experimentally determined ones. We check the quality for properties such as organic polymer refractive indices which are directly dependent on polarizabilities. Lattice energies and cell volumes of molecular crystals are evaluated using different literature known London dispersion correction schemes in comparison to the new approach. Furthermore, interfaces between organic molecules and inorganic surfaces are treated for which new properties can emerge.[29–31] Noncovalent London dispersion effects are essential for structural properties, but also for the quantification of the stability of such systems.[32–34] For this purpose we discuss adsorption energies of organic molecules on non-polar and ionic surfaces which either exhibit partly covalent or partly electrostatically driven binding motifs penetrating the particular description of noncovalent interactions (NCIs) between adsorbate and surface.

The next section shortly introduces the DFT-D4 theory, which defines the method under periodic boundary conditions. A detailed description of the methodologies—especially the generation of periodic EEQ partial charges and the periodic dispersion energy expression— is given in the Supplementary Material. A special focus shall be put on a scheme to calculate periodic reference polarizabilities. Afterwards the results for the above mentioned tests will be discussed. Finally, a summary and an outlook are given presenting the quality of the method on solid state properties and discuss possible future applications.

2 Theory

The theoretical framework of this work was presented in detail in Ref. 27 describing the basics of the molecular DFT-D4 model. For the calculation of periodic CNs, charge and geometry dependent dynamic polarizabilities $\alpha(i\omega)$, as well as for pair and triple-wise dispersion energies see the detailed theoretical description given in the Supplementary Material. A periodic EEQ model is developed within the present work for the efficient calculation of partial charges used in the ζ-scaling for atom-in-molecules polarizabilities (see Supplementary Material for further details). For this purpose, a cyclic cluster model is used to capture periodic boundary conditions. Since dispersion interactions are much faster decaying (leading order term is proportional to R^{-6}) than, *e. g.,* Coulomb interactions, we employ for this energy contribution a real-space cutoff within the periodic implementation. Analytical

gradients (nuclear forces) are available for the dispersion energy expression including pairwise and triple wise London dispersion interactions.

2.1 Dynamic polarizabilities from cluster extrapolation

A major problem for D3/D4 calculations of dense periodic systems has been the absence of suitable molecular reference systems. Consider for example the alkaline (earth) metals, for which there are only singly or doubly coordinated molecular references (KH, MgH2), but no references with CNs close to the ones realized, *e. g.*, in salts (KCl, $MgCl_2$, etc.). This problem persisted up to group 5, and causes too large polarizabilities to be used in these situations and in turn an over-stabilization of such systems. In order to eliminate this shortcoming, it is desirable to augment the database of references with dynamic polarizabilities of the atoms in exactly such situations. For this purpose, the polarizability per crystal unit cell and eventually that of the anions and cations has to be calculated, which can in principle be accomplished in two ways: (i) the calculation of polarizability per volume from the dielectric function of the solid (using, *e. g.,*, the Clausius–Mosotti equation[35] in the limit of ideal ionic crystals), and (ii) the calculation of polarizability per atom from a cluster extrapolation. Zhang *et al.* have shown that a cluster extrapolation gives satisfactory results for obtaining atomic polarizabilities within the solid state (maximum deviation of about 2% for the polarizability of silicon)[36]. Since this latter approach is moreover more similar with the method used to obtain molecular reference polarizabilities, it is most suitable for our purpose.

To simulate periodic boundary conditions of the crystal as close as possible we apply the PEECM.[37] Within this approach the entire periodic system is divided into three parts: the inner part covering the cluster (I), an optional shell (II) which is build from effective core potentials (ECPs), and the outer part (III) which describes its environment. Part (I) is treated quantum mechanically (QM), whereby in part of (II) ions are replaced by ECPs. Such an isolating ECP shell—surrounding the actual QM part—is necessary in order to prevent artificial polarization of the electron density by ions which would otherwise be in direct contact with the QM boundary. The outer part (III) is described by a periodic array of point charges, representing cationic and anionic sites of the perfect ionic crystal. The effect of the additional ECP shell is briefly shown in table 1 where a simple point charge embedding (PCE) model describes the cluster without additional ECPs.

As evident from table 1, the polarizability of lithium as part of an isolated Li_4Cl_4 cluster (i.e., no embedding) decreases significantly in PBC. Furthermore, when applying a PCE model, it is crucial to use adjusted point charges, since properties such as polarizabilities are strongly influenced by the present Coulomb field. The effect of unadjusted point charges within a PCE model ($q^{PC} = \pm 1$) results in artificially low polarizabilities and thus small dispersion coefficients. However, this can be overcome by means of an ECP shell, without having to manually adjust the point charge in advance (see "PEECM" embedding in table 1). Thus, it is advisable to embed the QM region into ECPs to get the right answer for the right reason. The electronic Coulomb energy term arising from

Table 1 Static polarizabilities (given in $Bohr^3$) and homoatomic dispersion coefficients (given in Hartree $Bohr^6$) as obtained by a cluster extrapolation for lithium inside the lithium chloride crystal. Here the effect of an ECP shell on the absolute size of the particular property within the crystal is investigated starting from a Li_4Cl_4 cluster. Furthermore, the point charges (termed q^{PC}) creating the outer part are given.

Embedding	q^{PC}	$\alpha^{Li}(0)$	C_6^{LiLi}
None	0.0	37.2	129.7
PCE	± 0.3	8.0	15.4
PCE	± 1.0	<0.1	0.1
PEECM	± 1.0	5.0	7.6

the periodic field of point charges surrounding the cluster has the following form

$$J = \sum_{\mu\nu} \sum_{k}^{N\in UC} \sum_{\mathbf{T}\in O}^{\infty} D_{\mu\nu} q_k \int d\mathbf{r}\, \frac{\mu(\mathbf{r})\nu(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_k - \mathbf{T}|}, \tag{1}$$

where UC denotes the unit cell of point charges, $D_{\mu\nu}$ are elements of the density matrix, μ and ν are basis functions, q_k and $\mathbf{R}_k$ denote charges and positions of point charges, and $\mathbf{T}$ denotes the direct lattice vector of the outer part III. The energy term is evaluated using the periodic fast multipole method[38] (PFMM) which—unlike the Ewald method—defines the lattice sums entirely in the direct space. Fig. 1 schematically shows the distribution of the cluster into three parts.

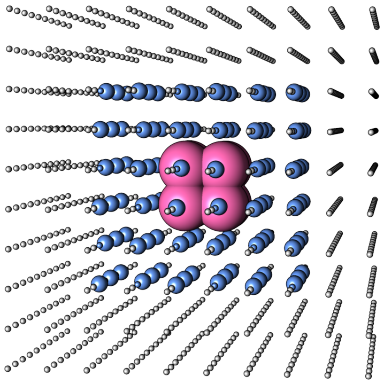

Fig. 1 The schematic construction of a cluster within the periodic electrostatic embedded cluster method. Part (I) consists out of the QM part (pink spheres) embedded into part (II) which is build from ECPs (blue atoms). Part (III) is embedding part (II) by periodic point charges (small gray atoms) representing cationic and anionic sites as in the perfect ionic crystal.

The correct representation of polarities for atoms inside solids without using QM information still remains a challenge. For the geometry dependent D3 method this is hardly possible, because only the CN is used to weight reference systems accordingly. This disadvantage was partly overcome with the development of the D4 method by including information from the electronic structure (atomic partial charges) in the calculation of the atomic dipole polarizabilities. Nevertheless, periodic dipole polarizabilities are significantly smaller than those present in molecules which rises

the need of reference systems that are especially designed for representing solid state properties. An addition of such references is easily feasible in the D4 method, since only dynamic polarizabilities of the new references are required, from which the corresponding dipole–dipole C_6 dispersion coefficients are obtained—by means of numerical Casimir–Polder integration—during the calculation itself. Figure 2 highlights the elements for which new periodic reference polarizabilities have been calculated. In particular, the elements of group 1–5 were supplemented with new references.

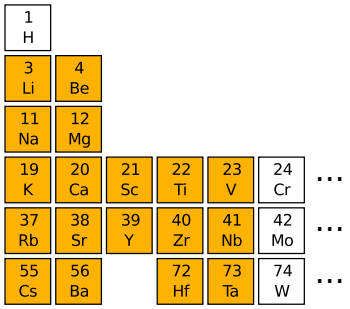

Fig. 2 The yellow highlighted elements received new periodic reference polarizabilities at the PEECM level. All these elements exposed relatively high polarizabilities for molecular reference systems.

All new reference system structures are given as a tarball within the Supplementary Material.

3 Results

3.1 Polarizabilities and refractive indices

Accurate polarizabilities are the foundation of many London-dispersion correction schemes. It is therefore of utmost importance to be able to calculate polarizabilities as exactly as possible. As demonstrated in previous works,[24,26,27] there is a direct correlation between the description of molecular polarizabilities and the accurate description of NCI energies. Therefore, this section will briefly discuss the quality of the new reference polarizabilities. In addition, ionic polarizabilities are shown representing the polarizabilities of cations and anions inside the crystal. In general, cations have a lower electron density and thus a lower local (atomic) polarizability and *vice versa*. The last part of this section covers the calculation of refractive indices of several optical organic polymers to validate the quality of polarizabilities in "organic" crystals where the new reference polarizabilities are supposed to have a small effect.

The experimental data were obtained by measuring the refractive index n (RI) at several wavelengths λ. In Ref. 35 all RIs are extrapolated to $\lambda = \infty$ (obtaining n_∞) and used within the Lorentz expression to calculate the particular experimental salt polarizability

$$\alpha_{\text{salt}}^{\text{exp}} = \frac{3V_m}{4\pi}\frac{n_\infty^2 - 1}{n_\infty^2 + 2}, \tag{2}$$

where V_m is the volume of the crystal divided by the number of molecules inside the crystal. All V_m values are taken from X-ray data[39] and all theoretical salt polarizabilities are obtained as the

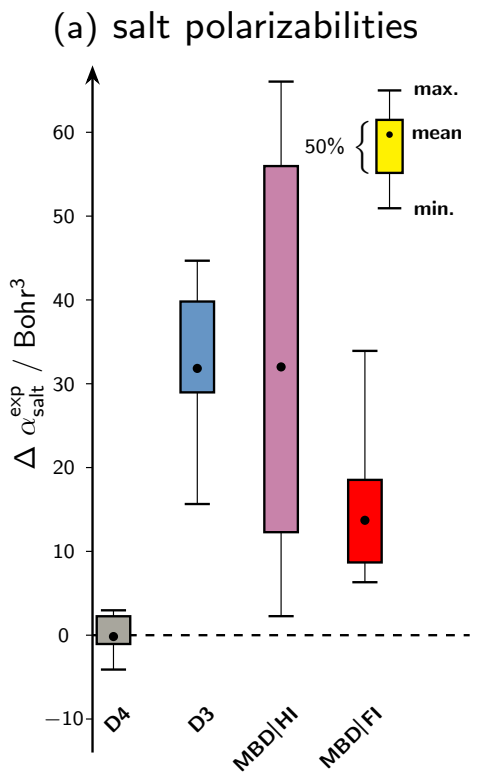

Fig. 3 Given are (a) deviations of salt polarizabilities with respect to experimentally obtained values taken from Ref. 35 and (b) schematic 2D-representation for anionic (yellow) and cationic (blue) polarizabilities within the KCl (top) and RbI (bottom) solid state. Given are the experimentally derived ionic polarizabilities (left) as computed by Tessman and coworkers,[35] the calculated atomic D4 polarizabilities (right), and the D3 polarizabilities (center). The radii of the spheres correspond to the absolute value of the polarizability.

sum of the cationic and the anionic polarizability

$$\alpha_{\text{salt}}^{\text{calc}} = \alpha_{\text{salt}}^{\text{cation}} + \alpha_{\text{salt}}^{\text{anion}}. \tag{3}$$

For D3, atomic polarizabilities are calculated as introduced in Ref. 24, hence we do not approximate polarizabilities as introduced in the empirical relationship for determining polarizabilities from homoatomic dispersion coefficients[40,41]. Furthermore, the Tkatchenko-Scheffler[10] method with iterative Hirshfeld (HI) partitioning[13,14] (termed TS/HI) as well as its many-body dis-

persion analogon[11,12] (termed MBD/HI) are listed for the sake of completeness[13,14,42]. Here an HI analysis is indispensable, since the TS and MBD variants calculate cationic polarizabilities that are too high in their absolute value (similar to neutral atomic polarizabilities) leading to artificially high salt polarizabilities. Furthermore, the MBD/FI (fractionally ionic) variant is used, which was developed by Gould *et al.*[43] This method promises in particular a drastic improvement for the calculation of polarizabilities in ionic systems. In figure 3 the quality of salt polarizabilities are given for the best performer methods namely for two DFT-D methods (DFT-D3 and DFT-D4) as well as for two MBD methods (MBD/HI and MBD/FI).

Extended statistical evaluations are given for the mean deviation (MD), the mean absolute deviation (MAD), the root mean squared deviation (RMSD), and the absolute maximum error (AMAX) and are given in Bohr^3 if not stated otherwise. The experimentally determined polarizabilities are reproduced with high accuracy by the DFT-D4 method (MD $= -0.2$, MAD $= 1.7$, RMSD $= 2.1$, AMAX $= 4.1$). MBD is not able to calculate reasonable polarizabilities for seven out of 16 different salts due to non-physical negative values after its screening procedure. Note that this failure is already literature-known.[43] MBD/HI calculates polarizabilities for all salts without such problems, however, this method gives only poor results (MD $= 32.0$, MAD $= 32.0$, RMSD $= 21.4$, AMAX $= 66.1$). This is slightly improved at the density-independent DFT-D3 level (MD $= 31.8$, MAD $= 31.8$, RMSD $= 9.5$, AMAX $= 44.7$). Compared to DFT-D3, further improvements are obtained when using the MBD/FI method where the fractional ionic reference systems are beneficial for describing salt polarizabilities (MD $= 14.6$, MAD $= 14.6$, RMSD $= 8.4$, AMAX $= 40.6$). TS and TS/HI show overall the largest deviations. Here even the iterative Hirshfeld partitioning cannot lower the absolute values of the neutral TS atomic polarizabilities enough to produce reasonable salt polarizabilities (TS: MD $= 192.0$, MAD $= 192.0$, RMSD $= 199.0$, AMAX $= 256.3$, and TS/HI: MD $= 114.1$, MAD $= 114.1$, RMSD $= 118.4$, AMAX $= 158.2$).

Until now we discussed salt polarizabilities whose values were obtained as the sum of the particular cationic and anionic polarizabilities. However, to check for ionic polarizabilities themselves, we use experimental ionic polarizabilities that have been created by partitioning salt polarizabilities to the contributions of their cations and anions using a least-squares fitting procedure.[35] Those experimental ionic polarizabilities are compared to theoretically derived polarizabilities representing the cation and the anion inside the crystal. The lower part of figure 3 schematically highlights how D4 and D3 obtain ionic polarizabilities for K^+ and Cl^- in KCl and for Rb^+ and I^- in RbI (both space group $Fm\bar{3}m$) compared to experimental ionic polarizabilities. Here, the absolute value of polarizabilities are expressed in sphere-radii, where large radii correspond to large polarizabilities and *vice versa*. Furthermore, cationic polarizabilities are represented in blue and anionic polarizabilities in yellow. The agreement to experimental ionic polarizabilities is significantly improved by adding periodic reference polarizabilities in combination with a charge scaling procedure as it is done within the periodic D4 method when comparing to the D3 method (*i.e.*, $\alpha_{exp}^{Rb\in RbI}/\alpha_{exp}^{I\in RbI} = 0.3$ versus $\alpha_{D4}^{Rb\in RbI}/\alpha_{D4}^{I\in RbI} = 0.4$, and $\alpha_{D3}^{Rb\in RbI}/\alpha_{D3}^{I\in RbI} = 1.8$). A comparison with the TS and the MBD model shows that the scaling of neutral atomic polarizabilities using Hirshfeld volumes is insufficient for the accurate calculation of ionic polarizabilities within RbI ($\alpha_{TS}^{Rb\in RbI}/\alpha_{TS}^{I\in RbI} = 8.0$ and $\alpha_{MBD}^{Rb\in RbI}/\alpha_{MBD}^{I\in RbI} = 13.1$). TS/HI and MBD/HI perform better in determining ionic polarizabilities within RbI compared to their non-iterative analogues ($\alpha_{TS/HI}^{Rb\in RbI}/\alpha_{TS/HI}^{I\in RbI} = 4.5$ and $\alpha_{MBD/HI}^{Rb\in RbI}/\alpha_{MBD/HI}^{I\in RbI} = 1.3$). The MBD/FI method significantly improves here ($\alpha_{MBD/FI}^{Rb\in RbI}/\alpha_{MBD/FI}^{I\in RbI} = 0.1$) whereby the cationic polarizability of Rb^+ is obtained somewhat too small and the anionic polarizability of I^- too large within this salt. In general it is assumed that improved ionic polarizabilities also yield improved interaction energies for ionic systems (*e. g.*,, adsorption processes on ionic surfaces) which will be discussed in more detail in section 3.3.

Another test case is the calculation of RI values for optical organic polymers, which is based on the work of Hachmann and co-workers.[44] In their work, polymer polarizabilities of organic compounds were calculated using hybrid TD-DFT (TD-PBE0/def2-TZVP) and RIs were obtained using the Lorentz equation. Contrary to their work, we calculate RIs from molecular polarizabilities of the monomer units of each polymer obtained by approximated methods (DFT-D, TS, or MBD). This approximation, however, is accompanied by a lack of many-body effects, so that the RIs to be calculated are known to be too large in absolute value. Furthermore, monomer structures are used, since the explicit polymer growth (including conformational analysis for each polymer) is beyond the scope of this work. However, this is intended to be covered in more detail in future works, *e.g.*, by interfacing the supramolecular toolkit of Jelfs and co-workers[45] to explicitly build polymer structures. Overall 73 monomer units of polymer structures have been extracted from Ref. 44 and re-optimized at the PBEh-3c[46] level of theory after searching for minimum conformers using the conformer-rotamer ensemble sampling tool[47] for each monomer unit. We calculate RI values as follows

$$n_\infty = \sqrt{\frac{3V_m + 8\pi\cdot\alpha}{3V_m - 4\pi\cdot\alpha}}, \tag{4}$$

where V_m values are taken from Ref. 44 assuming a constant packing fraction of the bulk polymer. Figure 4 shows the relative deviation of RIs from experimental values[44] (the MAD of TD-DFT RIs from experimental RIs is 2.3% its RMSD 3.0%). As can be seen from the graph, all methods generally calculate too large RI values, which is partly due to the approximation of using molecular polarizabilities α in equation 4. However, since this approximation is applied to all methods, they can be directly compared with each other. The smallest MD and RMSD values are obtained from the D3 and D4 methods, which are able to determine RI values with good accuracy. The slightly better performance of D3 compared to D4 is within statistical uncertainty and expected, as D3 already provides highly accurate polarizabilities for organic molecules and molecular crystals. Subsequently, the MBD/FI method from Gould and co-workers achieves adequate accuracy. The least suitable method for predicting RIs is the TS/HI method, which achieves improvements through addi-

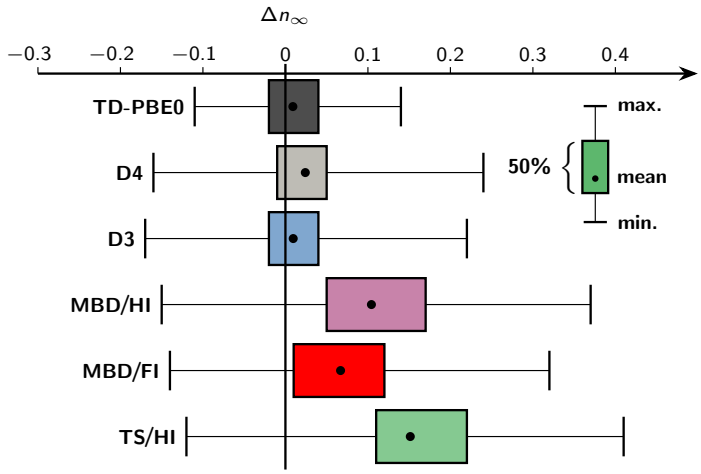

Fig. 4 Error ranges (relative to experimental data) for calculated refractive indices at different theoretical levels of theory (D4, D3, MBD/HI, MBD/FI, and TS/HI) for an organic polymer database.

tional many-body dispersion effects within the MBD/HI method.

3.2 Molecular crystals

3.2.1 Lattice energies

Molecular crystals are very relevant for material science as well as in pharmaceutical chemistry.[52,53] An important property is the lattice energy E_{lat}, which reflects how much energy is released per molecule upon resublimation. It is defined as

$$E_{\text{lat}} = \frac{1}{Z} E_{\text{crystal}} - E_{\text{gas}}, \tag{5}$$

where E_{crystal} is the energy of the crystal including overall Z molecules within the primitive cell and E_{gas} is the energy of the isolated molecule in the lowest energy conformation.

Recently, a set of eight highly accurate reference energies for molecular crystals have been published based on diffusion Monte-Carlo (DMC) calculations.[51] These reference systems include various binding motifs like, *e. g.*, hydrogen bonding in ice, electrostatic interactions in CO_2, as well as London-dispersion dominated unsaturated hydrocarbons. The statistical deviation (RMSD) of these high-level results from experimental data is small (0.004 eV). To provide a framework for the following discussion of the DFT results, it is useful to consider previous applications of this data set, which has been used to benchmark RPA and MP2.[54–56] With a computational cost between the DMC benchmark calculations and DFT-based methods, RPA and MP2 afford RMSDs of 0.08 eV and 0.06 eV, respectively. Moreover, by systematically studying the deviations from the reference, it was possible to devise an efficient GW-type singles correction to RPA (GWSE), which provides a much improved RMSD of only 0.01 eV.

In the following, we will explore the performance of DFT-D4 and D3(BJ) in combination with various functionals (PW91,[57] SCAN,[58,59] PBE,[60] rPBE,[61], revPBE,[62] TPSS,[63] B3LYP,[64,65] and PBE0[66]), and moreover compare D4 to other dispersion-corrected DFT methods. For the sake of completeness, we include almost all dispersion-corrected method implemented in the latest version of the prominent VASP program package, namely the methods of Tkatchenko/Scheffler (TS and MBD with fixed and iterative atomic charges)[10–13,13,14,14,43], Landgreth/Lundquist (original vdW-optPBE[15–20] and revised rev-vdW-DF2)[67] and Vydrov/Van Voorhis (revised VV10)[21,22].

We begin the discussion with the performance of D4 and D3(BJ) in combination with different functionals, which is depicted in 5a. The perhaps most impressive result of this survey is the dramatic improvement from the uncorrected DFT level to the results including D3 and D4 (*cf.* number given in parenthesis in the x-axis labels *vs* the lower purple and grey bars). In almost all cases, the error without any D is well above 0.5 eV, whereas the worst performance including the dispersion correction is 0.1 eV (revPBE-D3). Only the SCAN functional without D achieves an RMSD below 0.3 eV, indicating that it includes medium-range correlation effects to some extend, which is in agreement with previous reports.[58,59,68] The best over-all performance is provided by PBE0-D4 with an RMSD of 0.02 eV, which is well below that of much more demanding MP2 and RPA approaches and that of RPA(GWSE). Considering also the MDs of 0.08 eV, 0.06 eV and 0.01 eV for RPA, MP2 and PBE0-D4 shows that the error of PBE0-D4 is less systematic. The computationally even more efficient (no exact exchange) rPBE-D3(BJ) approach is almost as accurate as PBE0-D4, and closely followed by TPSS-D4, B3LYP-D3 and revPBE-D3, which all provide an RMSD ≤ 0.04 eV in the range of chemical accuracy (about 1 kcal mol^{-1}). For perspective, these DFT calculations take only hours (hybrid DFAs) to minutes (GGAs) on a reasonably modern compute node with 16 CPUs.

Comparing D3(BJ) and D4, the agreement with the reference improves significantly in some cases for D4 (*cf.*, revPBE and B3LYP), while for most DFAs it is only slightly more accurate, and in some cases slightly worse (*cf.*, PW91 and rPBE).

The results shown in figure 5b for the two-body models TS and TS/HI indicate a systematic under-binding. In particular for the crystals of the aromatics, one of the most common structural motif in organic chemistry, the deviations are as large as 50% of the total lattice energy. This is cured by the inclusion many-body effects in the MBD approaches, which afford an accuracy comparable to that of D3(BJ) and D4. Note, that D3(BJ) as employed here achieves this good result without including many-body effects. While the vdW-optPBE functional based on the original approach is not able to accurately describe the different interaction motifs, the revised variant rev-vdW-DF2 does a much better job, approaching the accuracy of PBE-D4. SCAN-D4 and SCAN-rVV10 perform very similar. Although they are slightly worse than PBE-D4 regarding their RMSD, the deviation within a given chemical compound class as indicated by similar colors in figure 5 is consistent. As a result, they provide by far the best relative energies of the ice (blue) and NH_3 forms.

In conclusion, these results suggest PBE0-D4 as the most accurate approach for studies on molecular crystals. At the computationally much more efficient (meta)GGA level, rPBE-D3 and TPSS-D4 are the next best options, providing only slightly worse lattice energies at much lower computational cost. Rev-vdW-DF2 and the MBD based approaches provide very similar accuracy. However, for MBD/HI and MBD/FI the computation

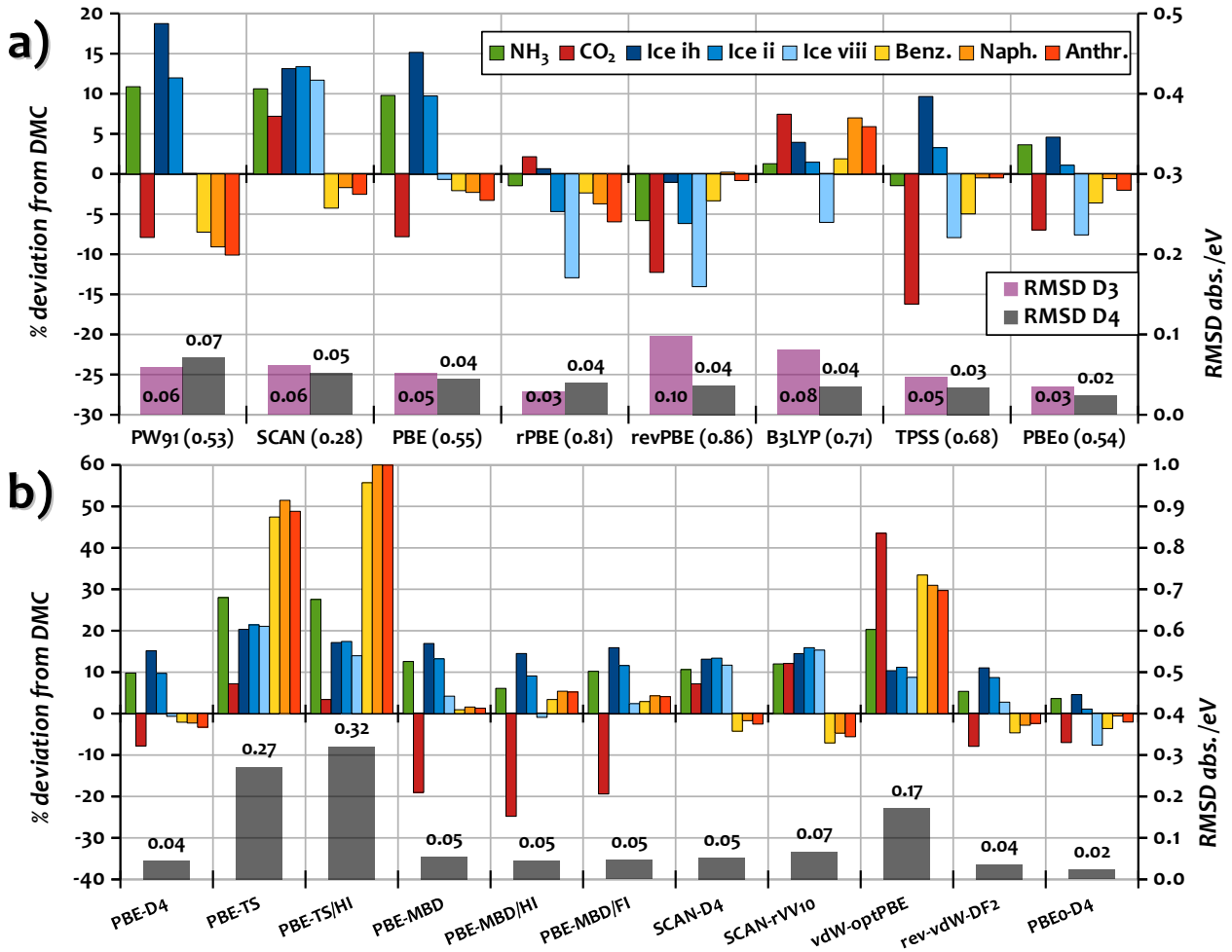

Fig. 5 a) Benchmark of DFT-D4 with several prominent DFs for the computation of lattice energies of eight molecular crystals (subset of the ICE10[48] and X23[49,50] benchmark set), for which high-level fixed node diffusion Monte-Carlo (FN-DMC) references are available.[51]. The RMSD of the uncorrected DFs is given in parenthesis (x-axis). **b)** PBE-D4 compared to other dispersion corrected approaches, including the original two-body TS correction,[10] TS/HI with iterative Hirschfeld charges,[13,14] TS/MBD with many-body effects,[11,12] the combination of MBD with HI,[13,14], as well as yet another revision that includes ionic references (MBD/FI).[43] Also shown are results obtained with SCAN-D4 compared to SCAN-rVV10,[21,22]), as well as with the original variant of vdW-DF,[15–20] and a more recent revised variant. The colored bars depict the relative (%) deviation from the reference, whereas the transparent bars at the bottom show the absolute root-mean square deviation (RMSD) from the reference including its numerical value. All calculations were conducted with VASP and employ the hardest PAW-PBE (GW) potentials available in the VASP library, an energy cut-off of 1000 eV and fine k-spacing. See Supplementary Material for more details.

of the dispersion-correction takes much longer than the GGA calculation itself (*cf.* timings below), limiting their applicability. If the variation in investigated chemical structures is small and the relative energies are more important than absolute energies, SCAN-D4 and similarly also SCAN-rVV10 can be recommended. However, due to the increased cost of the SCAN functional, also these approaches are computationally more demanding than GGA+D3/D4.

3.2.2 Geometries from experimental measurements

The packing of organic molecules into crystals is highly sensitive to an accurate treatment of NCIs and dispersion effects in particular. In 2012, Johnson and co-workers compiled a set of organic crystals from available low temperature X-ray structures (termed C21 benchmark[49]) which was refined and extended by Reilly and Tkatchenko resulting in the X23 benchmark set.[50] Various groups already used this benchmark set to test their electronic structure methods.[69,70] In their work Johnson *et al.* explicitly applied an artificial pressure to include volume expansion due to vibrational effects in the optimization procedure. However, as already discussed in earlier works[46] the zero point vibrational energy (ZPVE) and thermal contributions of the unit cell volume can be estimated and used to transform the experimental volume V_0 into back-corrected reference equilibrium volumes V_e which are directly comparable to optimizations on the electronic energy surface.

Here we have calculated unit cell volumes using PBE-D3(BJ), PBE-D4, PBE-TS, PBE-TS/HI, PBE-MBD, PBE-MBD/HI, PBE-MBD/FI, and pure PBE for the X23 systems. Figure 6 shows deviations from reference volumes per molecule for the 23 crystal structures. Uncorrected PBE only achieves a poor description by generally overestimating the cell volumes for these molecu-

lar crystals (black bar in figure 6, MD = 14.7, MAD = 14.7, and RMSD = 5.2 all given in %). Coupling PBE to any dispersion correction improves the description significantly but the respective methods perform differently well. PBE-TS performs worst which, however, can be improved by using an iterative Hirshfeld partitioning. Including many-body dispersion effects or adding an iterative Hirshfeld partitioning further improves the accuracy. Additional reference systems with fractional ionic character do not further increase the quality of the MBD results which is expected, since the considered molecular crystals have almost no ionic character. PBE-D3(BJ) volumes are comparable to those calculated with the TS/HI method. For the 23 tested molecular crystals, MBD gives a more accurate description compared to D3(BJ). However, PBE-D4 also improves upon PBE-D3(BJ) and furthermore is *on par* with the computationally more demanding MBD method in terms of all statistical evaluations (MD , MAD, and RMSD).

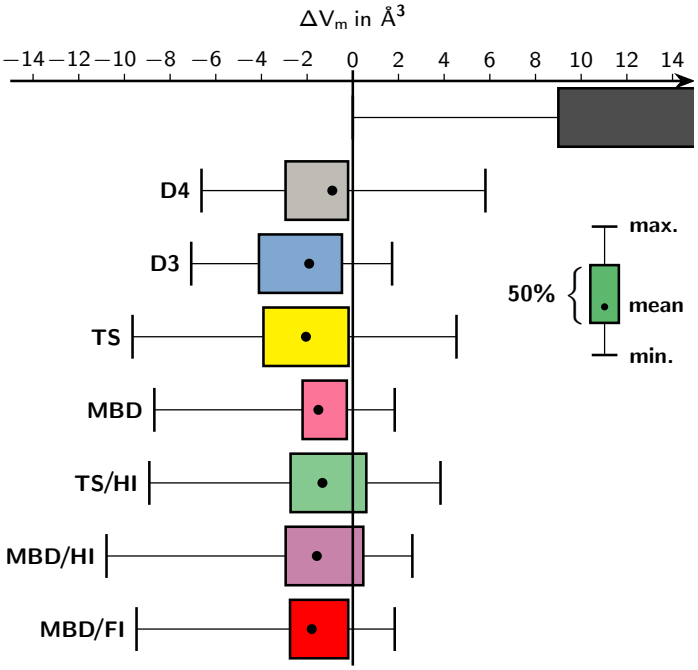

Fig. 6 Error ranges (relative to experimental data) for calculated cell volumes divided by the number of molecules inside the cell for 23 molecular crystal structures (X23 benchmark set). Shown are data for pure PBE (black) and PBE coupled to D4, D3(BJ), TS, MBD, TS/HI, MBD/HI, and MBD/FI. For all PAW calculations an 800 eV plane-wave energy cutoff is applied to minimize Pulay stress.

As another test, the D4 and MBD models are tested for their capability of reproducing strong hydrogen networks within different ice polymorphs. This important binding motif is underrepresented in the X23 benchmark set, so that the ICE10 benchmark[48] set is to be used for this purpose. A detailed description of crystallized water requires an accurate description of NCIs especially for reproducing structural properties. Overall, we investigate eight experimentally studied ice polymorphs. The measurements have been conducted at low temperatures (up to 100 K) where the thermal contribution to the ice density has been assumed to be rather small assuming that measured structures can be treated as equilibrium structures.[48] In the original ICE10 publication the ZPVE is estimated on each unit cell volume by correcting for thermal effects using the HF-3c[72] method applied with an ATM many-body dispersion treatment. In the present work we apply the quasi-harmonic approximation (QHA), see Supplementary Material for details) to eight out of ten ICE10 polymorphs to obtain back-corrected equilibrium volumes that are listed in table 2. For this purpose HSE-3c[71] calculations are performed within CRYSTAL17[73] to revise parts of this benchmark set. This composite DFT method has already proven its accuracy with respect to the reproducibility of volumes of molecular crystals[71] and is ideally suited for such computational intensive calculations.[74]

The direct comparison of PBE-D4 and PBE-MBD shows that both methods are well suited to describe equilibrium volumes of ice polymorphs. For the IX system, both methods calculate volumes that are too small, while for the VI and III systems D4 yields a too large volume while MBD does the opposite. The only "outlier" can be seen in the hexagonal ice Ih, where both methods drastically underestimate the equilibrium volume. The effect of the underlying density functional should be considered as well, since PBE is known to overbind hydrogen bonds[75]. For the Ih polymorph earlier studies already showed that PBE drastically underestimates the equilibrium volume[76] which could be corrected by explicitly including many-body correlation effects in terms of diagrammatic perturbation theory. However, this is beyond the scope of this work and will be neglected. The data of table 2 show that PBE-D4 is suitable to generate accurate volumes for ice polymorphs (MD = 0.3, MAD = 2.0, RMSD = 2.7 all given in %) when comparing to MBD corrected PBE (MD = −0.3, MAD = 2.3, RMSD = 3.4 all given in %). In general, pure PBE drastically overestimates all equilibrium volumes (PBE: MD = 5.6, MAD = 6.3, RMSD = 3.6 all given in %) indicating that use of an appropriate London dispersion correction is indispensable even in systems that are dominated by hydrogen bonding.

3.3 Adsorption on surfaces

To evaluate the performance of the D4 and related approaches for the calculation of adsorption energies, three chemically diverse model systems dominated by NCIs are considered.[77] These include the adsorption of non-polar benzene on gold (sparse), polar CO on polar MgO (saturated), as well as non-polar C_2H_2 on ionic NaCl (also saturated). For all of these, accurate experimental and/or high-level computational references are available. The respective adsorption models are shown in figure 7. In addition to D4, we consider its predecessors D3(BJ), D2, as well as other established methods, namely MBD (including variants), TS (including variants) and the vdw-DF2 and rVV10 dispersion functionals. All D and TS/MBD calculations are conducted in combination with the PBE functional.

The adsorption energies are calculated using

$$\Delta E_{\text{ads}} = \left(E_{\text{total}} - E_{\text{slab}} - nE_{\text{molecule}}\right)/n, \tag{6}$$

where n is the number of adsorbed molecules per unit cell, E_{slab} the energy of the vacant surface, E_{total} the energy of the adsorbed

Table 2 Refinement of the ICE10 benchmark set[48] (excluding polymorph II and XIII) with a correction of equilibrium volumes V_e to free-energy volumes V_0 due to ZPVE and thermal energies calculated at the HSE-3c level of theory[71] using a quasi-harmonic approximation. Structures and experimental temperatures are taken from Ref. 48. For all PAW calculations an 800 eV plane-wave energy cutoff is applied to minimize Pulay stress. All volumes are given in Å^3.

	HSE-3c			Experiment[a]		Δ V / %		
Polymorph	V_0	V_e	$\Delta V/V_0$ (%)	V_0	V_e^{ref}	PBE	-D4	-MBD
Ih	29.19	28.11	3.7	32.05	30.86	−3.0	−5.8	−6.5
III	25.11	23.69	5.7	25.69	24.23	7.3	0.9	−0.5
VI	23.36	21.44	8.2	22.84	20.97	5.0	0.5	−0.7
VII	23.03	21.32	7.4	20.26	18.76	7.4	1.8	0.8
VIII	23.03	21.38	7.2	20.09	18.64	8.2	2.7	2.2
IX	24.88	23.76	4.5	25.63	24.48	5.8	−1.0	−2.8
XIV	23.65	21.75	8.0	23.12	21.27	6.4	1.4	0.1
XV	23.45	21.60	7.9	22.45	20.68	7.2	1.8	5.0

[a] Experimental V_e estimated as $V_e^{ref} = V_0^{ref}\left(1 + \frac{V_e^{\text{HSE-3c}} - V_0^{\text{HSE-3c}}}{V_0^{\text{HSE-3c}}}\right)$.

molecule on the surface, and E_{molecule} the energy of the isolated molecule. To circumvent computationally expensive structure optimizations at each level, and yet avoid a bias by using one of the tested dispersion-correction schemes to conduct the optimizations, all calculations refer to single-point energies (cut-off 800 eV) on geometries obtained at the plain DFT/SCAN level (cut-off 500 eV). In these optimizations, the molecules as well as the first (complete) layer of the surfaces is relaxed, while the lower layers were kept fixed at default values provided by the atomic simulation environment (ASE).[78] SCAN has been chosen for this purpose because it provides reasonable agreement for all systems without any dispersion correction. In general, the impact of the geometric relaxation (with respect to ASE default parameters) on the binding energy has been investigated at the PBE-D3(BJ) and SCAN levels of theory and found to be small with about 0.01–0.02 eV for the polar surfaces, and 0.05 eV for benzene on gold.

Figure 7 displays the results of these calculations in the form of deviations from the experimental reference values. These data, the choice of the experimental references, as well as the details of the systems under considerations are discussed below. It should be pointed out that there is a systematic bias when comparing calculated adsorption energies to experimental enthalpy values as used here. Since the calculations do not account for zero-point and thermal contributions, both of which will reduce the interaction, a slight overestimation of the theoretical adsorption energy compared to the reference values is expected (and preferable to an underestimation).

Benzene on gold—The adsorption of aromatic molecules on transition metal surfaces is of particular interest in materials science because the catalytic conversion of aromatic substances is a key reaction in many petrochemical processes.[79] Our model is based on the common adsorption mode with benzene flat on the (111) surface of *fcc* Au as is shown in Fig. 7, and corresponds to coverage of $\theta = 1/25$.[80] Since gold provides one of the most inert surfaces, the adsorption is almost exclusive driven by London dispersion. The experimental estimate of the adsorption energy is −0.64 eV at a sub-monolayer coverage.[81] Plain PBE provides a qualitatively wrong, repulsive energy of (0.49 eV). While the old D2 model clearly over-corrects this failure of PBE with an adsorption energy of −1.4 eV, the newer D3(BJ) model provides a much improved energy of −0.76 eV, which can be considered in agreement with the experiment given the absence of zero-point and thermo-chemical corrections in the calculated values. The D4 model presented here is even closer to the experimental result with −0.69 eV. The original TS model performs very similar to D4 (−0.70 eV) irrespective if fixed or iterative Hirshfeld charges are used. Including many-body effects via the MBD approach significantly worsens the agreement, providing too small adsorption energies (−0.48 eV), which is surprising in face of the much better agreement of the MBD approach for the benzene crystal (*cf.* figure 5). Also for the MBD approach, the agreement is almost independent of the charge scheme. Only the fractional-ionic scheme (FI) leads to a slight improvement, which is surprising in face of the non-polar nature of this system. Plain SCAN finds, in contrast to PBE, an attractive interaction between adsorbate and surface. However, with −0.50 eV it is slightly below the experimental value. This confirms that SCAN describes parts of medium-range correlation effects intrinsically. However, as already seen for the molecular crystals, this can lead to double counting in combination with dispersion corrections. Accordingly, both SCAN-rVV10 and SCAN-D4 slightly over-bind the system with adsorption energies of −0.83 eV and −0.85 eV, respectively. Also the rev-vdw-DF2 functional correctly recovers the (attractive) interaction between adsorbate and surface, but slightly underestimates it with an adsorption energy of −0.56 eV.

Carbon monoxide on magnesium oxide—The adsorption of carbon monoxide on MgO(001) surfaces has been extensively studied both experimentally[82–84] and theoretically.[85–88] Experimental studies provided adsorption energies ranging from about −0.13 eV[83,89] up to −0.20 eV,[90] and showed that CO adsorbs

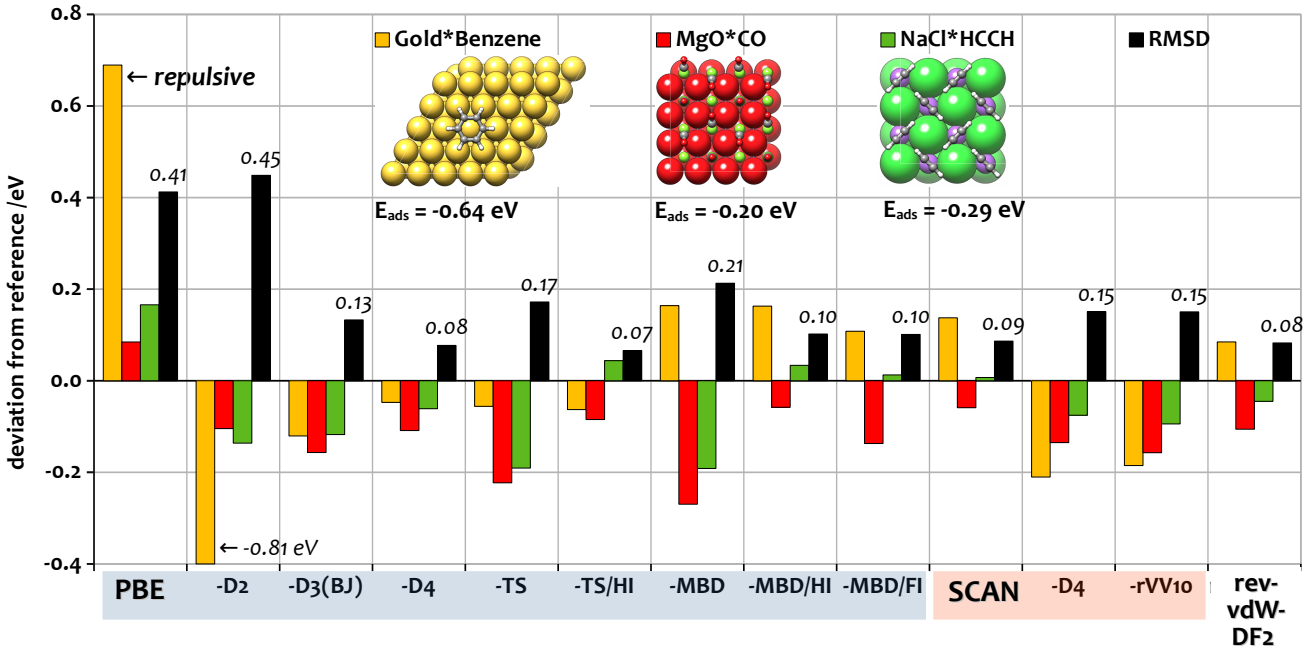

Fig. 7 Benchmark of various DFT approaches for the adsorption of small organic molecules on metallic and ionic surfaces.

in a C-down configuration on top of five-fold coordinated Mg atoms.[83] This is reflected in our model system with a coverage of 75%, which has been taken from Ref. 91 and is shown in figure 7. While previous DFT-based studies predict a wide range of adsorption energies from −0.10 eV to −0.56 eV,[86] coupled-cluster based approaches provide values in better agreement with the experiment of −0.15 eV[88] to−0.17 eV[92]. Considering the absence of zero-point and thermo-chemical corrections in our calculations, we selected the lowest experimental value of −0.20 eV to be used as reference. Already plain PBE is with a computed adsorption energy of −0.12 eV in qualitative agreement with the reference value. As noted previously, the D2 model over-binds CO with an adsorption energy of −0.30 eV. Surprisingly, this becomes even worse with D3(BJ) (−0.36 eV), and even D4 does not improves the result with −0.31 eV. In fact, all employed methods except plain PBE predict much too large adsorption energies for this system, which range from the worst value of −0.47 eV (MBD) to the best of −0.26 eV (SCAN, MBD/HI). This may be seen as a hint towards a more fundamental problem of DFT for this system, or a problem with the model system. In face of this general over-binding of dispersion-corrected DFT for this system, it is not surprising that SCAN and MBD/HI provide the best agreement, as these are also the approaches that provide the over-all smallest adsorption energies, and accordingly the only ones besides plain PBE with a positive mean average deviation.

Acetylene on sodium chloride—The adsorption of acetylene on sodium chloride has been studied by several groups experimentally and theoretically. Experimental adsorption energies[93,94] range from −0.25 to −0.31 eV corresponding to full and half coverage, respectively. Previous studies employing periodic DFT provide values ranging from −0.32 eV to −0.44 eV.[91] We employ the same model system, which corresponds to full coverage with the C_2H_2 molecules ordered in a T-shaped formation (*cf.* figure 7). We use a reference value of −0.29 eV, which is between the experimental estimates for full and half coverage. Plain PBE affords an attractive yet too small interaction between adsorbate and surface, as evident from the adsorption energy of −0.12 eV. Again, D2 significantly over-corrects to an adsorption energy of (−0.42 eV). This improves slightly with D3(BJ) providing −0.40 eV, and more significantly with the latest D4 approach to −0.35 eV. Most certainly, this improvement is related to the superior ionic polarizabilities of the D4 approach compared to D3(BJ). Also the results of the other methods demonstrate the necessity for accurate ionic polarizabilities. While the TS and MBD approaches with their fixed atomic charges both over-stabilize this system even more than D2 with an adsorption energies of −0,48 eV, the iterative schemes TS/HI and MBD/HI yield much better agreement with −0.24 eV and −0.25 eV. Adding ionic references in the MBD/FI schemes further improves this agreement, leading to an adsorption energy closer to both D4 and the reference with −0.27 eV. Also plain SCAN provides a good description, yielding an adsorption energy of −0.28 eV, which increases slightly when the functional is combined with a dispersion correction (SCAN-D4 −0.36 eV, SCAN-rVV10 −0.38 eV). Also the rev-vdW-DF2 functional performs well with a value of −0.33 eV.

In particular this last system nicely demonstrates the necessity to include charge and/or density information in the dispersion correction, as evident from the superiority of D4 over D3(BJ), as well as the charge-iterative Hirshfeld schemes over the standard models. Regarding the over-all statistical agreement, the data indicate that PBE-D4, PBE-TS/HI, SCAN and rev-vdW-DF2 are the most accurate approaches with an RMSD below 0.1 eV. Of these, in particular PBE-D4 stands out with the most consistent deviation in the form of a slight over-binding of the adsorbates.

As pointed out earlier, this is preferable to an under-binding and would be corrected by including zero-point and thermo-chemical contributions in the calculations. The perhaps most surprising result from this benchmark is the excellent performance of the uncorrected SCAN functional. However, this also comes at an increased computational cost, as will be discussed in detail in the next section.

3.4 Timings

In this section, we compare the computational cost of different dispersion-corrected DFT methods using the cyclohexanedione crystal structure as a test case. For this purpose, single-point energy calculations are conducted with VASP with settings for high numerical precision (cut-off of 800 eV, fine k-point grid) on a 16 core Intel(R) Xeon(R) CPU (E5-2660 v4 @ 2.00 GHz). With this set-up, the DFT/PBE single-point takes about 7.5 minutes walltime. Figure 8 provides the total walltime of the calculation, as well as (if possible) the timing for the dispersion-correction only, which is determined as the difference between the DFT/PBE calculation with and without the respective correction.

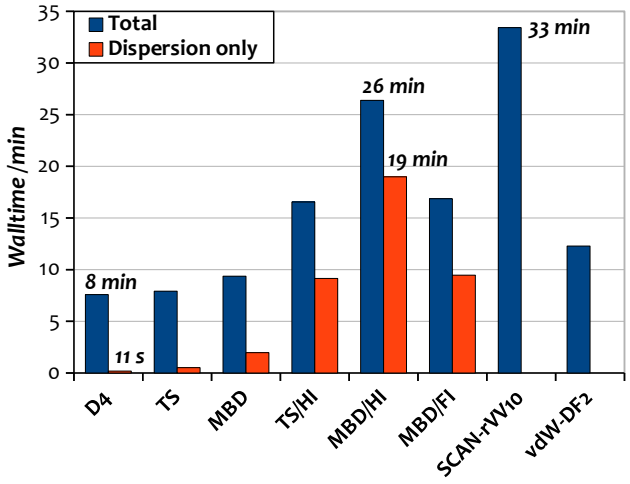

Fig. 8 Total walltime of a single-point calculation for the cyclohexanedione crystal with DFT/PBE and various dispersion-correction schemes, SCAN-rVV10 and rev-vdW-DF2 (blue bars). Time spent on the dispersion correction alone is shown in orange (not available for the self-consistent approaches).

Regarding the total walltime, DFT-D4, -TS and -MBD (the latter with fixed charges) are comparable with timings just under 10 minutes, followed by rev-vdW-DF2 just above 10 minutes. The charge-iterative schemes are significantly more expensive taking just above 15 (TS/HI, MBD/FI) up to almost 30 minutes (MBD/HI). Only the SCAN-rVV10 calculation takes even longer with 33 minutes, but this is mostly due to the increased computational cost of the SCAN functional.

Removing the overhead of the DFT calculation and focusing only on the timing of the dispersion-correction, D4 turns out as the fastest method by far with 11 s, followed by fixed-charge TS and MBD with 31 s and 118 s, respectively. All of these are in stark contrast to the charge-iterative schemes, which are orders of magnitude slower taking as long (9 min for TS/HI and MBD/FI), or twice as long (19 min for MBD/HI) as the underlying DFT/PBE calculation. As a result, the dispersion-correction becomes the bottleneck of the calculation, in spite of the high numerical precision settings for the DFT calculation. In combination with semi-empirical and/or tight-binding approaches, or just more sloppy DFT settings and a good initial guess – as is the typical scenario for DFT molecular-dynamics simulations – the charge-iterative schemes are impractical and only the fastest schemes (D4, TS) remain useful.

4 Summary and conclusions

We presented an extension of the molecular DFT-D4 model for describing dense condensed-phase systems under periodic boundary conditions. For this purpose, new periodic reference polarizabilities for alkaline, for alkaline-earth, and for early d-metals (group 3-5) were calculated using a pseudo-periodic electrostatic embedded cluster model.

The utility of the additional reference polarizabilities was demonstrated by calculating solid state polarizabilities of different salts and comparing them to experimental data. Here, the new references lead to a dramatic improvement for cationic and anionic polarizabilities, which now exhibit the correct ordering compared to the experiment and also show better quantitative agreement. For the same test set, related methods like MBD, MBD/HI and even MBD/FI exhibited much larger deviations. This is despite the explicit inclusion of ionic reference systems in the MBD/FI approach, which improves the description compared to MBH/HI, but does not approach the accuracy of the D4 model.

Eventually, the performance of the periodic DFT-D4 model was tested and compared to its predecessors and related approaches in three real-world application scenarios. These include (i) the calculation of lattice energies of molecular crystals, (ii) the determination of cell volumes of molecular crystals, as well as (iii) the modeling of the adsorption of chemically diverse substrates on non-polar, polar and ionic surfaces. For lattice energies of molecular crystals, various D3 and D4 corrected DFAs showed excellent agreement to high-level references. Especially “repulsive” density functionals (*e. g.,*, revPBE, TPSS, and PBE0) turned out to provide an accurate account for NCIs, whereas density functionals that already include mid-range correlation to some extent (SCAN) appear to suffer from possible double-counting issues. Other approaches showed to be competitive to D4 for describing such lattice energies like MBD/HI corrected DFAs or the rev-vdw-DF2 functional. The quality of geometries was investigated using cell volumes of 23 molecular crystals and eight different ice polymorphs. For those systems, DFT-D4 yielded accurate cell volumes with a deviation very similar to related approaches. In general, all tested approaches show a tendency to calculate slightly too small volumes. Regarding the calculation of adsorption energies for organic molecules, a direct comparison between DFT-D4 and electron density dependent (TS;TS/HI, MBD;MBD/HI;MBD/FI, vdw-DF2, rVV10) and independent (DFT-D3(BJ) and DFT-D2) models showed DFT-D4, TS/HI and rev-vdW-DF2 to provide the best agreement, followed by SCAN (without any vdW correction) and the iterative MBD schemes. Of these, D4 stands out as the approach provides the most systematic deviation in the form of a

slight over binding. Ultimately, an analysis of computational timings shows that DFT-D4 is orders of magnitude faster than the iterative MBD schemes which offer similar accuracy, and thus provides the best balance between cost and accuracy. With the development of the periodic DFT-D4 model it is routinely possible to describe chemically diverse molecular, periodic solids or surfaces accurately with a low computational effort.

5 Acknowledgments

EC thanks T. Bredow for valuable discussions. Furthermore, the authors thank J. G. Brandenburg for sharing molecular crystal structures of the DMC8 benchmark set and T. Bučko for help with the MBD calculations.

6 SUPPLEMENTARY MATERIAL

6.1 Periodic DFT-D4 methologies

Let us begin with a review of the theoretical framework of the D4 approach[27] as this will provide the context for the introduction of the new features. The underlying concept of D4 and its predecessors is to model the dispersion energy based on atomic pairwise dispersion coefficients C_6^{jk}, which are obtained from a Casimir–Polder integration of the respective atomic polarizabilities $\alpha^{\mathrm{eff}}(i\omega)$

$$C_6^{jk} = \frac{3}{\pi}\int_0^\infty d\omega\, \alpha_j^{\mathrm{eff}}(i\omega)\alpha_k^{\mathrm{eff}}(i\omega). \tag{7}$$

Note that atomic units are used throughout in this work. To account for the influence of the chemical environment, the atomic polarizabilities used in the D4 model are not fixed at the values of the isolated atoms $\alpha(i\omega)$, but depend (i) on geometric parameters captured by the atomic coordination number (CN_j, as in D3), as well as (ii) on effective atomic charges (z_j, new in D4) obtained *via* an electronegativity-equilibration (EEQ) scheme. The idea behind the introduction of this charge scaling is to allow for a more "natural" behavior of the effective polarizabilities, *i.e.*, to render atoms with negative partial charge more polarizable and *vice versa*.

In practice, the geometry and charge dependence of the polarizabilities is implemented using an interpolation based on a set of molecular reference systems. To do this efficiently and avoid an interpolation in two dimensions, the atomic polarizabilities of all reference systems are, in a first step, rescaled to match the effective charge of the atom under consideration of z_j *via*

$$\overline{\alpha}_i^{\mathrm{ref}}(i\omega) = \alpha_i^{\mathrm{ref}}(i\omega)\zeta(z_j, z_i^{\mathrm{ref}}), \tag{8}$$

where ζ describes an empirical relation between the polarizability of an atom and its effective charge. Its analytical form

$$\zeta(z_i, z_i^{\mathrm{ref}}) = \exp\left(\beta\left\{1-\exp\left[\gamma_i\left(1-\frac{z_i^{\mathrm{ref}}}{z_i}\right)\right]\right\}\right), \tag{9}$$

with β as a global parameter set to 3 and γ_i as the chemical hardness taken from Ref. 95 is discussed in more detail in Ref. 27. The calculation of the necessary effective charges z done with an EEQ model is described in section. 6.2

In a second step, the effective polarizability of the atom j is obtained *via* interpolation from the charge-scaled reference polarizabilities using a Gaussian weighting based on the coordination number (CN)

$$\alpha_j^{\mathrm{eff}}(i\omega) = \sum_{i,\mathrm{ref}=1}^{N^{i,\mathrm{ref}}} \overline{\alpha}_i^{\mathrm{ref}}(i\omega)\, W_j^{i,\mathrm{ref}}(\mathrm{CN}_i^{i,\mathrm{ref}}, \mathrm{CN}_j). \tag{10}$$

However, to use this CN-based approach in periodic systems, a different formula for the CN is used compared to the molecular implementation to avoid CN-divergences. The expression for the CN in periodic systems reads

$$\mathrm{CN}_i = \sum_{\mathbf{T}}\sum_j{}' \frac{\delta_{ij}^{\mathrm{EN}}}{2}\left(1+\mathrm{erf}\left(-k_0\left(\frac{|\mathbf{R}_{ij}+\mathbf{T}|-R_{ij}^{\mathrm{cov}}}{R_{ij}^{\mathrm{cov}}}\right)\right)\right) \tag{11}$$

$$\delta_{ij}^{\mathrm{EN}} = \left(k_1 \exp\left(|\mathrm{EN}_i - \mathrm{EN}_j| + k_2\right)^2\right)/k_3\,,$$

where, $\mathbf{T} = t_1\mathbf{a_1} + t_2\mathbf{a_2} + t_3\mathbf{a_3}$ denotes the translation vector with $\mathbf{a_1}$, $\mathbf{a_2}$, and $\mathbf{a_3}$ being the lattice vectors (t_1, t_2, and $t_3 \in \mathbb{Z}$). The primed sum over j indicates that the case $i = j$ is omitted for $\mathbf{T} = \mathbf{0}$. One of the central changes is the use of Pauling electronegativities (EN),[96] as well as the inter-nuclear distance R_{ij} of the pair ij, and the covalent atomic radii[97] ($R_{ij}^{\mathrm{cov}} = R_i^{\mathrm{cov}} + R_j^{\mathrm{cov}}$). Note that the CN has become EN-dependent to differentiate between covalent and ionic bonding (e.g., differentiate F_2 from HF). The parameters in equation 11 ($k_0 = 7.5$, $k_1 = 4.1$, $k_2 = 19.09$, and $k_3 = 254.56$) were taken from Ref. 27.

This charge and geometry dependent calculation of atomic polarizabilities from molecular reference systems, which may be described as an atom-in-molecule approach to polarizabilities, presumes the additivity of atomic polarizabilities,[98] which is reflected in the following equation

$$\alpha_i^{\mathrm{ref}}(i\omega) = \frac{1}{m}\left[\alpha^{I_m X_n}(i\omega) - \frac{n}{l}\,\alpha^{X_l}(i\omega)\zeta(z_X, z_X^{\mathrm{ref}})\right]. \tag{12}$$

Here, $\alpha^{I_m X_n}(i\omega)$ is the molecular polarizability of one I-reference, $\alpha^{X_l}(i\omega)$ refers to the homonuclear compound (e.g., $\alpha^{H_2}(i\omega)$ as dihydrogen) and m, n, and l are the particular stochiometric coefficients. All I atoms inside the reference molecules and the X atoms in the homonuclear compounds are electronically equal and thus symmetry equivalent. By exploiting this symmetry equivalence, the approximation of additive polarizabilities is justified. Furthermore, the charge scaling of all X atoms in the respective reference system is directly incorporated.

This more general scheme has no disadvantaged compared to the hydrogenated reference systems used in the D3 model. With it, any diatomic molecular polarizabilities, e.g., dihalide molecular polarizabilities (chlorine or fluorine) and oxygen molecular polarizabilities can be used in the subtraction scheme of equation 12 as briefly discussed in Ref. 91. This generalization of the approach opens up the possibility to provide specialized C_6 dispersion coefficients, which will be exploited here to properly describe interactions in ionic solids.

6.2 Periodic electronegativity equilibration model

For the generation of atomic partial charges q under periodic boundary conditions, a classical geometry dependent EEQ charge model is developed in the present work. For this purpose, a cyclic cluster model (CCM) is implemented which applies periodic boundary conditions to a cluster that uses a non-primitive unit cell of a solid, a surface, or an infinite chain by directly employing cyclic Born-van-Kármán boundary conditions. The environment of each atom is replaced by a notional cyclic arrangement of cluster atoms, where the interaction zone of each atom within the cyclic cluster is described by a Wigner-Seitz cell, constructed by the translation vectors of the unit cell and centered at the atom. The cluster is constructed as a supercell of the primitive unit cell, so that a repetition of this unit cell of N_1, N_2, and N_3 cells along the lattice vectors $\mathbf{a_1}$, $\mathbf{a_2}$, and $\mathbf{a_3}$ leads to a total cell number of $N = N_1 \cdot N_2 \cdot N_3$. In the CCM, the WSCs are stoichiometrically and symmetrically constructed to ensure local electroneutrality. This is guaranteed by the fact that each WSC central atom i is surrounded by Ξ_i neighbours j with an inverted partial charge. When setting up the cluster, the number of all neighbours of each WSC central atom is determined and weighting factors $w_{ij} = 1/\Xi_i$ for the respective neighbours are assigned, accordingly. Since the CCM is a finite-size method, the clusters used can also carry a net charge without running into convergence problems occurring within the employed Ewald sums. In contrast to a supercell model, no summation over special $\mathbf{k}$-points has to be carried out. Instead, a discrete number of $\mathbf{k}$-points is contained implicitly by placing them equally distributed in space.

$$\{\mathbf{k}\} = \prod_j^3 \frac{g_j}{N_j}\mathbf{b_j} \text{ with } g_j = 0,\ldots,N_j^{-1} \tag{13}$$

Here, we introduce the reciprocal lattice vectors $\mathbf{b}$. The periodic charge density $\rho(r)$ of the system is supposed to be a superposition of spherically symmetric Gaussian functions centered at the atoms position, each normalized to the corresponding nuclear charge q_i given by the following expression

$$\rho_i(r) = \sum_{\mathbf{T}} \frac{q_i}{a_i^3\pi^{3/2}} \exp\left(-\frac{|\mathbf{r}-\mathbf{R}_i-\mathbf{T}|^2}{a_i^2}\right). \tag{14}$$

Here, the atomic van der Waals radii a_i are introduced. By choosing such atomic charge densities the total isotropic electrostatic (IES) energy is amenable by the following expression given in matrix notation

$$E_{\text{IES}} = \mathbf{q}^T\left(\frac{1}{2}\text{A}\cdot\mathbf{q} - \mathbf{X}\right). \tag{15}$$

The interaction matrix A contains all periodic Coulomb interactions, which are developed in Ewald sums by splitting the Coulomb operator into short-range and long-range contributions (Ewald splitting parameter $\xi = \sqrt{\pi}/V^{1/3}$). Here, the previously determined weighting factors w_{ij} (as obtained from the CCM) are applied for all off-diagonal elements

$$\begin{aligned}
A_{ij}^{\text{rec}} &= \frac{4\pi}{V}\sum_{\mathbf{k}\neq 0}\cos\left(\mathbf{k}\cdot(\mathbf{R}_{ij}+\mathbf{T})\right)\exp\left\{-\frac{k^2}{4\xi^2}\right\}\frac{w_{ij}}{k^2}\\
A_{ii}^{\text{rec}} &= \frac{4\pi}{V}\sum_{\mathbf{k}\neq 0}\exp\left\{-\frac{k^2}{4\xi^2}\right\}\frac{1}{k^2}\\
A_{ij}^{\text{dir}} &= \sum_{\mathbf{T}}\left(\frac{\text{erf}(\gamma_{ij}|\mathbf{R}_{ij}+\mathbf{T}|)}{|\mathbf{R}_{ij}+\mathbf{T}|} - \frac{\text{erf}(\xi|\mathbf{R}_{ij}+\mathbf{T}|)}{|\mathbf{R}_{ij}+\mathbf{T}|}\right)w_{ij}\\
A_{ii}^{\text{dir}} &= \sum_{\mathbf{T}\neq 0}\frac{\text{erf}(\gamma_{ii}|\mathbf{T}|)}{|\mathbf{T}|} - \frac{\text{erf}(\xi|\mathbf{T}|)}{|\mathbf{T}|}\\
A_{ii}^{\text{self/back}} &= J_{ii} + \frac{2\gamma_{ii}}{\sqrt{\pi}} - \frac{\pi}{\xi^2 V}.
\end{aligned} \tag{16}$$

Furthermore, we define γ_{ij} to be equal to $\left(a_i^2+a_j^2\right)^{-1/2}$. The Lagrangian is constructed under the constraint that the sum of the atomic charges conserves the total charge of the cluster, *i.e.*,

$$\begin{aligned}
L &= E_{\text{IES}} + \lambda\left(\sum_k q_k - q_{\text{cluster}}\right)\\
&\text{with}\quad \frac{\partial L}{\partial \mathbf{q}} = \mathbf{0} \;\wedge\; \frac{\partial L}{\partial \lambda} = \sum_i q_i - q_{\text{cluster}} = 0,
\end{aligned} \tag{17}$$

which leads to a set of $(N+1)$ linear equations. The right-hand side (RHS) of this set of equations is given by $X_i = -\chi_i$, where χ_i consists of the fitted atomic electronegativity EN_i which is shifted according to the following expression

$$\chi_i = \text{EN}_i - \Omega_i. \tag{18}$$

The molecular EEQ model uses for this shift the square root of a modified error function CN as described in Ref. 27. Since high coordination numbers can be reached very quickly in a periodic system, artificial polarity reversals can occur. An instructive example is displayed in figure 9, where the polarity between cations and anions within the sodium chloride crystal is reversed (*i.e.*, sodium formally becomes anionic and chlorine cationic). This is an artifact of the definition of the CN, which depends on the covalent radius R^{cov} of the respective atom. As a result, the atom with the larger covalent radius also gets a higher CN (here $R_{\text{Na}}^{\text{cov}} = 3.5\,\text{Bohr}$ and $R_{\text{Cl}}^{\text{cov}} = 2.5\,\text{Bohr}$) and thus a higher EN shift resulting in non-physical polarity changes. In order to avoid such artifacts, the procedure is modified for periodic boundary conditions. Here, Ω_i is used which includes the global parameter $\gamma = 8$ and is given as

$$\Omega_i = \kappa_i \log\left(\frac{1+\exp(\gamma)}{1+\exp(\gamma)-CN_i}\right). \tag{19}$$

This classical charge model requires overall five empirical parameters (J_{ii}, a_i, EN_i, κ_i, and R_i^{cov}) per element and achieves for molecules across the entire periodic table of elements an average deviation of about 0.04 e^- (0.03 e^- for organic molecules) with respect to PBE0 based Hirshfeld charges.[27]

By using the definition of the Lagrangian given in equation 17

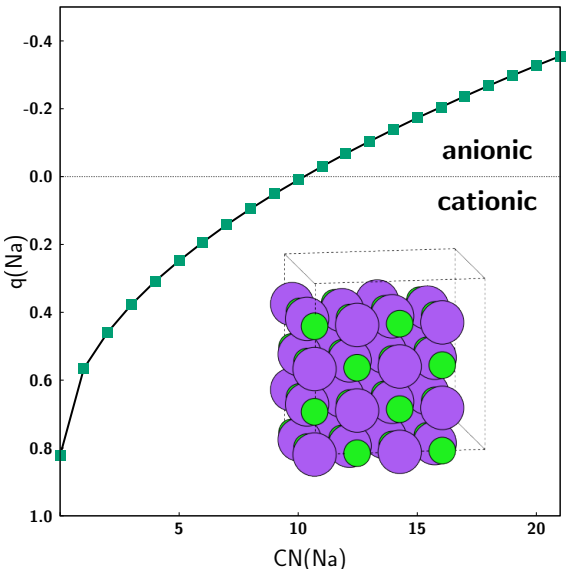

Fig. 9 Functional dependence on the sodium partial charge q(Na) with increasing CN(Na) using the molecular EEQ model within the sodium chloride crystal.

the analytical charge gradients is derived as

$$\frac{\partial \mathbf{q}}{\partial \mathbf{R}_j} = \widetilde{A}^{-1}\left[-\frac{\partial(\mathrm{A}\cdot\mathbf{q})}{\partial \mathbf{R}_j} + \frac{\partial \mathbf{X}}{\partial \mathbf{R}_j}\right] \tag{20}$$

where the inverse of the indefinite $(N+1)$ matrix has been obtained by a Bunch–Kaufman factorization[99] and inversion.

6.3 Dispersion energy

The periodic DFT-D4 energy expression is constructed as follows

$$E_{\mathrm{disp}}^{D4} = E_{\mathrm{disp}}^{(6,8)} + E_{\mathrm{disp}}^{(9),\mathrm{ATM}}. \tag{21}$$

The left part of equation 21 corresponds to the pairwise dispersion energy which is given by

$$E_{\mathrm{disp}}^{(6,8)} = -\frac{1}{2}\sum_i \sum_{\mathbf{T}} \sum_j{}' \sum_{n=6,8} s_n \frac{C_{(n)}^{ij}}{R_{ij_{\mathbf{T}}}^{(n)}} f_{\mathrm{damp}}^{(n)}\left(R_{ij_{\mathbf{T}}}\right). \tag{22}$$

Here, the primed sum over j indicates that the case $i=j$ is omitted for $\mathbf{T}=\mathbf{0}$. In equation 22, s_n scales the individual multipolar contributions—s_6 and s_8 for the dipole–dipole and dipole–quadrupole term—and $f_{\mathrm{damp}}^{(n)}$ denotes the rational Becke–Johnson (BJ) damping function (denoted as BJ-damping (BJD) in the following) which is used to couple this approach to standard DFAs.

$$f_{\mathrm{BJD}}^{(n)}(R_{ij_{\mathbf{T}}}) = \frac{R_{ij_{\mathbf{T}}}^{(n)}}{R_{ij_{\mathbf{T}}}^{(n)} + \left(a_1 R_0^{ij} + a_2\right)^{(n)}} \tag{23}$$

Equation 23 incorporates the DFA-specific parameters a_1 and a_2 and the cutoff-radii defined as

$$R_0^{ij} = \sqrt{\frac{C_8^{ij}}{C_6^{ij}}}, \tag{24}$$

where the recursive relation between dipole–dipole and dipole–quadrupole dispersion coefficients is used. Furthermore, we define the following expression for the rational damping term

$$R_{0,\mathrm{BJ}}^{ij} = \left(a_1 R_0^{ij} + a_2\right). \tag{25}$$

The simplest way to include three-body effects uses the well-known Axilrod–Teller–Muto[100,101] (ATM) term (*cf.*, right side of equation 21) which is defined as the sum over ijk energy contributions each defined by

$$E^{ijk} = \frac{C_9^{ijk}\left(3\cos\theta_i\cos\theta_j\cos\theta_k + 1\right)}{\left(R_{ij}R_{jk}R_{ki}\right)^3}. \tag{26}$$

Here, θ_i, θ_j, and θ_k are the internal angles of the triangle formed by R_{ij}, R_{jk}, and R_{ki} while C_9^{ijk} is the triple-dipole constant given by

$$C_9^{ijk} \approx \sqrt{C_6^{ij}C_6^{jk}C_6^{ki}}. \tag{27}$$

The C_9^{ijk} coefficients are derived from C_6 coefficients which are obtained from charge-neutral atomic polarizabilities (i.e., neutral atoms with $z_i = Z_i$). The finally used three-body dispersion energy expression is as follows

$$E_{\mathrm{disp}}^{(9),\mathrm{ATM}} = -k\sum_i \sum_{\mathbf{T}} \sum_j{}' \sum_{\mathbf{T}'}{}' \sum_k{}'' f_{damp}^{(9)}(\overline{R}_{ij_{\mathbf{T}}k_{\mathbf{T}'}}) E^{ij_{\mathbf{T}}k_{\mathbf{T}'}}, \tag{28}$$

where the sum is over all atom triples ijk applied with a zero-damping scheme proposed by Chai and Head-Gordon[102]

$$f_{damp}^{(9)}(\overline{R}_{ij_{\mathbf{T}}k_{\mathbf{T}'}}) = \frac{1}{1+6\left(\overline{R}_{ij_{\mathbf{T}}k_{\mathbf{T}'}}\right)^{-16}}. \tag{29}$$

To avoid multiple counting of three-body interactions the factor k is set to $\frac{1}{3}$ if atoms i, j, k are within the reference cell and to $\frac{1}{2}$ in all other cases. As previously, the primed sum over j indicates that the case $i=j$ is omitted for $\mathbf{T}=\mathbf{0}$, the double primed sum over k indicates that the case $j=k$ is omitted for $\mathbf{T}=\mathbf{T}'$ and $i=k$ is omitted for $\mathbf{T}=\mathbf{0}$. Equation 29 includes the averaged interatomic distance

$$\overline{R}_{ij_{\mathbf{T}}k_{\mathbf{T}'}} = \left(R_{ij_{\mathbf{T}}}R_{j_{\mathbf{T}}k_{\mathbf{T}'}}R_{k_{\mathbf{T}'}i} \Big/ R_{0,\mathrm{BJ}}^{ij}R_{0,\mathrm{BJ}}^{jk}R_{0,\mathrm{BJ}}^{ki}\right)^{1/3}, \tag{30}$$

which incorporates $R_{0,\mathrm{BJ}}^{ij/jk/ki}$ (*cf.* equation 25). Since dispersion interactions are much faster decaying (leading order term $\propto R^{-6}$) than, e.g., Coulomb interactions we employ for this energy contribution a real-space cutoff within the periodic implementation. Furthermore, analytical gradients are available for the dispersion energy expression in equation 21.

6.4 Theory of the Quasi-harmonic approximation

In order to account for thermodynamic properties to crystals, the knowledge of phonon modes is required over the complete first Brillouin zone (FBZ) of the system. The easiest way to account for ZPVE contributions includes harmonic lattice dynamics where

each **k**-point in the FBZ is associated with 3M harmonic oscillators (*i. e.* phonons) which are labeled by a phonon band index n ($n = 1,\ldots,3N$) and whose energy levels are given by the usual harmonic expression as

$$\varepsilon_m^{n,\mathbf{k}} = \left(m + \frac{1}{2}\right)\omega_{\mathbf{k}n}, \tag{31}$$

where m is an integer, $\omega_{\mathbf{k}n} = 2\pi\nu_{\mathbf{k}n}$, and N is the number of atoms per primitive cell. The overall vibrational canonical partition function of a crystal at a given temperature T is given as

$$Q_{vib}(T) = \prod_{\mathbf{k}}\prod_{n=1}^{3N}\sum_{m=0}^{\infty}\exp\left(-\frac{\varepsilon_m^{n,\mathbf{k}}}{k_B T}\right), \tag{32}$$

where k_B is Boltzmann's constant. From this expression it is straightforward to obtain harmonic expressions to the internal energy given as

$$\mathscr{E}(T) = k_B T^2\left(\frac{\partial\log(Q_{vib})}{\partial T}\right) = \sum_{\mathbf{k}n}\hbar\omega_{\mathbf{k}n}\left(\frac{1}{2} + \frac{1}{\exp\left(\frac{\hbar\omega_{\mathbf{k}n}}{k_B T}\right) - 1}\right). \tag{33}$$

However, the harmonic approximation has its limitations like zero thermal expansion, temperature independence of elastic constants and bulk modulus, equality of constant-pressure and constant-volume specific heats, as well as infinite thermal conductivity and phonon lifetimes.[103] To overcome such drawbacks, the simplest way includes quasi-harmonic quantities in the sense of the quasi-harmonic approximation (QHA).[104–107] According to the QHA the Helmholz free energy of a crystal is written retaining the same harmonic expression but introducing an explicit dependence of vibrational phonon frequencies on volume as given by

$$F^{QHA}(T,V) = U_0(V) + F_{vib}^{QHA}(T,V), \tag{34}$$

where $U_0(V)$ is the zero-temperature internal energy of the crystal without any vibrational contribution (similar to volume constrained geometry optimizations) and the vibrational part is given by

$$F_{vib}^{QHA}(T,V) = \sum_{\mathbf{k}n}\frac{\hbar\omega_{\mathbf{k}n}(V)}{2} + k_B T\left[\ln\left(1 - \exp\left(-\frac{\hbar\omega_{\mathbf{k}n}(V)}{k_B T}\right)\right)\right], \tag{35}$$

where the first part refers to the zero-point energy of the system. The equilibrium volume at a given temperature T is obtained by minimizing $F^{QHA}(T,V)$ with respect to volume V while keeping T constant.

6.5 Technical details

All molecular dynamic dipole polarizabilities $\alpha(i\omega)$ were calculated using time-dependent density functional theory (TD-DFT).[108,109] A variant of the PBE0 hybrid functional was used, with a Fock-exchange admixture of 37.5% (dubbed PBE38). This method has already proved its accuracy and robustness in previous works.[24,26,27] The atomic orbital (AO) basis sets used in the TD-DFT calculations are of def2-QZVP[110,111] quality closely representing the complete basis set (CBS) limit for this property. The following def2-ECPs are used: ECP-28[112–114] covering 28 core electrons (for Rb, Sr, Y-Cd, In-SB, Te-Xe, Ce-Lu), ECP-46 covering 46 core electrons (for Cs, Ba, La), and ECP-60 covering 60 core electrons (for Hf-Hg, Tl-Bi, Po-Rn) as defined in Ref. 110. Crystal structures have been extracted from MATERIALS PROJECT[115] (for alkali metals Li, Na, K, Rb, and Cs; for alkaline earth metals Be, Mg, Ca, Sr, and Ba; for *d*-block elements Sc and Y) and used within the PEECM to obtain dynamic polarizabilities. Since some elements have experimental crystal structures that exhibit high dipole moments, the PEECM calculation could not be successfully converge (e.g, for Ti, Zr, Hf, V, Nb, and Ta). For these elements closed-shell monomers without dipole moments were extracted from the crystal structure and dynamic polarizabilities were calculated using the presented level of theory. Furthermore, periodic coordination numbers were assigned to those six extracted clusters to be used as approximated "periodic" reference system. The ECPs used to create the shells within part (II) of the particular clusters have been extracted from the TURBOMOLE[116,117] basis set library (nomenclature reads as follows element/ecp-electrons in core-name; for alkali metals we used: Li/ecp-2-sdf, Na/ecp-10-sdf, K/ecp-18-sdf, Rb/ecp-36-sdf, and Cs/ecp-54-sdf; for earth alkali metals we used: Be/ecp-2-sdf, Mg/ecp-10-sdf[118], Ca/ecp-18-sdf, Sr/ecp-36-sdf, and Ba/ecp-54-sdf; for *d*-block elements we used: Sc/ecp-10-mdf[119], Y/ecp-28-mwb[120]; for halogen we applied Cl/ecp-10-sdf and F/ecp-2-sdf and for oxygen we applied O/ecp-2-sdf).

6.6 Computational details

6.6.0.1 X23: Solid state volumina For the determination of the 23 molecular crystal structure volumes we applied the VASP 6.0.8 software package. All PBE PAW calculations used an 800eV plane-wave cutoff (convergence criteria: energy difference $\propto 10^{-6}$). The DFT conjugated gradient method has been used within the optimization where all atomic positions and the cell has been relaxed. For all calculations standard pseudopotentials have been used.

6.6.0.2 ICE10: QHA calculations Quasi-harmonic approximation calculations have been performed for eight different ice polymorphs. For this purpose the QHA implementation within the CRYSTAL17 code has been applied in combination with HSE-3c. Here, four different volumina have been used (steps of 2.5%) for which overall ten different temperatures have been applied (ranging from 10 K to 100 K applying 10 K steps). V_0 has been extracted from Helmholtz free energy calculations at a pressure of 0 GPa.

6.6.0.3 Timings for the cyclohexanedione crystal A self-consistent field (SCF) calculation has been performed for the cyclohexanedione crystal using the PBE/800eV setup in VASP 6.0.8 (convergence criteria: energy difference $\propto 10^{-6}$). The converged wave function has been applied to determine the pure timing arising from each dispersion correction. For SCAN-rVV10 and vdW-DF2 we applied an PAW cutoff of 800eV. For all calculations standard pseudopotentials have been used.

6.6.0.4 Refractive indices Refractive indices have been calculated for an organic polymer database. For D3 and D4 we use polarizabilities from the DFTD3 and DFTD4 standalone programs. For other dispersion corrections we have calculated polarizabilities using VASP 6.0.8 with a PBE/500eV setup using standard pseudopotentials.

6.6.0.5 Salt polarizabilities Salt polarizabilities have been calculated for several alkali halides. For D3 and D4 we use polarizabilities from the DFTD3 and DFTD4 standalone programs. For other dispersion corrections we have calculated polarizabilities using VASP 6.0.8 with a PBE/500eV setup using standard pseudopotentials. All values are given in table 3 together with statistical measures.

7 References

Notes and references

1 R. G. Parr, *W. Yang Density-functional theory of atoms and molecules*, 1989.

2 W. Kohn, *Rev. Mod. Phys.*, 1999, **71**, 1253.

3 K. Burke, *J. Chem. Phys.*, 2012, **136**, 150901.

4 S. Grimme, A. Hansen, J. G. Brandenburg and C. Bannwarth, *Chem. Rev.*, 2016, **116**, 5105–5154.

5 Y. Zhang and W. Yang, *J. Chem. Phys.*, 1998, **109**, 2604–2608.

6 P. Mori-Sànchez, A. J. Cohen and W. Yang, *J. Chem. Phys.*, 2006, **125**, 201102.

7 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.

8 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2005, **122**, 154104.

9 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2005, **123**, 154101.

10 A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.*, 2009, **102**, 073005.

11 A. Tkatchenko, R. A. DiStasio Jr, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.

12 R. A. DiStasio, V. V. Gobre and A. Tkatchenko, *J. Phys.: Condens. Matter*, 2014, **26**, 213202.

13 T. Bucko, S. Lebegue, J. Hafner and J. G. Angyan, *J. Chemical Theory Comput.*, 2013, **9**, 4293–4299.

14 T. Bučko, S. Lebègue, J. G. Ángyán and J. Hafner, *J. Chem. Phys.*, 2014, **141**, 034114.

15 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2004, **92**, 246401.

16 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2005, **95**, 109902.

17 K. Berland and P. Hyldgaard, *Phys. Rev. B*, 2014, **89**, 035412.

18 D. C. Langreth, B. I. Lundqvist, S. D. Chakarova-Käck, V. R. Cooper, M. Dion, P. Hyldgaard, A. Kelkkanen, J. Kleis, L. Kong and S. Li, *J. Phys.: Condens. Matter*, 2009, **21**, 084203.

19 K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard and B. I. Lundqvist, *Rep. Prog. Phys.*, 2015, **78**, 066501.

20 G. A. DiLabio and A. Otero-de-la-Roza, *ArXiv E-Prints*, 2014.

21 O. A. Vydrov and T. Van Voorhis, *J. Chem. Phys.*, 2009, **130**, 104105.

22 O. A. Vydrov and T. Van Voorhis, *Phys. Rev. Lett.*, 2009, **103**, 063004.

23 A. Stone, *The theory of intermolecular forces*, oUP oxford, 2013.

24 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

25 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

Table 3 Static polarizabilities as obtained by experimental measurements (polarizabilities are extracted from Ref. 35), and theoretical values for D4, D3, TS, TS/HI, MBD, MBD/HI, and MBD/FI. All polarizabilities and statistical measures are given in Bohr^3. Statistical values are given in %. All crystal structures have the Fm$\bar{3}$m space group.

					Salt polarizabilities				
#	Compound	α_{exp}	α_{D4}	α_{D3}	α_{TS}	α_{MBD}	$\alpha_{TS/HI}$	$\alpha_{MBD/HI}$	$\alpha_{MBD/FI}$
1	LiF	6.1	9.0	11.9	144.4	—	8.4	8.4	15.1
2	LiCl	19.6	20.3	18.7	163.9	132.0	27.6	27.6	32.5
3	LiBr	27.9	27.1	22.2	171.4	142.4	42.7	42.2	51.3
4	LiI	42.0	39.4	28.2	189.4	158.1	81.9	77.6	82.6
5	NaF	7.8	10.6	16.3	139.7	117.2	18.0	17.9	16.2
6	NaCl	22.0	22.4	23.1	158.3	143.2	34.6	34.5	31.0
7	NaBr	29.6	29.0	26.6	166.2	151.2	41.5	41.4	43.4
8	NaI	42.3	41.6	32.6	184.5	167.5	62.1	62.0	61.9
9	KF	13.6	16.2	21.0	244.6	219.9	58.7	57.8	27.4
10	KCl	28.2	28.0	27.9	261.8	247.2	68.3	67.6	34.8
11	KBr	35.7	34.6	31.4	270.7	256.3	75.1	74.4	47.1
12	KI	49.9	47.2	37.3	291.0	274.1	95.5	94.7	65.5
13	RbF	17.4	19.6	24.5	269.3	249.5	85.8	84.6	36.2
14	RbCl	31.8	31.4	31.3	282.7	270.7	91.3	90.2	38.2
15	RbBr	40.0	37.8	34.8	290.8	279.4	98.0	97.0	50.4
16	RbI	54.6	50.5	40.8	310.9	297.3	118.4	117.2	68.8
		MD	5.7	156.2	873.7	694.0	126.1	123.6	62.7
		MAD	10.1	156.2	873.7	694.0	126.1	123.6	62.7
		RMSD	16.2	134.5	577.2	402.0	103.0	101.2	39.5
		AMAX	47.1	568.7	2267.2	1516.8	393.2	386.1	147.4

26 E. Caldeweyher, C. Bannwarth and S. Grimme, *J .Chem. Phys.*, 2017, **147**, 034112.

27 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *J. Chem. Phys*, 2019, **150**, 154122.

28 M. Bursch, E. Caldeweyher, A. Hansen, H. Neugebauer, S. Ehlert and S. Grimme, *Acc. Chem. Res.*, 2018, **52**, 258–266.

29 D. Cahen, R. Naaman and Z. Vager, *Adv. Funct. Mater.*, 2005, **15**, 1571–1578.

30 L. Kronik and N. Koch, *MRS Bull.*, 2010, **35**, 417–421.

31 S. R. Forrest and M. E. Thompson, *Chem. Rev.*, 2007, **107**, 923–925.

32 N. Atodiresei, V. Caciuc, P. Lazić and S. Blügel, *Phys. Rev. Lett.*, 2009, **102**, 136809.

33 G. Mercurio, E. R. McNellis, I. Martin, S. Hagen, F. Leyssner, S. Soubatch, J. Meyer, M. Wolf, P. Tegeder, F. S. Tautz and K. Reuter, *Phys. Rev. Lett.*, 2010, **104**, 036102.

34 A. Tkatchenko, L. Romaner, O. T. Hofmann, E. Zojer, C. Ambrosch-Draxl and M. Scheffler, *MRS Bull.*, 2010, **35**, 435–442.

35 J. R. Tessman, A. H. Kahn and W. Shockley, *Phys. Rev.*, 1953, **92**, 890.

36 G. X. Zhang, A. Tkatchenko, J. Paier, H. Appel and M. Scheffler, *Phys. Rev. Lett.*, 2011, **107**, 245501.

37 A. M. Burow, M. Sierka, J. Döbler and J. Sauer, *J. Chem. Phys.*, 2009, **130**, 174710.

38 K. N. Kudin and G. E. Scuseria, *Chem. Phys. Lett.*, 1998, **289**, 611–616.

39 J. E. Mayer and M. G. Mayer, *Phys. Rev.*, 1933, **43**, 605.

40 S. Grimme, *J. Chem. Theory Comput.*, 2014, **10**, 4497–4514.

41 H. Schröder and T. Schwabe, *J. Comput. Chem.*, 2016, **37**, 2052–2059.

42 A. Ambrosetti, A. M. Reilly, R. A. DiStasio Jr and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.

43 T. Gould, S. Lebègue, J. G. Ángyán and T. Bučko, *J. Chem. Theory Comput.*, 2016, **12**, 5920–5930.

44 M. A. F. Afzal, C. Cheng and J. Hachmann, *J. Chem. Phys.*, 2018, **148**, 241712.

45 L. Turcani, E. Berardo and K. E. Jelfs, *J. Comput. Chem.*, 2018, **39**, 1931–1942.

46 S. Grimme, J. G. Brandenburg, C. Bannwarth and A. Hansen, *J. Chem. Phys.*, 2015, **143**, 054107.

47 S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 2847–2862.

48 J. G. Brandenburg, T. Maas and S. Grimme, *J. Chem. Phys.*, 2015, **142**, 124104.

49 A. Otero-De-La-Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **137**, 054103.

50 A. M. Reilly and A. Tkatchenko, *J. Chem. Phys.*, 2013, **139**, 024705.

51 A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè and A. Michaelides, *Proc. Natl. Acad. Sci.*, 2018, **115**, 1724–1729.

52 A. J. Cruz-Cabeza, S. M. Reutzel-Edens and J. Bernstein,

Chem. Soc. Rev., 2015, **44**, 8619–8635.
53 A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse *et al.*, *Nature*, 2017, **543**, 657.
54 J. Klimeš, *J. Chem. Phys.*, 2016, **145**, 094506.
55 S. Wen and G. J. O. Beran, *J. Chem. Theory Comput*, 2011, **7**, 3733–3742.
56 M. Cutini, B. Civalleri, M. Corno, R. Orlando, J. G. Brandenburg, L. Maschio and P. Ugliengo, *J. Chem. Theory Comput.*, 2016, **12**, 3340–3352.
57 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671.
58 J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.
59 J. G. Brandenburg, J. E. Bates, J. Sun and J. P. Perdew, *Phys. Rev. B*, 2016, **94**, 115144.
60 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
61 B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B*, 1999, **59**, 7413.
62 Y. Zhang and W. Yang, *Phys. Rev. Lett.*, 1998, **80**, 890–890.
63 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
64 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
65 P. J. Stephens, F. J. Devlin, C. F. N. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
66 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
67 K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. B*, 2010, **82**, 081101.
68 L. Trombach, S. Ehlert, S. Grimme, P. Schwerdtfeger and J.-M. Mewes, *Phys. Chem. Chem. Phys.*, 2019.
69 J. Moellmann and S. Grimme, *J. Phys. Chem. C*, 2014, **118**, 7615–7621.
70 D. J. Carter and A. L. Rohl, *J. Chem. Theory Comput.*, 2014, **10**, 3423–3437.
71 J. G. Brandenburg, E. Caldeweyher and S. Grimme, *Phys. Chem. Chem. Phys.*, 2016, **18**, 15519–15523.
72 R. Sure and S. Grimme, *J. Comput. Chem.*, 2013, **34**, 1672–1685.
73 R. Dovesi, R. Orlando, A. Erba, C. M. Zicovich-Wilson, B. Civalleri, S. Casassa, L. Maschio, M. Ferrabone, M. De La Pierre, P. d'Arco *et al.*, *Int. J. Quantum Chem.*, 2014, **114**, 1287–1317.
74 E. Caldeweyher and J. G. Brandenburg, *J. Phys.: Condens. Matter*, 2018, **30**, 213001.
75 L. Goerigk, H. Kruse and S. Grimme, *ChemPhysChem*, 2011, **12**, 3421–3433.
76 M. Macher, J. Klimeš, C. Franchini and G. Kresse, *J. Chem. Phys.*, 2014, **140**, 084502.
77 P. Lazić, V. Caciuc, N. Atodiresei, M. Callsen and S. Bluegel, *J. Phys. Condens. Matter*, 2014, **26**, 263001.
78 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus *et al.*, *J. Phys. Condens. Matter*, 2017, **29**, 273002.
79 A. Stanislaus and B. H. Cooper, *Catal. Rev.-Sci. Eng.*, 1994, **36**, 75–123.
80 J. Wellendorff, A. Kelkkanen, J. J. Mortensen, B. I. Lundqvist and T. Bligaard, *Top. Catal.*, 2010, **53**, 378–383.
81 D. Syomin, J. Kim, B. E. Koel and G. B. Ellison, *J. Phys. Chem. B*, 2001, **105**, 8387–8394.
82 J. Heidberg, M. Kandel, D. Meine and U. Wildt, *Surf. Sci.*, 1995, **331**, 1467–1472.
83 G. Spoto, E. N. Gribov, G. Ricchiardi, A. Damin, D. Scarano, S. Bordiga, C. Lamberti and A. Zecchina, *Prog. Surf. Sci.*, 2004, **76**, 71–146.
84 M. Sterrer, T. Risse and H. J. Freund, *Appl. Catal. A*, 2006, **307**, 58–61.
85 P. Ugliengo and A. Damin, *Chem. Phys. Lett.*, 2002, **366**, 683–690.
86 R. Valero, J. R. B. Gomes, D. G. Truhlar and F. Illas, *J. Chem. Phys.*, 2008, **129**, 124710.
87 B. Civalleri, L. Maschio, P. Ugliengo and C. M. Zicovich-Wilson, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6382–6386.
88 V. Staemmler, *J. Phys. Chem. A*, 2011, **115**, 7153–7160.
89 R. Wichtendahl, M. Rodriguez-Rodrigo, U. Härtel, H. Kuhlenbeck and H. J. Freund, *Surf. Sci.*, 1999, **423**, 90–98.
90 Z. Dohnalek, G. A. Kimmel, S. A. Joyce, P. Ayotte, R. S. Smith and B. D. Kay, *J. Phys. Chem. B*, 2001, **105**, 3747–3751.
91 S. Ehrlich, J. Moellmann, W. Reckien, T. Bredow and S. Grimme, *ChemPhysChem*, 2011, **12**, 3414–3420.
92 A. D. Boese and J. Sauer, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16481–16493.
93 S. K. Dunn and G. E. Ewing, *J. Phys. Chem.*, 1992, **96**, 5284–5290.
94 A. G. Cabello-Cartagena, J. Vogt and H. Weiss, *J. Chem. Phys.*, 2010, **132**, 074706.
95 D. C. Ghosh and N. Islam, *Int. J. Quantum Chem.*, 2010, **110**, 1206–1213.
96 L. Pauling, *Cornell University, Press, Ithaca, New York*, 1960, 17.
97 P. Pyykkö and M. Atsumi, *Chem.-Eur. J.*, 2009, **15**, 186–197.
98 K. J. Miller, *J. Am. Chem. Soc.*, 1990, **112**, 8533–8542.
99 J. R. Bunch and L. Kaufman, *Linear Algebra Its Appl.*, 1980, **34**, 341–370.
100 B. M. Axilrod and E. Teller, *J. Chem. Phys.*, 1943, **11**, 299–300.
101 Y. Muto, Proc. Phys. Math. Soc. Jpn., 1943, pp. 629–631.
102 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2000, **10**, 6615–6620.
103 M. T. Dove, *Introduction to lattice dynamics*, Cambridge university press, 1993, vol. 4.
104 A. Erba, *J. Chem. .Phys.*, 2014, **141**, 124115.
105 A. Erba, M. Shahrokhi, R. Moradian and R. Dovesi, *J. Chem. Phys.*, 2015, **142**, 044114.

106 A. Erba, J. Maul, M. De La Pierre and R. Dovesi, *J. Chem. Phys.*, 2015, **142**, 204502.
107 A. Erba, J. Maul, M. Itou, R. Dovesi and Y. Sakurai, *Phys. Rev. Lett.*, 2015, **115**, 117402.
108 S. J. A. van Gisbergen, J. G. Snijders and E. J. Baerends, *J. Chem. Phys.*, 1995, **103**, 9347–9354.
109 V. P. Osinga, S. J. A. van Gisbergen, J. G. Snijders and E. J. Baerends, *J. Chem. Phys.*, 1997, **106**, 5091–5101.
110 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
111 F. Weigend, F. Furche and R. Ahlrichs, *J. Chem. Phys.*, 2003, **119**, 12753–12762.
112 D. Andrae, U. Haeussermann, M. Dolg, H. Stoll and H. Preuss, *Theoretica chimica acta*, 1990, **77**, 123–141.
113 B. Metz, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2000, **113**, 2563–2569.
114 K. A. Peterson, D. Figgen, E. Goll, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2003, **119**, 11113–11123.
115 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, *APL Materials*, 2013, **1**, 011002.
116 F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, *WIREs Comput. Mol. Sci.*, 2014, **4**, 91–100.
117 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.
118 P. Fuentealba, L. Von Szentpaly, H. Preuss and H. Stoll, *Journal of Physics B: Atomic and Molecular Physics*, 1985, **18**, 1287.
119 M. Dolg, U. Wedig, H. Stoll and H. Preuss, *J. Chem. Phys.*, 1987, **86**, 866–872.
120 D. Andrae, U. Haeussermann, M. Dolg, H. Stoll and H. Preuss, *Theor. Chem. Acc.*, 1990, **77**, 123–141.