Understanding the role of vibrations, exact exchange, and many-body van der Waals interactions in the cohesive properties of molecular crystals

Anthony M. Reilly and Alexandre Tkatchenko

Citation: J. Chem. Phys. **139**, 024705 (2013); doi: 10.1063/1.4812819
View online: http://dx.doi.org/10.1063/1.4812819
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v139/i2
Published by the AIP Publishing LLC.

Understanding the role of vibrations, exact exchange, and many-body van der Waals interactions in the cohesive properties of molecular crystals

Anthony M. Reilly[a] and Alexandre Tkatchenko[b]
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

(Received 14 May 2013; accepted 19 June 2013; published online 12 July 2013)

The development and application of computational methods for studying molecular crystals, particularly density-functional theory (DFT), is a large and ever-growing field, driven by their numerous applications. Here we expand on our recent study of the importance of many-body van der Waals interactions in molecular crystals [A. M. Reilly and A. Tkatchenko, J. Phys. Chem. Lett. **4**, 1028 (2013)], with a larger database of 23 molecular crystals. Particular attention has been paid to the role of the vibrational contributions that are required to compare experiment sublimation enthalpies with calculated lattice energies, employing both phonon calculations and experimental heat-capacity data to provide harmonic and anharmonic estimates of the vibrational contributions. Exact exchange, which is rarely considered in DFT studies of molecular crystals, is shown to have a significant contribution to lattice energies, systematically improving agreement between theory and experiment. When the vibrational and exact-exchange contributions are coupled with a many-body approach to dispersion, DFT yields a mean absolute error (3.92 kJ/mol) within the coveted "chemical accuracy" target (4.2 kJ/mol). The role of many-body dispersion for structures has also been investigated for a subset of the database, showing good performance compared to X-ray and neutron diffraction crystal structures. The results show that the approach employed here can reach the demanding accuracy of crystal-structure prediction and organic material design with minimal empiricism. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4812819]

I. INTRODUCTION

The structure and properties of molecular crystals have long been of great interest, not just for fundamental reasons of understanding molecular aggregation but also due to their numerous applications. As the preferred form of active pharmaceutical ingredients for oral administration, the dissolution and morphology of drug-molecule crystals are very important for bio-availability and processing.[1] For these reasons alone the prediction of molecular crystal structures is of the utmost importance.[2] In addition, molecular crystals can also have a wide range of optical, electronic, and mechanical properties,[3] which in some cases can be tuned based on environmental variables[4] or composition.[5]

The importance of molecular crystals has lead to various computational approaches being employed to understand and elucidate their properties. While crystal-structure prediction and pharmaceuticals are very active and important areas of research,[6,7] many computational studies have also focused on energetic materials,[8,9] proton-transfer systems,[4,10] vibrational properties such as phonons and thermal displacement parameters,[11,12] as well as electronic properties.[13] The methods employed range from tailor-made force fields[8] to generic force fields such as OPLS[14] and CHARMM[15] to first-principles methods such as density-functional theory (DFT).

The growth in the use of DFT for studying molecular crystals has gone in tandem with the development of dispersion-inclusive DFT methods,[16] as semi-local functionals (as well as hybrid functionals) neglect long-range dispersion or van der Waals (vdW) interactions, which can be essential for the formation of molecular solids. There are numerous dispersion-inclusive DFT methods, including the pairwise approaches of DFT-D,[17,18] DFT+vdW,[19] and the exchange-diplole moment (XDM) method[20] as well as "vdW density functionals" such as vdW-DF1[21] and vdW-DF2.[22]

Given the important applications of molecular crystals, considerable effort has been put into benchmarking these approaches. Crystal-structure prediction trials show encouraging results for DFT-D approaches,[6] but the small energy differences involved[7] mean that different methods can still yield conflicting predictions of the stable structure of crystal.[23] The most common type of benchmark for these systems is to calculate lattice energies,

$$E_{\text{lat}} = \frac{E_{\text{s}}}{Z} - E_{\text{g}}, \tag{1}$$

where E_{s} is the total energy of a unit-cell, Z is the number of molecules per unit-cell, and E_{g} is the total energy of the molecule in the gas phase, in its lowest-energy conformation. In the absence of an established computational standard, the calculated lattice energies are often compared with sublimation enthalpies.[24–27] This comparison requires accounting for a number of vibrational contributions. In many instances these contributions were either ignored or considered in the high-temperature limit, whereby following the Dulong-Petit law the relationship reduces to[25,28]

$$\Delta H_{\text{sub}}(T) = -E_{\text{lat}} - 2RT. \tag{2}$$

[a]Electronic mail: reilly@fhi-berlin.mpg.de
[b]Electronic mail: tkatchen@fhi-berlin.mpg.de

A realistic benchmark of DFT methods requires the accurate estimation of the vibrational contributions to the sublimation enthalpy to avoid systematic bias towards methods that over- or underestimate lattice energies. It is equally important to properly assess the role of the density functional used to calculate the lattice energies. While many studies have focused on the role of dispersion,[25,29,30] few have critically assessed the shortcomings of semi-local functionals in detail. For example, the de-localisation or self-interaction errors in DFT[31] can often have a significant effect on hydrogen-bonded systems.[32]

Recent work by Otero-de-la-Roza and Johnson has shown that apart from an implementation of XDM,[33] many dispersion-inclusive methods systematically overestimate lattice energies. In subsequent work we have shown that the origin of this overestimation for DFT+vdW (also known as the Tkatchenko-Scheffler or TS method) clearly lies in the absence of many-body collective response and energy contributions.[34] The recently developed many-body dispersion (MBD) method[35,36] was shown to correctly model these contributions, halving the mean absolute error (MAE) compared to the pairwise method and reducing the MAE between DFT and experiment for 16 molecular crystals to 3.8 kJ/mol, within the coveted "chemical accuracy" target of 4.2 kJ/mol. Crucially, such good agreement hinged on using MBD with a non-empirical hybrid functional (Perdew-Burke-Ernzerhof (PBE0)[37]). The corresponding semi-local functional, PBE,[38] achieved an accuracy of 6.4 kJ/mol.

In the present contribution we expand on our recent work in three ways. First, we consider the computational evaluation of the vibrational contributions to sublimation enthalpies in detail (Sec. III A), focusing on the role of phonon dispersion and anharmonicity. Secondly, predicted lattice energies for a larger set of 23 molecular crystals will be presented and discussed in terms of many-body dispersion and exact-exchange contributions (Sec. III B). Finally, we explore the role of many-body dispersion in determining the structure of molecular crystals, presenting optimized lattice constants and structures for some representative systems (Sec. III C). The computational methodology is outlined in Sec. II.

II. METHODOLOGY

A. Density-functional theory calculations

All DFT calculations were performed using the CASTEP plane-wave DFT code (version 6.01).[39] A plane-wave cut-off energy of 1000 eV was used in all calculations, sufficient to converge both norm-conserving and ultra-soft pseudopotential calculated total energies to less than 1 meV/atom. For solid-state calculations a k-point sampling of at least 0.06 $Å^{-1}$ was used to sample the electronic reciprocal space. Norm-conserving pseudo-potentials were employed for calculating total energies used in the lattice-energy calculations, while the phonon calculations were performed using ultra-soft pseudo-potentials, with phonon calculations starting from the ultra-soft pseudo-potential minimized geometry.[40] To ensure convergence, the Fourier-transform grid scales for the density and its high-frequency components were set to 2.25 and 2.50 times the wavefunction value, respectively. Calculations were performed with the PBE[38] and hybrid PBE0[37] functionals. Pairwise dispersion-interaction contributions were added using the DFT+vdW (or TS method),[19] using the built-in functionality of CASTEP.[41]

Initial structures for the crystals were obtained either from Otero-de-la-Roza and Johnson[30] or the Cambridge Structural Database.[42] Space-group symmetry was exploited for solid-state calculations. Isolated molecules, representing the gas-phase reference for lattice-energy calculations, were modeled in cubic supercells of at least 15 Å in length, with some of the larger molecules, such as adamantane and anthracene being modeled in a supercell of 20 Å length. In most cases the initial structure for the isolated molecule was obtained directly from the optimized crystal structure. For oxalic acid, care was taken to use the correct gas-phase conformation, which features intra-molecular hydrogen bonds that substantially alter the gas-phase stability of the molecule.[43] Succinic acid is known to have a number of gas-phase conformers around 445 K,[44] but here we use only the most stable C_2 conformer, which is likely to be the main conformer at the lower temperature range (360–375 K) of the sublimation measurement.[45] The full geometries (lattice parameters for crystals and coordinates for both crystals and molecules) were optimized at the PBE+TS level, with convergence criteria of 5×10^{-6} eV/atom for total energies, 5×10^{-3} eV/Å for forces, 1×10^{-3} Å for displacements, and 2×10^{-2} GPa for stresses. All subsequent PBE0 and MBD single-point calculations were performed using these PBE+TS geometries.[46]

B. Many-body dispersion method

The MBD method is described in detail elsewhere.[34–36] Briefly, the atoms in the system are represented as quantum harmonic dipole oscillators defined by the TS polarizabilities. In a two stage process, the frequency-dependent electrodynamic response of the system is first obtained using the self-consistent screening (SCS) equations.[47,48] The resulting screened polarizabilities and static frequencies are used in a Hamiltonian that explicitly considers all orders of dipole interactions. The calculation of MBD energies has been implemented in a development version of CASTEP. For the solid-state calculations of both the SCS equations and MBD energies, interactions with periodic replicas of the unit-cell have been included in the dipole-coupling elements within a spherical cut-off distance of ≈50 Å. The SCS equations are solved for the unit-cell only, as there is no dependence on the size of the simulation cell, as long as sufficient periodic interactions are included. However, the many-body Hamiltonian must be evaluated for a supercell (typically 25 Å in each direction) to properly sample the reciprocal space of the explicit many-body interactions. For isolated molecules the MBD calculations are performed with a single molecule in an aperiodic calculation.

The derivation of analytical forces and stresses for the MBD method is currently in progress,[49] although analytical gradients exist for the contributions from the SCS equations.[50] In the present work, finite-difference forces and stresses have

been used to optimize a small number of the crystal structures studied. The convergence criteria were the same as those for the PBE+TS optimizations.

C. Estimation of vibrational contributions to sublimation enthalpies

The enthalpy of sublimation can be related to the lattice energy via a number of vibrational and thermodynamic quantities:

$$\Delta H_{\mathrm{sub}}(T) = -E_{\mathrm{lat}} + (E_{\mathrm{ZPE,g}} - E_{\mathrm{ZPE,s}}) + \int_0^T \Delta C_p(T)\mathrm{d}\,T + \sum \Delta H_{\mathrm{trans}}, \quad (3)$$

where s and g refer to the solid and gas-phase respectively, E_{ZPE} is the zero-point vibrational energy (ZPE), ΔC_p is the difference between the gas-phase and solid-state heat capacities, and $\Delta H_{\mathrm{trans}}$ is the enthalpy of transformation for any phase transition that occurs between 0 K and T. This expression is only valid when the isolated molecule has a single well-defined conformation at the temperature of the measurement. When this is not the case additional terms accounting for the conformation ensemble must be included. By design, the present database includes relatively rigid molecules, which are known to have a single conformation at the temperature of the experimental sublimation-enthalpy measurements. In the limit of an ideal gas, taking the rigid-rotor and harmonic approximations and assuming no solid-state phase transformations, Eq. (3) reduces to

$$\Delta H_{\mathrm{sub}}(T) = -E_{\mathrm{lat}} + \Delta E_{\mathrm{vib}}(T) + 4RT, \quad (4)$$

where $\Delta E_{\mathrm{vib}}(T)$ is the total (thermal and zero-point) vibrational-energy difference between the gas and the solid. For a linear molecule, such as CO_2, the final term becomes $7/2RT$ to account for the missing rotational degree of freedom. For a more detailed discussion of these approximations see, e.g., Otero-de-la-Roza and Johnson.[30] The vibrational quantities can be readily calculated in the harmonic limit using DFT phonon calculations,[30,51]

$$E_{\mathrm{vib}}(T) = \sum_q \left(\sum_p \left(\frac{\hbar\omega_{p,q}}{2} + \frac{\hbar\omega_{p,q}}{\exp\left(\frac{\hbar\omega_{p,q}}{k_{\mathrm{b}}T}\right) - 1} \right) \right), \quad (5)$$

where $\omega_{p,q}$ is the pth phonon frequency at wavevector q, T is the temperature, and k_{b} is Boltzmann's constant. The first part of Eq. (5) is the ZPE, while the second part gives the thermal energy based on a Planck distribution. In the absence of phonon dispersion or for an isolated-molecule supercell calculation, the summation can be performed using only the Γ-point frequencies. In the present work we have determined the vibrational terms using both unit-cell and supercell finite-displacement phonon calculations[52,53] (at the PBE+TS level of theory) to interpolate to arbitrarily fine q-point grids, using the in-built functionality of CASTEP. The thermal contribution was evaluated at 298 K, the temperature to which most sublimation enthalpies are extrapolated.[54,55] The application of density functional perturbation theory[51] is not possible as it would require a fully self-consistent implementation of the TS vdW term, which is not currently available in the CASTEP code. The vibrational quantities for isolated molecules were also determined using phonon calculations. In such isolated-molecule calculations we expect to find six zero-frequency modes (or five for a linear molecule), which correspond to the translational and rotational degrees of freedom. Three such zero-frequency modes are expected at the Γ point in the solid state, corresponding to complete translations of the crystal. In any numerical calculations, minor deviations with small negative or positive non-zero values will occur.[56] Care was taken to ignore the spurious contributions from these frequencies. In no case were "true" negative frequencies found, indicating that all of the optimized molecular geometries and crystal structures represented true minima on the potential-energy surface.

For adamantane,[57] anthracene,[58] hexamine,[57] naphthalene,[59] succinic acid,[60] trioxane,[61] and urea[62] experimental solid-state C_p data are available. These data have been used with DFT-calculated gas-phase C_p values and the DFT ZPE values to evaluate Eq. (3) directly, yielding a semi-anharmonic value for the lattice energy. Taking the derivative of Eq. (5) with respect to temperature, the constant-volume heat capacity for an isolated molecule can be calculated as

$$C_V(T) = k_{\mathrm{b}} \sum_p \left(\frac{\left(\frac{\hbar\omega_p}{k_{\mathrm{b}}T}\right)^2 \exp\left(\frac{\hbar\omega_p}{k_{\mathrm{b}}T}\right)}{\left(\exp\left(\frac{\hbar\omega_p}{k_{\mathrm{b}}T}\right) - 1\right)^2} \right). \quad (6)$$

The calculated values of C_V were then converted to C_p by adding R. To account for the rotational heat capacity of the isolated molecule in the rigid-rotor approximation, $3RT$ was added to the value of the integrated heat-capacity difference.

In the case of adamantane, examination of the literature revealed that a phase transformation from the ordered trigonal phase to a disordered fcc phase occurs at 208 K.[57] The low-temperature structure used as the basis of the DFT calculations here is not appropriate for use with Eq. (4) as this ignores the enthalpy of transition and the highly anharmonic nature of phase transitions. Experimental C_p values are available for the solid phase over the range of 0–298 K.[57] By extrapolating the C_p data through the second-order phase-transition an estimate for the vibrational contribution has been made on the basis of the low-temperature trigonal crystal structure.

III. RESULTS AND DISCUSSION

A. Vibrational contributions to theoretical lattice energies

To enable a direct and fair comparison between theoretical lattice energies and experimental sublimation enthalpies requires careful evaluation of Eqs. (3) or (4). In the harmonic limit, this can be readily achieved using phonon calculations as outlined above. In the solid state, interactions over distances larger than the unit-cell lead to a wavevector dependence of the phonon frequencies. Using finite-displacements within a unit-cell neglects this dispersion, sampling only the Γ point and ignoring the acoustic modes, which are purely translation ($\omega = 0$) at the zone center. By exploiting the

TABLE I. The zero-point energy and total vibrational energy at 298 K (per molecule) of phase-I ammonia as calculated using PBE+TS phonon calculations with three different supercells of length a.

Supercell	$E_{ZPE,s}$ (kJ/mol)	$E_{vib,s}$(298 K) (kJ/mol)	a (Å)
$1 \times 1 \times 1$	98.04	104.21	4.9548
$2 \times 2 \times 2$	98.49	105.86	9.9096
$3 \times 3 \times 3$	98.53	105.88	14.8644

short-range nature of the force-constant matrix the supercell or direct method can extrapolate unit-cell phonon calculations to arbitrary wavevectors, presuming the simulation cell is of sufficient size so that force constants decay to zero.[52,53]

To assess the importance of phonon dispersion and the size of simulation cell for the supercell method we have performed phonon calculations using both a crystallographic unit-cell and supercells of phase-I ammonia. The PBE functional (without a vdW term) is already known to model the phonons of phase-I ammonia well at the Γ point.[11] Table I shows the calculated ZPE and total vibrational energy at 298 K using different sized cells to calculate the phonon frequencies at the PBE+TS level of theory. Changing from a unit-cell to a $3 \times 3 \times 3$ supercell leads to a change in the ZPE of 0.5 kJ/mol and a change in the thermal energy, and hence vibrational contribution, of nearly 1.7 kJ/mol at 298 K, 5% of the sublimation enthalpy. The larger contribution of phonon dispersion to the total energy can be understood as the contribution of the acoustic and low-frequency phonon modes, which are more sensitive to the simulation cell size. While their low frequencies contribute less to the ZPE, they are thermally populated more easily. Figure 1 shows the phonon dispersion for ammonia as calculated using the $2 \times 2 \times 2$ supercell. It is clear that there is appreciable phonon dispersion in the lattice and acoustic modes. Encouragingly, even with a cell of only 9 Å, the acoustic modes are well represented, with only some minor artefacts as they approach the Γ point.

The difference in energies between the $2 \times 2 \times 2$ and $3 \times 3 \times 3$ supercells is of the order of 0.1 kJ/mol, suggesting that supercells around the size of 9–10 Å are of sufficient size to obtain reliable estimates of the vibrational contributions to the sublimation enthalpy. This is supported by

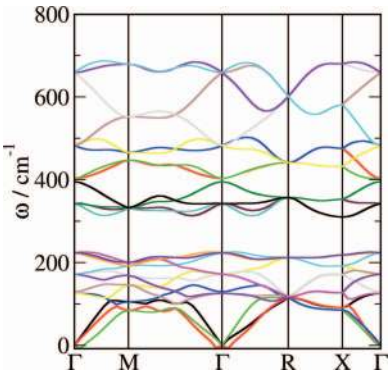

FIG. 1. The phonon dispersion of the lattice modes of phase-I ammonia, calculated with PBE+TS using the supercell method, with a $2 \times 2 \times 2$ simulation cell. The isolated molecule vibrations appear at higher wavenumbers.[11]

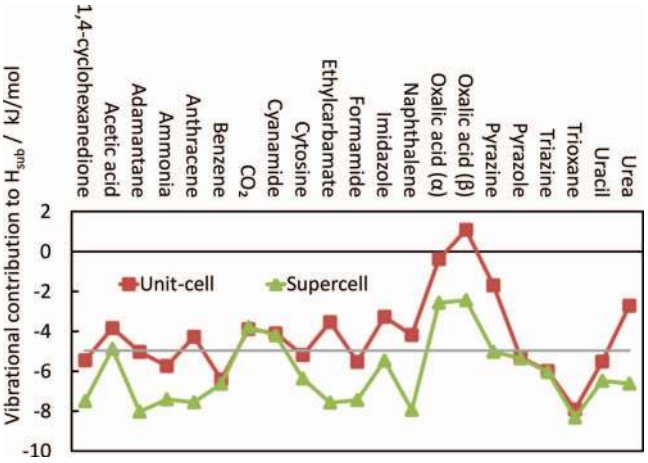

FIG. 2. Vibrational contributions ($\Delta E_{vib} + 4RT$) to the sublimation enthalpy in the harmonic limit at 298 K, as derived from PBE+TS unit-cell and supercell phonon calculations. The gray line indicates the $2RT$ value.

supercell phonon calculations of the polyacene crystals.[63] To study the broad importance of phonon dispersion on the vibrational contributions unit-cell and supercell (of at least 9–10 Å length in each direction) calculations have been performed for the whole database of molecular crystals. The resulting vibrational contributions at 298 K are shown in Figure 2. For the majority of cases there is a substantial increase in the vibration contributions. Those systems where there are only small changes are typically ones with larger unit-cells (e.g., the smallest lattice constant for triazine is ≈7 Å). In the case of oxalic acid (both α and β) and succinic acid the unit-cell calculations give spuriously small values. Both oxalic acid polymorphs have small unit-cells and feature hydrogen bonding over cell boundaries. It is likely that the interpolation of the force-constant matrix to arbitrary wavevectors is not at all valid for such small simulations cells. The relative rigidity will also reduce the contribution from internal degrees of freedom.

Recently, Otero-de-la-Roza and Johnson have also carried out phonon calculations to obtain vibrational contributions for many of the systems studied here.[30] They employed the Einstein model to approximate the zero-point and thermal energy of each crystal. In this model the phonon dispersion is ignored and each phonon branch is approximated with a single frequency obtained at the Γ point. The ZPE contribution from the acoustic modes is then neglected, while their thermal contribution is approximated in the Dulong-Petit limit as $3RT$. This approach was validated based on the lack of any phonon dispersion in the CO_2 crystal. Figure 2 confirms this assertion for CO_2, but it is clear that for the majority of systems there is some contribution from dispersion. However, this may be limited solely to the contribution of the acoustic modes, which are poorly modeled in small simulation cells. To assess this, a comparison between the full phonon treatment and the Einstein model is given for a selection of systems in Table II. Generally, the Einstein model underestimates zero-point energies by 0.5–1.0 kJ/mol, partly due to neglecting the contribution of the acoustic vibrations and phonon dispersion in the low-frequency translational vibrations, which typically

TABLE II. Comparison between zero-point and thermal (298 K) vibrational energies (per molecule) calculated using supercell phonon calculations and the Einstein model using only Γ-point frequencies. A contribution of $3RT$ has been added to the Einstein-model thermal energies (before converting into per molecule values) to account for the acoustic modes.

	Einstein model		Supercell method	
Crystal	E_{ZPE} (kJ/mol)	E_{vib} (kJ/mol)	E_{ZPE} (kJ/mol)	E_{vib} (kJ/mol)
Adamantane	624.52	649.14	625.05	649.15
Ammonia	98.22	105.77	98.53	105.88
Benzene	261.66	277.98	261.81	278.00
Cyanamide	92.87	103.61	93.11	103.66
Ethylcarbamate	281.69	303.27	282.48	303.35
Urea	170.02	184.65	171.19	184.30

TABLE III. Harmonic and semi-anharmonic estimates of the vibrational contribution ($\Delta E_{vib} + 4RT$) to the sublimation enthalpy at 298 K. It should be noted that the semi-anharmonic value for adamantane includes the contribution from the phase transformation known to occur at 208 K.

	$\Delta E_{vib} + 4RT$ (kJ/mol)	
Crystal	Harmonic	Anharmonic
Adamantane	− 8.0	− 11.0
Anthracene	− 7.6	− 10.9
Hexamine	− 9.9	− 10.4
Naphthalene	− 7.9	− 10.5
Succinic acid	− 4.3	− 7.2
Trioxane	− 8.3	− 10.1
Urea	− 6.6	− 8.7

increases the wavevector-averaged zero-point energy. However, overall there is some error cancellation as the final thermal energies at 298 K are in remarkable agreement, with deviations of less than 0.5 kJ/mol.

Given the good agreement between the supercell and Einstein models, we would expect the vibrational contributions of Otero-de-la-Roza and Johnson[30] to agree well with those obtained here. However, in general the values of Otero-de-la-Roza and Johnson are smaller than those obtained here, by around 1.4 kJ/mol on average, although in some cases the difference is as much as 5 kJ/mol. It should be noted that their calculations were performed with the PBE functional using PBE+XDM geometries. In the absence of a vdW contribution, the PBE+XDM geometry will be far from equilibrium, a necessary prerequisite for phonon calculations. This could substantially affect the low-frequency lattice modes and the vdW term is the most likely source of the remaining discrepancies between the two sets of vibrational contributions. Performing the phonon calculations with PBE+MBD might further increase the vibrational contributions but given the difference between the PBE and PBE+TS values it is unlikely to alter the values significantly. This is the case for ammonia, where the Γ-point ZPE increases by less than 0.2 kJ/mol when using MBD instead of TS.

The harmonic treatment ignores effects due to anharmonic thermal motion and cell expansion. To partly understand the effect of these approximations, experimental solid-state heat capacities have been used with DFT-calculated ZPEs and gas-phase heat capacities [Eq. (6)] to evaluate Eq. (3) in a semi-anharmonic fashion. A comparison between the harmonic and semi-anharmonic vibrational contributions is given in Table III. In most cases the semi-anharmonic correction is appreciably larger than the harmonic one by around 2–3 kJ/mol. The small correction for hexamine is likely due to its high-symmetry crystal structure ($I\bar{4}3m$) and rigid molecular structure. Figure 3 shows the experimental and theoretical heat capacities for naphthalene. The experimental curve clearly has additional contributions to the heat capacity in the low temperature region (<75 K) and also diverges from the theoretical data at higher temperatures, indicating the anharmonic contributions of the low-frequency modes, which are populated at low-temperatures and the high-frequency modes, which are only populated sufficiently as the system approaches room temperature.

For naphthalene and anthracene there are experimental gas-phase heat capacities that have been extrapolated to low temperature.[64] For naphthalene using the experimental gas-phase values changes the vibrational contribution by less than 0.2 kJ/mol, a clear indication that in taking the difference between a theoretical and experimental quantity we do not lose any fortuitous cancellation of errors in the computational quantities. Performing the phonon calculations with experimental lattice parameters might account for part of the thermal expansion contributions to the ZPE, but it is not clear to what extent such contributions are contained within the C_p term.

Anharmonic corrections of the order of 2–3 kJ/mol are not negligible in pursuit of chemical accuracy (4.2 kJ/mol). Give the magnitude of the contributions and fact that even where C_p data are available the anharmonicity in the ZPE is ignored, we must always attach an uncertainty to experimental lattice energies, even when there is little uncertainty in the value of the underlying sublimation enthalpy.

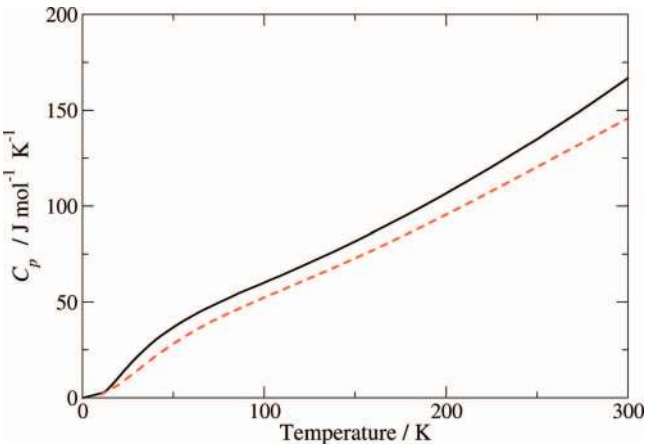

FIG. 3. Solid-state heat capacity of naphthalene as a function of temperature obtained from experiment (black solid line, Ref. 59) and PBE+TS supercell phonon calculations (red dashed line).

TABLE IV. Experimental sublimation enthalpies (ΔH^0_{sub}; at 298 K), vibrational contributions ($\Delta E_{vib} + 4RT$; calculated using supercell phonon calculations with experimental C_p data were available), the resulting experimental lattice energies ($E_{lat,exp}$), DFT-predicted lattice energies ($E_{DF+disp}$, where "DF" is the density functional and "disp" the dispersion method), and the contributions of many-body and exact-exchange effects to the PBE0+MBD binding energies, for each of the crystals studied. Experimental sublimation enthalpies are taken from Otero-de-la-Roza and Johnson,[30] except for those of anthracene,[55] succinic acid,[45] and hexamine.[64,66] All quantities are in kJ/mol.

			Lattice energies						
Molecules	ΔH^0_{sub}	$\Delta E_{vib} + 4RT$	$E_{lat,exp}$	E_{PBE+TS}	$E_{PBE0+TS}$	$E_{PBE+MBD}$	$E_{PBE0+MBD}$	ΔE_{MBD}[a]	ΔE_{PBE0}[b]
1,4-cyclohexanedione	81.1	− 7.5	− 88.6	− 105.9	− 101.2	− 92.2	− 88.2	13.0	4.1
Acetic acid	68.0	− 4.9	− 72.8	− 82.6	− 79.0	− 78.3	− 74.6	4.4	3.6
Adamantane	58.4	− 11.0	− 69.4	− 108.0	− 105.0	− 81.0	− 78.6	26.4	2.4
Ammonia	29.8	− 7.4	− 37.2	− 45.4	− 42.4	− 42.9	− 40.2	2.2	2.7
Anthracene	101.9	− 10.9	− 112.7	− 134.4	− 133.5	− 121.8	− 119.1	14.5	2.8
Benzene	45.1	− 6.6	− 51.7	− 66.3	− 62.0	− 55.0	− 51.0	11.0	4.0
CO_2	24.6	− 3.8[c]	− 28.4	− 25.2	− 24.4	− 21.7	− 21.2	3.2	0.5
Cyanamide	75.5	− 4.2	− 79.7	− 94.3	− 88.8	− 94.3	− 88.8	0.0	5.5
Cytosine	163.4	− 6.4	− 169.8	− 172.6	− 167.9	− 170.0	− 164.5	3.5	5.6
Ethylcarbamate	78.7	− 7.6	− 86.3	− 99.2	− 94.0	− 92.1	− 87.1	6.9	5.0
Formamide	71.8	− 7.4	− 79.2	− 86.3	− 84.0	− 82.8	− 80.7	3.3	2.1
Hexamine	75.8	− 10.4	− 86.2	− 114.9	− 109.8	− 86.9	− 83.4	26.3	3.5
Imidazole	81.4	− 5.5	− 86.8	− 101.9	− 96.3	− 97.1	− 91.4	4.9	5.7
Naphthalene	71.3	− 10.5	− 81.7	− 99.9	− 98.4	− 87.4	− 85.4	13.0	2.0
Oxalic acid (α)	93.7	− 4.7	− 96.3	− 100.7	− 98.5	− 98.1	− 95.7	2.8	2.3
Oxalic acid (β)	93.6	− 2.4	− 96.1	− 104.3	− 100.1	− 98.6	− 94.8	5.3	3.9
Pyrazine	56.3	− 5.0	− 61.3	− 76.2	− 72.0	− 67.3	− 63.0	9.0	4.3
Pyrazole	72.4	− 5.4	− 77.7	− 88.5	− 83.1	− 82.8	− 77.6	5.5	5.2
Succinic acid	123.1	− 7.2	− 130.3	− 147.1	− 143.4	− 138.7	− 135.3	8.1	3.4
Triazine	55.7	− 6.0	− 61.7	− 68.9	− 65.6	− 58.7	− 55.7	10.0	3.0
Trioxane	56.3	− 10.1	− 66.4	− 75.9	− 72.0	− 62.4	− 59.3	12.7	3.2
Uracil	129.2	− 6.5	− 135.7	− 149.0	− 144.6	− 145.9	− 140.4	4.2	5.5
Urea	93.8	− 8.7	− 102.5	− 113.1	− 111.4	− 111.2	− 109.7	1.7	1.5

[a] $\Delta E_{MBD} = (E_{PBE0+MBD} - E_{PBE0+TS})$.
[b] $\Delta E_{PBE0} = (E_{PBE0+MBD} - E_{PBE+MBD})$.
[c] Calculated as $\Delta E_{vib} + 7/2RT$.

B. Prediction of lattice energies

1. Overview of the X23 database

Having determined the vibrational contributions to the sublimation enthalpy, we can now directly compare the DFT predicted lattice energies with experimental estimates. In this work we consider 23 systems, the 21 systems of Otero-de-la-Roza and Johnson,[30] and two additional systems, hexamine and succinic acid. This database (the X23 database) contains a mix of systems that have largely vdW or hydrogen bonding, together with three systems that feature both interactions. These two types of interactions are among the most important and common in molecular crystals, making the database a useful and broad benchmark. Before comparing theory and experiment it is worth briefly commenting on uncertainties in experimental sublimation enthalpies. One estimate of the general uncertainty in sublimation enthalpies is 4.9 kJ/mol, based on a statistical assessment of experimental deviations.[65] Other broad estimates are ±10%.[25] Some of the systems in the present database, such as naphthalene and anthracene, are established thermochemical standards with well-defined uncertainties of around 1–2 kJ/mol.[55] For other systems larger uncertainties, up to 10 kJ/mol, are not uncommon and there can often be a wide variation between different experimental measurements.[64] As such, care must be taken when interpreting experimental data. The majority of the systems in the present database have multiple experimental values and small or moderate uncertainties. Therefore, in the present work we can quantitatively assess the performance of computational methods to an accuracy of around 1–3 kJ/mol.

The predicted lattice energies using PBE and PBE0 with TS and MBD are given in Table IV, while a statistical summary is give in Table V. Figure 4 shows the relative error in the lattice energy, with the database divided into mainly vdW interactions and hydrogen-bonding interactions, although many systems will also have appreciable electrostatic interactions. CO_2 is discussed separately in Sec. III D. The general trends for the performance of the different

TABLE V. The mean error (ME), mean absolute error (MAE), standard deviation (SD), and mean absolute relative error (MARE) of different vdW-inclusive DFT methods with respect to experimental lattice energies. The ME, MAE, and SD are in kJ/mol, while the MARE is in (%).

	ME	MAE	SD	MARE
PBE+TS	13.13	13.40	8.62	17.22
PBE0+TS	9.51	10.02	8.53	12.94
PBE+MBD	4.72	5.91	5.13	8.04
PBE0+MBD	1.17	3.92	4.79	5.51

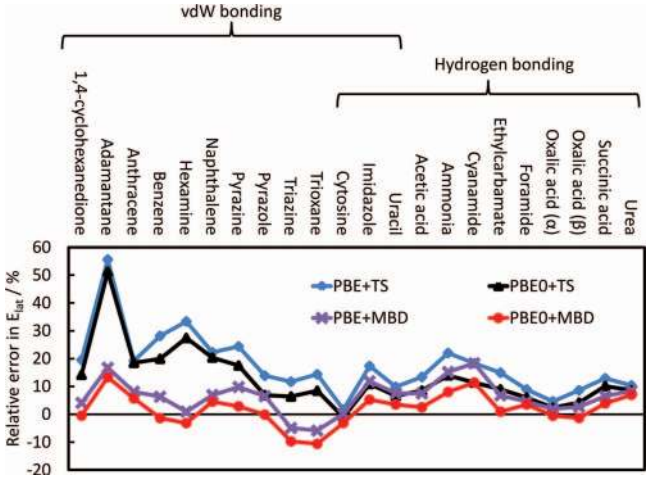

FIG. 4. Relative error in the lattice energy prediction of 22 molecular crystals with four different combinations of vdW terms and functionals. As the relative error is calculated as $100\% \times (E_{\mathrm{lat,DFT}} - E_{\mathrm{lat,exp}})/E_{\mathrm{lat,exp}}$, a positive value indicates overbinding.

functionals are the same as seen with our smaller study of the X16 database of 16 molecular crystals:[34] the inclusion of MBD generally halves the MAE and standard deviation (SD) in the pairwise TS method, while the inclusion of a fraction of exact exchange shifts the lattice energies down, reducing the MAE but barely changing the SD. While with this larger database there is a slight increase in the MAE and SD compared to the X16, the MAE is still within the chemical accuracy target of 4.2 kJ/mol, and the SD within one estimate of experimental uncertainties, 4.9 kJ/mol.[65]

In absolute terms, the worst agreement is for adamantane, a globular molecule whose lattice energy is overestimated by 9.2 kJ/mol with PBE0+MBD, although this is a significant improvement on the PBE+TS value. Interestingly, hexamine, which also has a similar globular structure, has a much smaller absolute error of less than 2.8 kJ/mol. In relative terms, the worst agreement is for CO_2, with a relative error of 25%. This value is very sensitive to any contributions from MBD and PBE0 as the experimental lattice energy is the smallest in the database at −28.4 kJ/mol. Possible origins for this large deviation will be discussed in Sec. III D.

2. Many-body dispersion contributions

It is evident from Table IV that the role of MBD is to reduce lattice energies, with positive values for the contribution of MBD to the lattice energy in all cases bar cyanamide, where there is practically no difference between the lattice energies with MBD and TS. Many previous studies at capturing many-body effects in dispersion have employed some variation on the Axilrod-Teller-Muto (ATM) three-body dispersion interaction.[67] It is well known that the contribution of the three-body term to cohesive energies or lattice energies is normally repulsive.[68] In their recent work Otero-de-la-Roza and Johnson employed an XDM-based ATM term for 21 of the 23 systems studied here.[69] They found that the ATM term had only relatively small contributions to the total dispersion energies and marginal contributions to lattice energies, although this may in part stem from the damping functions employed. Overall the ATM terms *increased* the MAE marginally for their dataset. This is in contrast to the MBD contributions to the 23 systems studied here. Risthaus and Grimme have also shown recently that an ATM term can give significant improvements for the modeling of supramolecular complexes.[70]

It is clear from the results here that many-body contributions for molecular crystals are significant, not just for the relative energies but also for the absolute value of the dispersion energy, e.g., for the acetic acid crystal the pairwise energy is −39.6 kJ/mol, while the MBD energy is −45.1 kJ/mol, an increase of 12%. Even for adamantane, in which MBD reduces the lattice energy by 26.4 kJ/mol, the MBD energy of the lattice is 29.0 kJ/mol *larger* than the pairwise energy. However, the corresponding MBD energy of the isolated molecules are also substantially larger and in relative terms the isolated molecules become more stable, reducing lattice energies.

These different effects, increases in absolute energies, decreases in relative energies, stem primarily from polarization effects. MBD considers not just many-body energies, like the ATM expression, but also explicitly couples the various dipole interactions, leading to a number of different types of polarization effects within the molecules and solids. The first effect is short-range intramolecular polarization, whereby the polarizability of an atom increases appreciably along covalent chemical bonds. Obtaining C_6 coefficients from the self-consistently screening polarizabilities can illustrate this effect, with, for example, the C-atom C_6 of β oxalic acid increasing from 34 a.u. with TS to 43.0 a.u. This increase in polarizability will increase the dispersion energy, but it is present in both the solid and the isolated molecules.

The second contribution is long-range de-polarization or screening of the dispersion interactions. In a solid, the presence of many atoms and molecules in between two (or more) interacting atoms will typical reduce the polarizabilty. We have noted previously that the screened C-atom C_6 coefficients of form-I aspirin are 10% smaller than those of the isolated molecules.[34] The MBD Hamiltonian will also give additional long-range screening as it is equivalent to a random-phase approximation treatment of the model coupled dipole-oscillators system it considers.[49] The third type of polarization effect is intermolecular polarization. Hydrogen-bonded solids have strong interactions in the solid state, through which further polarization of atoms can occur. As such polarization is missing in the isolated molecules, any such additional polarization will increase the polarizability and hence relative stability of the solid form. Competition between short-range intermolecular polarization and long-range de-polarization may explain the more complex behavior of MBD for the hydrogen-bonded solids. All of the vdW systems show appreciable screening effects, with large reductions in the lattice energy, whereas some hydrogen-bonded systems (e.g., cyanamide and cytosine) show barely any change. The importance of polarization effects has also recently been demonstrated using solely the SCS equations by Bučko and co-workers.[50]

These competing effects are illustrated for oxalic acid in Figure 5. The ellipsoids represent the anisotropic

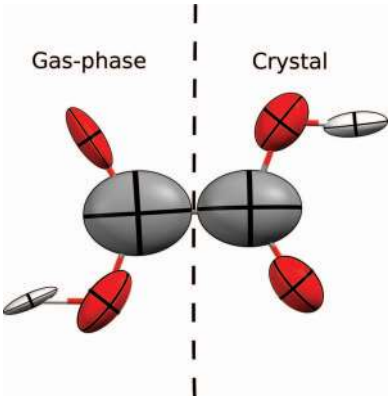

FIG. 5. Screened anisotropic polarizability tensors for a molecule of the β oxalic acid crystal structure without and with solid-state interactions. In β oxalic acid the molecules form chains of hydrogen-bonded dimers. Note that the correct gas-phase conformation of oxalic acid has intramolecular hydrogen bonds.[43]

polarizability tensors obtained from the self-consistent screening equations, with half of the molecule shown with the solid-state tensors and half with the gas-phase ones. A C_2 axis relates the two parts. The intramolecular polarization is evident with highly anisotropic polarizabilities directed primarily along chemical bonds, although in the gas phase the H atoms point also towards the more polarizable C atoms. In standard pairwise methods all of the atoms would have spherical polarizabilities. Examining the C-atom tensors we can see that the gas-phase one is slightly larger ($C_6 = 43.0$ a.u. in the solid and 45.5 a.u. in the isolated molecule), partly illustrating the long-range screening effect. Going from the gas phase to the solid state there is a significant re-orientation of the H-atom tensor, shifting from being orientated towards the C atom, to being orientated along the hydrogen bonds that are formed in the solid state.

3. Role of exact exchange

Inclusion of 25% exact exchange through the PBE0 functional has a small but still important contribution to the lattice energy. In general the contribution of PBE0 is to shift the lattice energies down, as can be seen from the reduction in the mean error (ME) and the slight change in the standard deviation. The individual contributions vary between 1 and 5 kJ/mol, with many of the hydrogen-bonded and lone-pair systems having a larger contribution. However, even benzene has a large change in its lattice energy of the order of 8%, which likely stems from the better modeling of its multipole moment and density with a hybrid functional.[71] The primary reason for applying hybrid functionals for cohesive properties is to correct for artefacts due to electron self-interaction or de-localisation errors.[31] Such errors are less significant in the solid state, where there is natural de-localisation, especially in hydrogen-bonded systems. The role of PBE0 is therefore most likely related to reducing de-localisation errors in isolated molecules, stabilising them relative to the solid. The comparable relative contribution of MBD and exact exchange for many system suggests that caution must be employed when fitting or scaling dispersion contributions. As both contributions reduce the lattice energy, spurious and inconsistent results may be obtained by fitting or scaling pairwise dispersion terms for semi-local functionals to experimental or benchmark theoretical data.

Hybrid functionals are often not used in the study of cohesive properties of molecular crystals,[29,30] largely due to their additional computational cost, which particularly in a plane-wave basis can reach more than an order of magnitude larger than the corresponding semi-local functional. The present database of crystals suggests that hybrids may be important for accurate work. Indeed, the correct polymorphic ordering of the polymorphs of oxalic acid and glycine is only recovered when coupling MBD with PBE0.[72] In their study of 21 molecular crystals Otero-de-la-Roza and Johnson employed the B86b functional,[30] as it best mimicked the Hartree-Fock repulsive wall when compared to other semi-local functionals.[33] However, B86b only marginally improves over PBE by 0.6 kJ/mol, whereas the hybrid functional used here (PBE0) yields contributions of around 3 kJ/mol.

The choice of PBE0 is driven by its non-empirical nature. While there is no definitive value for the amount of exact exchange that should be added to a hybrid, there is some motivation for a value of 25%,[73] and the important aspect of PBE0 (and PBE) is that it is not fitted to any experimental or theoretical data and is therefore not artificially biased towards or against specific systems. PBE0 also has a small short-range many-body correlation error,[74] which allows it to cleanly couple with the MBD method.

There are numerous other exchange functionals available, both with and without exact exchange and different fractions thereof. Indeed, recently it has been suggested that a fraction of 1/3 exact exchange should be used as the PBE-based hybrid functional.[75] In addition, screened hybrids such as the Heyd, Scuseria, and Ernzerhof (HSE) functional[76] might be more computationally efficient and accurate in some cases, although their effect on molecular-crystal lattice energies is largely unknown. It is likely that for many of the hydrogen-bonded systems the lattice energy could be tuned with differing amounts of exact exchange and/or screening, as different systems have different hydrogen-bond strengths and therefore de-localisation. We refrain from such a study in the present work, in part due to the substantial computational cost of hybrid functionals in a plane-wave implementation. Crystal-structure prediction and material design can also only work when computational methods are fully transferable to different geometries and even compositions, making an *ab initio* approach essential.

4. Comparison with (effective) pairwise treatments

The results presented in Subsections III B 1 and III B 2 show that the MBD approach represents a physically motivated and systematic improvement over the TS pairwise method. There are a number of other pairwise or two-body approaches reported in the literature. Using the data for 21 crystals of Otero-de-la-Roza and Johnson, we compare PBE0+MBD with XDM and the vdW-DF2 functional in Table VI. To make a direct comparison it should be noted that for oxalic acid the gas-phase structure was assumed

TABLE VI. The ME, MAE, SD, and MARE of different vdW-inclusive DFT methods for the lattice energies of the 21 molecule crystals of Otero-de-la-Roza and Johnson.[30] Note that an approximate correction has been applied to the oxalic acid lattice energies of Ref. 30 and that the reference lattice energies from the present work (Table IV) have been employed for calculating the different metrics. The ME, MAE, and SD are in kJ/mol, while the MARE is in (%).

	ME	MAE	SD	MARE
B86b+XDM	−3.47	4.71	5.49	5.97
vdW-DF2	4.80	6.40	5.77	8.49
PBE0+MBD	1.17	3.92	4.87	5.69

to have the solid-state conformation in their work. At the PBE and PBE0 levels (with and without dispersion contributions) the difference between the correct conformer and the solid-state conformation is approximately 21.5 kJ/mol and a correction of this amount has been applied. As this is averaged over 21 systems variations in this correction due to the different functionals are liable to be minimal. Of the three methods PBE0+MBD performs the best for all four metrics. The MAE of B86b+XDM is still quite good at 4.71 kJ/mol. However, with a negative ME, correctly adding exact-exchange or many-body contributions would make the XDM performance worse, as has in part been seen with the use of the ATM term.[69] The vdW-DF2 functional performs well compared to the TS pairwise method but still appreciably overbinds the crystals, almost twice as much as PBE0+MBD. This overbinding is surprising given the known underestimation of small-molecule C_6 coefficients by vdW-DF2.[77]

That an effective pairwise approach can give seemingly good performance for molecular crystals is no surprise. There are numerous studies of condensed phases of molecular systems employing empirical potentials. Fitting or scaling of C_6 parameters for *specific* systems can often give good results but their transferability and suitability for arbitrary molecular configurations is a concern.[78] Given that the overall role of MBD is to reduce the lattice energy, it is natural to investigate to what extend could a scaled version of the TS method capture this behavior. Fitting the expression $E_{\text{lat}} = (E_{\text{lat,PBE0}} + A_{\text{scr.}} E_{\text{TS,solid}} - E_{\text{TS,gas}})$ to the experimental lattice energies yields a pairwise screening coefficient of $A_{\text{scr.}} = 0.85(2)$ and a MAE of 4.54 kJ/mol.

While a scaled version of TS might yield good predictions for lattice energies, there are a number of situations where the fine detail of many-body dispersion is essential. The lattice energy of cytosine is predicted well by both TS and MBD, likely due to cancellation of the different polarization effects in MBD. Scaling the TS value therefore gives a very poor prediction, in error by 14 kJ/mol. Constant scaling would also not affect the relative ordering of polymorphs, the correct ordering of which is the critical challenge for crystal-structure prediction. In the present database, TS and other pairwise approaches[30] systematically overestimate the stability of β oxalic acid, with near degeneracy only restored at the PBE0+MBD level.[34,72] The importance of including many-body effects goes beyond cohesive energies; it has been recently shown that the prediction of the dielectric constants of the polyacene crystals is also significantly improved by the MBD method.[63]

C. Role of MBD in crystal structure geometries

All of the lattice energy and phonon calculations present in Secs. III A and III B have employed PBE+TS optimized geometries. Compared to PBE without any vdW contributions, PBE+TS yields very good lattice parameters, often within a few percent of experimental values.[29,30] To understand the role of MBD in determining the structure of molecular crystals, the lattice parameters and internal geometries of a selection of molecular crystals have been optimized at the PBE+MBD level. The resulting lattice parameters are compared with PBE+TS and experiment in Table VII.

In many cases the MBD and TS lattice vectors are smaller than experimental values, with mean relative errors of −0.55% and −0.75% for TS and MBD, respectively. It is interesting to note that although MBD reduces the binding energy, it leads to a more compact structure than the stronger bound pairwise method. None of the systems studied are known to feature negative thermal expansion and therefore we expect the equilibrium optimized lattice parameters to be smaller than experimental values, as even at very low temperatures zero-point vibrations will lead to an expansion of the experimental cell, both in terms of contributions to the free-energy surface and anharmonic averaging over the lattice vibrations. In non-cubic systems variations in the orientation of the molecules will also lead to changes in the cell parameters and their relative sizes. Quantifying how much smaller the lattice parameters should be in the absence of zero-point and thermal contributions is difficult. For simple systems the expansion contribution from quasi-harmonic effects can be readily determined using phonon calculations,[51,79] but for orthorhombic or monoclinic systems accurate calculations would be much more demanding. Estimation of the anharmonic contributions would also require *NPT* molecular-dynamics simulations, which would be similarly demanding, especially at low temperatures, where path-integral methods would be required. However, the ZPE contributions to covalently bound cubic solids are known to be of the order of 0.25%–0.5% of the lattice constant.[79] For weaker-bound molecular crystals values around 1% cannot be seen as surprising, and this is consistent with the range of values observed here.

In a few cases, MBD yields lattice vectors longer than experiment. TS also overestimates the length of the benzene c lattice vector. The overestimated lattice vectors in acetic acid and pyrazine both occur for relatively short lattice vectors, with sheets or stacking of molecules occurring in these directions. In particular, the c vector of pyrazine corresponds to the direction in which the molecules stack in sheets, tilted at an angle. Such geometries can lead to repulsive three-body interactions.[68] Over such small distances the repulsive contributions likely dominate the other attractive interactions. This may be sensitive to the range-separation parameter used to match MBD to the DFT functional, which is fitted solely

TABLE VII. The unique equilibrium lattice parameters and unit-cell volumes of 6 molecular crystals obtained using PBE+TS and PBE+MBD optimizations, together with relative errors and experimental values determined at the temperature $T_{exp.}$. Zero-point vibrational contributions are not included in the unit-cell relaxations but are known for simple cubic covalently bound solids to amount 0.25%–0.5% of the lattice constant.[79] Lattice parameters are in (Å) and (°), while volumes are in (Å^3) and relative quantities are in (%).

Molecule	$T_{exp.}$ (K)		PBE+TS	ΔPBE+TS	PBE+MBD	ΔPBE+MBD	Exp.
Naphthalene[80]	10	a	8.109	0.30	7.970	− 1.42	8.0846
		b	5.884	− 0.90	5.868	− 1.17	5.9375
		c	8.660	0.30	8.570	− 0.74	8.6335
		β	124.05	− 0.50	123.11	− 1.26	124.67
		V	342.32	0.44	335.70	− 1.50	340.83
Ammonia[81]	2	a	4.962	− 1.70	4.987	− 1.21	5.048
		V	122.18	− 5.02	124.03	− 3.58	128.63
Urea[82]	12	a	5.559	− 0.11	5.508	− 1.03	5.565
		c	4.683	− 0.02	4.662	− 0.47	4.684
		V	144.71	− 0.24	141.42	− 2.51	145.06
Benzene[83]	4	a	7.368	0.24	7.307	− 0.60	7.351
		b	9.225	− 1.48	9.293	− 0.76	9.364
		c	6.839	2.15	6.814	1.78	6.695
		V	464.88	0.87	462.69	0.40	460.84
Acetic acid[84]	40	a	13.157	0.05	12.988	− 1.24	13.151
		b	3.907	− 0.40	4.005	2.10	3.923
		c	5.712	− 0.86	5.686	− 1.31	5.762
		V	293.66	− 1.21	295.82	− 0.49	297.27
Pyrazine[85]	184	a	9.292	− 0.35	9.181	− 1.54	9.325
		b	5.717	− 2.28	5.689	− 2.74	5.850
		c	3.683	− 1.33	3.763	0.80	3.733
		V	195.66	− 3.92	196.55	− 3.48	203.64

considering a database of 22 small gas-phase dimers.[35] However, the differences between PBE and PBE+TS/MBD dwarf those between TS and MBD alone. Both methods capture the vast majority of the contribution of dispersion, and given its far superior prediction of lattice energies, MBD is clearly a better choice for DFT calculations of molecular crystals.

D. Higher-order multipole contributions to van der Waals interactions

The many-body dispersion method accounts for all orders of dipole interactions in a seamless fashion. While these dipole interactions represent the leading order contribution to the dispersion energy, a number of two-body methods also employ higher-order dipole-quadrupole (C_8) and quadrupole-quadrupole (C_{10}) interactions in the pairwise energy expansion.[18,20] These higher-order terms are much more short-ranged than the leading order dipole-dipole term and can become correlated with the semi-local functional, making their parameterization difficult.[18] While several groups advocate the use of these higher-order terms, a number of studies suggest that with non-empirical semi-local functionals the leading-order dipole term is the most important in the pairwise expansion of the dispersion energy,[86,87] while for water-sulfate clusters the higher-order terms in XDM have been found to worsen the prediction of binding energies.[88] The present work, focusing on ambient-pressure crystal structures, strengthens the suggestion that the dipole contributions are dominant when they are coupled with non-empirical DFT functionals.

At higher pressures or for very close-packed structures, atoms may approach at much shorter distances, making the short-range multipole terms more important for describing their properties. This has been seen for the high-pressure phases of ice, which cannot be sufficiently modeled using only a dipole-dipole term.[32] In the present database we believe CO_2 to be an example of a close-packed structure where multipoles are important. The lattice has high symmetry ($Pa3$) and is essentially fcc packing of angled ellipsoids. The TS method yields a reasonable prediction for the lattice energy, however when many-body contributions are accounted for, the lattice energy is underestimated by nearly 7 kJ/mol (PBE0+MBD), a relative error of 25%. The TS lattice parameter for CO_2 is 5.790 Å, while the experimental value, extrapolated to 0 K, is 5.55 Å.[89] Even with a lattice constant 4.3% too large all of the CO nearest neighbor distances (3.21 Å) are within the CO vdW radius of 3.32 Å. Optimizing the lattice with MBD yields little improvement with $a = 5.786$ Å. With such high symmetry, many of the dipole contributions cancel out, making an accurate treatment of the multipole terms important. We note though that the XDM method, which does include C_8 and C_{10} terms in the pairwise expansion, yields a very similar lattice constant to TS,[30] suggesting that such an accurate treatment is not trivial. Indeed, as they are so short ranged in nature, the higher-order effects might be better represented by using improved functionals. For instance, it is known that PBE0 has smaller short-range errors for interactions in van der Waals systems compared to PBE.[74]

Systems such as CO_2 and adamantane clearly require additional contributions, whether higher-order multipole terms or improved density functionals and they therefore are

useful targets for method development. However, the excellent performance of PBE0+MBD for the remainder of the X23 database demonstrates that it is a powerful approach for predicting lattice energies and stabilities of organic molecular crystals.

IV. CONCLUSION

In the present work, a detailed assessment of the different contributions to the lattice energies of molecular crystals has been performed. The vibrational contributions are not strictly contributions to the lattice energy, but rather are important for converting experimental sublimation enthalpies into lattice energies that can be compared with theory. In the harmonic approximation, supercell phonon calculations show significant deviations from the widely used Dulong-Petit law, as noted elsewhere.[30] A semi-anharmonic treatment using experimental heat-capacity data suggests that the anharmonic contributions are not negligible and care should be taken in comparing experimental lattice energies with DFT values.

The contribution of many-body dispersion and exact exchange to the lattice energy has been assessed for a database of 23 molecular crystals, building on our previous work with 16 systems.[34] Both contributions are found to be important. The use of MBD as compared to the TS pairwise method reduces the MAE by more than half from 10.04 kJ/mol with PBE0+TS to 3.92 kJ/mol with PBE0+MBD. Using a single scaling parameter, the pairwise approach can reach within 0.6 kJ/mol of the many-body lattice energies. However, such scaling will not correct any systematic deviations, such as the incorrect ordering of polymorph stabilities. The use of exact exchange has a smaller but still important contribution of the order of 3 kJ/mol, reducing the impact of de-localisation errors, particularly in hydrogen-bonded systems.

The role of MBD in determining the lattice parameters has also been explored for a selection of systems using finite-difference forces and stresses. Compared to PBE, which is known to often substantially overestimate lattice constants,[29] both the pairwise and many-body approaches give good agreement with experiment, with many of the differences likely to be of the order of zero-point and thermal vibrational contributions to the lattice constants. The estimation of these contributions for any crystal is non-trivial and will form part of our future work in this area.

Overall, it has been shown that assessing predictions of lattice energies requires careful consideration of vibrational, many-body dispersion and exact-exchange contributions. Neglecting any of these contributions would lead to an overestimation of the importance of other contributions, hindering efforts to develop more accurate and physically motivated computational approaches. While the need for exact-exchange functionals may at present pose computational difficulties, developments in localized-basis set codes[90] and screened-exchange functionals[76] should provide a way forward. The excellent performance of MBD, both for lattice energies and structures, suggests that it can be a powerful tool for the investigation of structure-properties relationships of molecular crystals. Ongoing development of analytical forces and stresses should enable us to fully explore the role of MBD in governing geometries and phonons, as well as related properties. Studies at high pressure should also lead to a better understanding of multipole terms in the dispersion energy.

ACKNOWLEDGMENTS

The authors acknowledge support from the European Research Council (ERC Starting Grant VDW-CMAT).

[1] N. Blagden, M. de Matas, P. Gavan, and P. York, Adv. Drug Delivery Rev. **59**, 617 (2007).
[2] S. L. Price, Adv. Drug Delivery Rev. **56**, 301 (2004).
[3] C. M. Reddy, G. Rama Krishna, and S. Ghosh, Cryst. Eng. Comm. **12**, 2296 (2010).
[4] D. M. S. Martins, D. S. Middlemiss, C. R. Pulham, C. C. Wilson, M. T. Weller, P. F. Henry, N. Shankland, K. Shankland, W. G. Marshall, R. M. Ibberson, K. Knight, S. Moggach, M. Brunelli, and C. A. Morrison, J. Am. Chem. Soc. **131**, 3884 (2009).
[5] S. Karki, T. Friščcí, L. Fábián, P. R. Laity, G. M. Day, and W. Jones, Adv. Mater. **21**, 3905 (2009).
[6] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti, and I. K. Zhitkov, Acta Crystallogr., Sect. B: Struct. Sci. **67**, 535 (2011).
[7] S. L. Price, Acc. Chem. Res. **42**, 117 (2009).
[8] D. C. Sorescu and B. M. Rice, J. Phys. Chem. C **114**, 6734 (2010).
[9] S. Hunter, T. Sutinen, S. F. Parker, C. A. Morrison, D. M. Williamson, S. Thompson, P. J. Gould, and C. R. Pulham, J. Phys. Chem. C **117**, 8062 (2013).
[10] C. A. Morrison, M. M. Siddick, P. J. Camp, and C. C. Wilson, J. Am. Chem. Soc. **127**, 4042 (2005).
[11] A. M. Reilly, D. S. Middlemiss, M. M. Siddick, D. A. Wann, G. J. Ackland, C. C. Wilson, D. W. H. Rankin, and C. A. Morrison, J. Phys. Chem. A **112**, 1322 (2008).
[12] A. M. Reilly, S. Habershon, C. A. Morrison, and D. W. H. Rankin, J. Chem. Phys. **132**, 094502 (2010).
[13] S. Sharifzadeh, A. Biller, L. Kronik, and J. B. Neaton, Phys. Rev. B **85**, 125307 (2012).
[14] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).
[15] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).
[16] J. Klimeš and A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).
[17] S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
[18] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[19] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[20] A. D. Becke and E. R. Johnson, J. Chem. Phys. **127**, 154108 (2007).
[21] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004); T. Thonhauser, V. R. Cooper, S. Li, A. Puzder, P. Hyldgaard, and D. C. Langreth, Phys. Rev. B **76**, 125112 (2007).
[22] K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, Phys. Rev. B **82**, 081101 (2010).
[23] S. Wen and G. J. O. Beran, J. Chem. Theory Comput. **8**, 2698 (2012).
[24] G. J. O. Beran and K. Nanda, J. Phys. Chem. Lett. **1**, 3480 (2010).
[25] L. Maschio, B. Civalleri, P. Ugliengo, and A. Gavezzotti, J. Phys. Chem. A **115**, 11179 (2011).
[26] Z. Zheng, J. Zhao, Y. Sun, and S. Zhang, Chem. Phys. Lett. **550**, 94 (2012).
[27] J. C. Sancho-García, J. Aragó, E. Ortí, and Y. Olivier, J. Chem. Phys. **138**, 204304 (2013).
[28] A. Gavezzotti and G. Filippini, "Energetic aspects of crystal packing: experiment and computer simulations," in *Theoretical Aspects and Computer*

Modeling of the Molecular Solid State (Wiley and Sons, Chichester, 1997), pp. 61–97.
[29]W. A. Al-Saidi, V. K. Voora, and K. D. Jordan, J. Chem. Theory Comput. **8**, 1503 (2012).
[30]A. Otero-de-la-Roza and E. R. Johnson, J. Chem. Phys. **137**, 054103 (2012).
[31]A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).
[32]B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).
[33]A. Otero-de-la-Roza and E. R. Johnson, J. Chem. Phys. **136**, 174109 (2012).
[34]A. M. Reilly and A. Tkatchenko, J. Phys. Chem. Lett. **4**, 1028 (2013).
[35]A. Tkatchenko, R. A. DiStasio, Jr., R. Car, and M. Scheffler, Phys. Rev. Lett. **108**, 236402 (2012).
[36]R. A. DiStasio, Jr., O. A. von Lilienfeld, and A. Tkatchenko, Proc. Natl. Acad. Sci. U.S.A. **109**, 14791 (2012).
[37]C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[38]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[39]S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Z. Kristallogr. **220**, 567 (2005).
[40]The choice of ultra-soft pseudo-potentials was motivated by initial computational difficulties that have since been resolved such that ultra-soft and norm-conserving pseudo-potentials give very similar results. All-electron calculations have also been used to validate selected systems.
[41]E. R. McNellis, J. Meyer, and K. Reuter, Phys. Rev. B **80**, 205414 (2009).
[42]F. H. Allen, Acta Crystallogr., Sect. B: Struct. Sci. **58**, 380 (2002).
[43]S. A. Blair and A. J. Thakkar, Chem. Phys. Lett. **495**, 198 (2010).
[44]N. Vogt, M. A. Abaev, A. N. Rykov, and I. F. Shishkov, J. Mol. Struct. **996**, 120 (2011).
[45]M. A. R. da Silva, M. J. Monte, and J. R. Ribeiro, J. Chem. Thermodyn. **33**, 23 (2001).
[46]See supplementary material at http://dx.doi.org/10.1063/1.4812819 for an archive file containing the PBE+TS optimized solid-state and gas-phase structures of the 23 systems, along with the PBE Hirshfeld volumes.
[47]A. Mayer, Phys. Rev. B **75**, 045407 (2007).
[48]B. T. Thole, Chem. Phys. **59**, 341 (1981).
[49]A. Tkatchenko, A. Ambrosetti, and R. A. DiStasio, Jr., J. Chem. Phys. **138**, 074106 (2013).
[50]T. Bučko, S. Lebègue, J. Hafner, and J. G. Ángyán, Phys. Rev. B **87**, 064110 (2013).
[51]S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).
[52]W. Frank, C. Elsässer, and M. Fähnle, Phys. Rev. Lett. **74**, 1791 (1995).
[53]K. Parlinski, Z. Q. Li, and Y. Kawazoe, Phys. Rev. Lett. **78**, 4063 (1997).
[54]J. S. Chickos and W. E. Acree, J. Phys. Chem. Ref. Data **31**, 537 (2002).
[55]M. V. Roux, M. Temprado, J. S. Chickos, and Y. Nagano, J. Phys. Chem. Ref. Data **37**, 1855 (2008).
[56]M. Johnson, K. Parlinski, I. Natkaniec, and B. Hudson, Chem. Phys. **291**, 53 (2003).
[57]S.-S. Chang and E. F. Westrum, J. Phys. Chem. **64**, 1547 (1960).
[58]P. Goursot, H. L. Girdhar, and E. F. Westrum, J. Phys. Chem. **74**, 2538 (1970).
[59]J. P. McCullough, H. L. Finke, J. F. Messerly, S. S. Todd, T. C. Kincheloe, and G. Waddington, J. Phys. Chem. **61**, 1105 (1957).
[60]C. E. Vanderzee and E. F. Westrum, Jr., J. Chem. Thermodyn. **2**, 681 (1970).
[61]M. V. Bommel, J. V. Miltenburg, and A. Schuijff, J. Chem. Thermodyn. **20**, 397 (1988).
[62]O. Andersson, T. Matsuo, H. Suga, and P. Ferloni, Int. J. Thermophys. **14**, 149 (1993).
[63]B. Schatschneider, J.-J. Liang, A. M. Reilly, N. Marom, G.-X. Zhang, and A. Tkatchenko, Phys. Rev. B **87**, 060104 (2013).
[64]See http://webbook.nist.gov/chemistry/name-ser.html for "NIST Web-Book, and references within" (accessed Feb. 2013).
[65]J. S. Chickos, Netsu Sokutei **30**, 125 (2003).
[66]H. de Wit, J. V. Miltenburg, and C. D. Kruif, J. Chem. Thermodyn. **15**, 651 (1983).
[67]B. M. Axilrod and E. Teller, J. Chem. Phys. **11**, 299 (1943).
[68]O. A. von Lilienfeld and A. Tkatchenko, J. Chem. Phys. **132**, 234109 (2010).
[69]A. Otero-de-la-Roza and E. R. Johnson, J. Chem. Phys. **138**, 054103 (2013).
[70]T. Risthaus and S. Grimme, J. Chem. Theory Comput. **9**, 1580 (2013).
[71]E. J. Meijer and M. Sprik, J. Chem. Phys. **105**, 8684 (1996).
[72]N. Marom, R. A. DiStasio, Jr., V. Atalla, S. Levchenko, A. M. Reilly, J. R. Chelikowsky, L. Leiserowitz, and A. Tkatchenko, Angew. Chem., Int. Ed. **52**, 6629 (2013).
[73]J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
[74]A. Tkatchenko and O. A. von Lilienfeld, Phys. Rev. B **78**, 045116 (2008).
[75]C. A. Guido, E. Brémond, C. Adamo, and P. Cortona, J. Chem. Phys. **138**, 021104 (2013).
[76]J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); **124**, 219906 (2006).
[77]O. A. Vydrov and T. Van Voorhis, Phys. Rev. A **81**, 062708 (2010).
[78]A. J. Stone, Science **321**, 787 (2008).
[79]P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew, Phys. Rev. B **85**, 014111 (2012).
[80]S. C. Capelli, A. Albinati, S. A. Mason, and B. T. M. Willis, J. Phys. Chem. A **110**, 11695 (2006).
[81]A. W. Hewat and C. Riekel, Acta Crystallogr., Sect. A: Found. Crystallogr. **35**, 569 (1979).
[82]S. Swaminathan, B. M. Craven, and R. K. McMullan, Acta Crystallogr., Sect. B: Struct. Sci. **40**, 300 (1984).
[83]C. J. Craven, P. D. Hatton, C. J. Howard, and G. S. Pawley, J. Chem. Phys. **98**, 8236 (1993).
[84]R. Boese, D. Bläser, R. Latz, and A. Bäumen, Acta Crystallogr., Sect. C: Cryst. Struct. Commun. **55**, IUC9900001 (1999).
[85]G. de With, S. Harkema, and D. Feil, Acta Crystallogr., Sect. B: Struct. Sci. **32**, 3178 (1976).
[86]N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler, and L. Kronik, J. Chem. Theory Comput. **7**, 3944 (2011).
[87]S. N. Steinmann and C. Corminboeuf, J. Chem. Theory Comput. **8**, 4305 (2012).
[88]N. Mardirossian, D. S. Lambrecht, L. McCaslin, S. S. Xantheas, and M. Head-Gordon, J. Chem. Theory Comput. **9**, 1368 (2013).
[89]W. Keesom and J. Köhler, Physica **1**, 655 (1934); A. Curzon, *ibid.* **59**, 733 (1972).
[90]V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).