STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

ISSN: 2052-5206
journals.iucr.org/b

Organic crystal polymorphism: a benchmark for dispersion-corrected mean-field electronic structure methods

Jan Gerit Brandenburg and Stefan Grimme

Acta Cryst. (2016). B**72**, 502–513

Acta Cryst B
STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS
ISSN 2052-5206

Organic crystal polymorphism: a benchmark for dispersion-corrected mean-field electronic structure methods

Jan Gerit Brandenburg*‡ and Stefan Grimme*

Mulliken Center for Theoretical Chemistry, University of Bonn, Beringstrasse 4-6, 53115 Bonn, Germany.
*Correspondence e-mail: gerit.brandenburg@thch.uni-bonn.de, grimme@thch.uni-bonn.de

Received 12 February 2016
Accepted 13 May 2016

Edited by G. M. Day, University of Southampton, England

‡ Current address: Department of Chemistry, University College London, 20 Gordon Street, London WC1H0AJ, England.

Keywords: crystal structure prediction; London dispersion; density functional theory; polymorphism; benchmark.

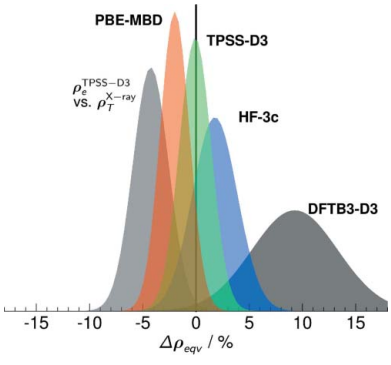
Supporting information: this article has supporting information at journals.iucr.org/b

We analyze the energy landscape of the sixth crystal structure prediction blind test targets with various *first principles* and *semi-empirical* quantum chemical methodologies. A new benchmark set of 59 crystal structures (termed POLY59) for testing quantum chemical methods based on the blind test target crystals is presented. We focus on different means to include London dispersion interactions within the density functional theory (DFT) framework. We show the impact of pairwise dispersion corrections like the semi-empirical D2 scheme, the Tkatchenko–Scheffler (TS) method, and the density-dependent dispersion correction dDsC. Recent methodological progress includes higher-order contributions in both the many-body and multipole expansions. We use the D3 correction with Axilrod–Teller–Muto type three-body contribution, the TS based many-body dispersion (MBD), and the nonlocal van der Waals density functional (vdW-DF2). The density functionals with D3 and MBD correction provide an energy ranking of the blind test polymorphs in excellent agreement with the experimentally found structures. As a computationally less demanding method, we test our recently presented minimal basis Hartree–Fock method (HF-3c) and a density functional tight-binding Hamiltonian (DFTB). Considering the speed-up of three to four orders of magnitudes, the energy ranking provided by the low-cost methods is very reasonable. We compare the computed geometries with the corresponding X-ray data where TPSS-D3 performs best. The importance of zero-point vibrational energy and thermal effects on crystal densities is highlighted.

1. Introduction

The polymorphism of molecular crystals is important in various areas of chemistry and physics with possible applications for pharmaceutical compounds, pigments, explosives and metal–organic framework materials (Price, 2014; Mooij *et al.*, 2000). There are many examples of well known generic drugs such as aspirin and paracetamol with multiple polymorphs (Buvar *et al.*, 2015). In order to release a new drug, pharmaceutical companies have to screen the polymorph landscape. Those experimental screenings are rather expensive and cannot guarantee to cover the complete polymorph space. Simulation techniques can help and guide experimentalists by supplying them with computed crystal energy landscapes. This crystal structure prediction (CSP) task is sketched in Fig. 1. The most stable crystal structures shall be predicted from knowledge of their molecular composition (Lewis structure). The computed (possibly free) energy landscape can then be used to identify the measured structures and to suggest plausible alternative polymorphs.

The Cambridge Structural Database organizes CSP blind challenges every few years to test and trigger methodological



progress (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005, 2009; Bardwell *et al.*, 2011). In 2015 the sixth blind test of organic crystal structure prediction methods took place following the same general spirit (Reilly *et al.*, 2016). In contrast to previous blind tests, each group could submit two lists of 100 structures in order to obtain a better picture of the computed energy landscapes. Furthermore, the molecular targets were more complicated, including a hydrated salt and some rather flexible and large molecules. In this regard the targets are close to relevant 'real life' applications. More details on the general blind test procedure and discussion of the different systems and methodological approaches can be found in the main report (Reilly *et al.*, 2016).

A variety of methods for generating, optimizing and ranking possible polymorph candidates were applied. Since an increasing number of density functional approximations (DFAs) were used in the final energy evaluation, a critical analysis of their performance is mandatory. A number of studies benchmark DFAs for their ability to calculate lattice energies (Otero-de-la-Roza & Johnson, 2012; Reilly & Tkatchenko, 2013; Carter & Rohl, 2014; Moellmann & Grimme, 2014; Brandenburg & Grimme, 2014*b*; Brandenburg *et al.*, 2015). To what extent the lowest achieved mean absolute deviation (MAD) in these works of ~ 4 kJ mol^{-1} for absolute lattice energies transfers to a corresponding accuracy for relative energy differences of crystal polymorphs has to be studied in more detail. Early investigations of typical organic crystal structure prediction blind test targets analyzed the balance between intra- and intermolecular interactions (Karamertzanis *et al.*, 2008). A few studies used more recent method developments to analyze the polymorphism of glycine or aspirin (Wen & Beran, 2012; Marom *et al.*, 2013; Reilly & Tkatchenko, 2014; Beran *et al.*, 2014), but this analysis has to be done more rigorously for an extended set of data. A comprehensive overview of the previously investigated polymorph stabilities is given in a recent review article (Beran, 2016).

Here, we use the generated crystal energy landscapes from the sixth blind test to compile a benchmark set of relative crystal stabilities. Because we are testing the ability of methods to compute relative crystal energies that should transfer to polymorph stabilities, this set will be dubbed POLY59. The set consists of all nine experimentally realised structures, which are complemented with the ten energetically lowest computed structures per target molecule (from our blind test submission). On this POLY59 set, several electronic structure methods including density functional tight-binding (DFTB3-D3; Brandenburg & Grimme, 2014*a*), minimal basis set Hartree-Fock (HF-3c; Sure & Grimme, 2013) and dispersion-corrected DFAs are evaluated. The main focus is on methods that include the London dispersion interaction within the density functional theory (DFT) framework. As the semi-classical dispersion corrections are most prominent, we compare several schemes that compute only the leading order contribution, *i.e.* C_6 only (Grimme, 2006; Tkatchenko & Scheffler, 2009; Steinmann *et al.*, 2009) with modern variants that include higher-order terms in both the multipole and the many-body sense (Grimme *et al.*, 2010; Tkatchenko *et al.*, 2012; Lee *et al.*, 2010). The accuracy of a method is judged by the ranking of the experimentally observed polymorph, which is expected to be the energetically lowest structure. This assumption is actually the main limitation of this study. On the one hand, a low-energy polymorph may exist that is experimentally not accessible due to high energy barriers. On the other hand, finite temperature contributions may be important that are not included in our rankings shown below. We try to estimate the methodological effects and give an explicit route to improve this in future calculations (§4.3).

We give a short overview of the applied electronic structure methods in §2, describe the compilation of the benchmark set in §3, and discuss the method performance in §4. While the discussion focuses on the energy ranking (§4.2), we will also present and discuss the equilibrium structures of the two most successful methods, namely TPSS-D3 and PBE-MBD (§4.1). In addition, the discussion includes an analysis of zero-point vibrational energy (ZPVE) and thermal contributions to both the energies and geometries.

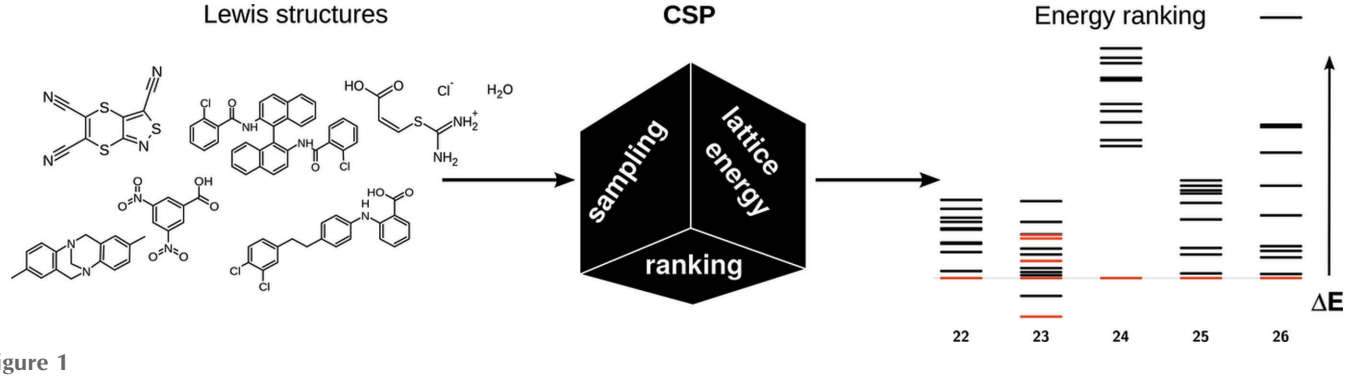


Figure 1
An exemplary crystal structure prediction (CSP) task is sketched with the blind test structures as target molecules. Solely knowing the Lewis structure CSP has to sample the polymorph space of the target molecule, calculate the corresponding structures and relative energies (possibly including thermal effects), and result in a realistic landscape of possible structures. In combination with experiment, the measured polymorphs can be identified and possible alternative structures can be proposed.

2. Methodology

2.1. Dispersion-corrected mean-field methods

Effective mean-field electronic structure methods are an irreplaceable tool in computational chemistry and solid-state physics. The most widely applied are density functional approximations (DFAs), but grid-free Hartree–Fock (HF) can have advantages too (Parr & Yang, 1989; Stone, 1997). As most prominent DFA, we apply the Perdew–Burke–Enzerhof (PBE) generalized gradient approximated exchange and correlation functionals (Perdew *et al.*, 1996) and the Tao–Perdew–Staroverov–Scuseria (TPSS) meta-GGA functional (Tao *et al.*, 2003). The DFAs are normally expanded in properly converged single-particle basis sets, and for periodic systems the projector-augmented plane wave (PAW) method of Blöchl and Kresse is very efficient (Blöchl, 1994; Kresse & Joubert, 1999). Recently, a minimal basis set Hartree–Fock method (expanded in Gaussian-type orbitals with basis set corrections) HF-3c has been developed that has a speedup of approximately one to two orders of magnitudes compared with converged DFA computations (Sure & Grimme, 2013). For further speedup of two orders of magnitudes, the minimal basis DFA energy expression can be expanded in charge density fluctuations, resulting in the density functional tight-binding (DFTB) approximation (Aradi *et al.*, 2007; Elstner, 2007).

All the above methods are not capable of describing London dispersion interactions (also known as attractive van der Waals forces). For an accurate description of organic crystals, accounting for these interactions is mandatory as recognized extensively in the last decade. For reviews or overviews on the 'dispersion problem in DFT', see Burns *et al.* (2011), Klimeš & Michaelides (2012) and Grimme (2011). Correction schemes to include the missing London dispersion have been developed, two recent review articles describe their different construction principles (Grimme *et al.*, 2016; Woods *et al.*, 2016).

The computation of the electronic lattice energy of a crystal can be decomposed into the electronic mean field (MF) and the dispersion part

$$E_{\rm tot} = E_{\rm MF} + E_{\rm disp}. \tag{1}$$

The leading order fluctuating-dipole–fluctuating-dipole contribution to the dispersion energy between two fragments A and B can be expressed *via* the Casimir–Polder relation

$$E_{\rm disp}^{AB} = -\frac{C_6^{AB}}{R_{AB}^6},$$
$$C_6^{AB} = \frac{3}{\pi}\int_0^\infty {\rm d}\omega\, \alpha_A(i\omega)\alpha_B(i\omega). \tag{2}$$

The typical $1/R_{AB}^6$ dependence on the distance R_{AB} is recovered and the corresponding C_6 coefficients can be computed by the dynamical polarizabilities $\alpha_{A/B}$ (Casimir & Polder, 1948). Dispersion corrections that use an atomic partitioning of the polarizabilities are nowadays classified as semi-classical schemes. Several dispersion corrections only compute this leading order term and mainly differ in the way the C_6 coefficients are estimated. We consider the DFT-D2 scheme (Grimme, 2006), the Tkatchenko–Scheffler TS correction (Tkatchenko & Scheffler, 2009), and the density dependent dDsC method (Steinmann *et al.*, 2009). While D2 uses an empirical C_6 estimate, the TS dispersion coefficients are based on atomic reference values that are scaled with relative Hirshfeld volumes, and in dDsC the local exchange dipole is integrated and related to the C_6.

Modern variants compute higher-order terms, namely higher pairwise contributions (dipole–quadrupole, dipole–octupole, quadrupole–quadrupole *etc.*) and higher many-body contributions (dipole–dipole–dipole *etc.*). Here, we mainly use the latest development from our group, the DFT-D3 scheme (Grimme *et al.*, 2010), that can be expressed as

$$E_{\rm disp}^{\rm (D3)} = -\frac{1}{2}\sum_{n=6,8}\sum_{A\neq B}^{\rm atom\ pairs}\frac{C_n^{AB}}{R_{AB}^n}f_n^d(R_{AB}) - \frac{1}{6}\sum_{A\neq B\neq C}^{\rm atom\ triples}\frac{C_9^{ABC}\left(1+3\cos\theta_A\cos\theta_B\cos\theta_C\right)}{R_{ABC}^9}f_9^d(R_{ABC}). \tag{3}$$

The damping functions f_n^d are introduced to combine the D3 dispersion interaction with the semilocal correlation contribution from the DFA (Grimme *et al.*, 2011). In contrast to the empirical D2 method, in D3 a modified Casimir–Polder integration [similar to equation (2)] yields the *ab initio* C_6^{AB} for reference systems (hydrogenated atoms). Higher-order dipole–quadrupole pair terms and dipole–dipole–dipole three-body terms of Axilrod–Teller–Muto type (Axilrod & Teller, 1943; Muto, 1943) are calculated *via* recursion relations and averages, respectively, from the corresponding C_6 coefficients. In the following, the D3 method is always applied in the above given form, *i.e.* including C_6, C_8 and C_9 contributions. Tkatchenko and coworkers extended the TS method in order to estimate the full many-body series in the dipole approximation, which is dubbed many-body dispersion MBD (Tkatchenko *et al.*, 2012). A modern method not considered here but worth mentioning is the exchange-dipole model (XDM) of Becke and Johnson introduced in a series of articles (Becke & Johnson, 2005*a*,*b*, 2006, 2007). The dDsC method is a modification of XDM with a density-dependent zero-damping function instead of the rational damping used in XDM (and D3). Similar to the D3 method, in XDM higher-order multipole contributions (C_8 and C_{10}) are included and an Axilrod–Teller–Muto-type three-body term (C_9) has been considered. The importance of the higher-order multipole and many-body terms is an active debate, see DiStasio *et al.* (2014), Dobson (2014), Kennedy *et al.* (2014) and Otero-de-la-Roza & Johnson (2015) for a detailed discussion of many-body dispersion effects.

While we typically prefer the TPSS DFA, all other semi-classical dispersion schemes are combined with the PBE functional. Note that the PBE functional typically overestimates simple hydrogen-bonded systems (Brandenburg *et al.*, 2015; Bryantsev *et al.*, 2009; Thanthiriwatte *et al.*, 2011), but

Table 1
Accuracy estimation of various methods to describe exchange–repulsion (EXR), electrostatic (ES), induction (IND) and dispersion (DISP).

The scale indicates (0) no contribution, (1) bad, (2) acceptable, (3) good and (4) very good accuracy.

DFA-DISP	EXR	ES	IND	DISP
PBE (+ semi-classical DISP)				
PBE	3	4	3	0
PBE-D2	3	4	3	2
PBE-TS	3	4	3	3
PBE-dDsC	3	4	3	3
PBE-D3	3	4	3	4
PBE-MBD	3	4	3	4
Empirical and non-local DFA				
M06L	3	4	3	2
vdW-DF2	3	4	3	4
D3-based hierarchical methods				
TPSS-D3	3	4	3	4
HF-3c	3	3	1.5	4
DFTB3-D3	1	2	1.5	4

this behavior is not systematically observed for systems with multiple strong hydrogen bonds as *e.g.* in supramolecular complexes (Sure & Grimme, 2015). In Table 1 we roughly compare the estimated accuracy of the different methods regarding the different intermolecular interactions, namely exchange–repulsion (EXR), electrostatic (ES), induction (IND) and dispersion (DISP). This accuracy estimation is based on empirical benchmark performances as summarized in Grimme *et al.* (2016), Goerigk & Grimme (2011), Peverati & Truhlar (2014) and Řezáč *et al.* (2011). Note that only approximate and mean relative trends are discussed. The table is intended for non-experts in the field of noncovalent interactions from electronic structure theory to help interpret the results presented here. For instance, a bad performance of plain PBE highlights missing DISP interactions, while large errors of DFTB3-D3 highlight the strong effects of ES and IND. The good error compensation between the various energy components for HF on a minimal basis has been analyzed before (Kruse & Grimme, 2012).

In addition to the semi-classical dispersion corrections, we apply the M06L meta-GGA (Zhao & Truhlar, 2006) and the vdW-DF2 (Lee *et al.*, 2010) methods. The Minnesota functional M06L covers medium range correlation by a very flexible meta-GGA functional form and empirical fitting to large training sets. In vdW-DF2 an approximation to the charge density susceptibility (that can be related to polarizability) is used to construct a nonlocal functional kernel that properly describes the long-range dispersion interaction. While vdW-DF2 has no many-body contributions, it integrates the true charge density and thus represents the dynamic multipoles to all orders. Because the single-point computation of electronic energies on the structures provided by Price and co-workers was one of the weakest points in our blind test submission, we test the transferability of optimized structures between the low-cost HF-3c method and the TPSS-D3 DFA. The TPSS-D3 single-point energies on a HF-3c structure will be dubbed TPSS-D3//HF-3c.

Table 2
Experimentally realised polymorphs from the sixth crystal structure prediction blind test.

No.	Space group	Z(unit)	$\rho_T^{\text{X-ray}}$†	ρ_e‡	T (K)
22-0	$P2_1/n$	4	1.754	1.772	150
23-0α	$P2_1/c$	4	1.351	1.420	299
23-0β	$P\bar{1}$	2	1.380	1.444	299
23-0γ	$P\bar{1}$	4	1.394	1.483	301
23-0δ	$P2_1/n$	4	1.335	1.406	301
23-0ϵ	$P\bar{1}$	4	1.350	1.420	292
24-0	$P2_1/c$	4	1.549	1.604	240
25-0	$P2_1/n$	4	1.402	1.471	298
26-0	$P\bar{1}$	2	1.346	1.408	298

† At temperature T measured mass density in g cm^{-3}. ‡ ZPVE exclusive mass density at T = 0 K in g cm^{-3} (estimated at the DFTB3-D3 level).

2.2. Generation of crystal energy landscapes

The generation of the TPSS-D3 crystal lattice energy landscapes is described in the supporting information of the blind test report (Reilly *et al.*, 2016), here we only give a short summary.

Our initial set of 1000 structures for each target was provided by the Price group (see the supporting information of Reilly *et al.*, 2016) based on the approach that uses electronic structure calculations on the molecule only (Price, 2009). It was successful for the large flexible molecule in the previous blind test (Kazantsev *et al.*, 2011). The program packages *CrystalPredictor* (Karamertzanis & Pantelides, 2007), *DMACRYS* (Price *et al.*, 2010) and *CrystalOptimizer* (Kazantsev *et al.*, 2010) had been used in their generation and refinement. The lattice energies used to determine the 1000 lowest energies used empirical exp-6 intermolecular potentials, atomic multipoles and intramolecular energies from PBE0 calculations on the isolated molecules in many conformations. We filtered these structures in three steps. First, we calculated a semi-empirical HF-3c (Sure & Grimme, 2013) single-point energy at the force-field structure. In the second step the energetically lowest structures are fully optimized at the HF-3c level. Finally, the energetically lowest HF-3c structures are reoptimized at the TPSS-D3 level.

3. Benchmark set

The target systems of the sixth crystal structure prediction blind test are summarized in Table 2 and shown in Fig. 2. We give the unit-cell symmetry, the measurement temperature and the crystal mass density from the X-ray measurement $\rho_T^{\text{X-ray}}$. We estimate the ZPVE and thermal contribution to the unit cell for each crystal at the DFTB3-D3 level (Brandenburg & Grimme, 2014*a*) and use these values to back correct the density $\rho_T^{\text{X-ray}}$ to ZPVE exclusive equilibrium (0 K) mass densities ρ_e. The values ρ_e can be used for direct comparison to results of geometry relaxations on the electronic energy surface for convenient benchmarking. Note that for target 23, five different polymorphic forms have been found (phase α-ϵ). For the other four targets (22, 24, 25, 26) only one polymorph has been measured, although other metastable phases might

exist. More details on the measurement conditions are given in the main blind test report (Reilly *et al.*, 2016).

We supplement the nine measured systems with ten additional low-energy structures per target molecule obtaining in total 59 benchmark systems. This set is dubbed POLY59 in the following. The additional systems are taken from our blind test submission and correspond to the energetically lowest structures computed at the TPSS-D3 level. In order to remove redundant structures, we use the Crystal Packing Similarity Tool available through *Mercury* 3.6 (Macrae *et al.*, 2008). The crystal structure is represented by a 15-molecule cluster which is considered identical when all molecules agree to within 25% of the distances, angles and triangles, and the overall root mean square deviation (RMSD) of the Cartesian coordinates is below 0.5 Å. The full POLY59 set is given in Table 3.

TPSS-D3 provides highly accurate geometries (as shown below). Therefore, the provided POLY59 structures can be used to calculate single-point energies with all methods under consideration without relaxing the geometry. As long as the potential energy surfaces of the methods are roughly parallel to the TPSS-D3 one, this approximation should have negligible effects on relative energies. Note that our original submission contains a few redundant structures. Additionally, our submitted structures have been symmetrized (*via PLATON*; Spek, 2009) that slightly modified the geometries. For consistent single-point energy evaluations of the POLY59 set the structures provided here have to be used. These are given in space group P_1 as freely optimized at the TPSS-D3 level.

Table 3
Summary of all systems contained in the POLY59 benchmark set.

The suffix zero indicates the experimentally found structure, the mass density (in g cm^{-3}) is estimated at the TPSS-D3 level.

No.	Sym.	Z	$\rho_e^{\text{TPSS-D3}}$	No.	Sym.	Z	$\rho_e^{\text{TPSS-D3}}$
22-0	$P2_1/n$	4	1.77	24-4	$P2_1/c$	4	1.55
22-1	$P2_1/c$	4	1.76	24-5	$P2_1/c$	4	1.54
22-2	$P2_1/n$	4	1.83	24-6	$P2_12_12_1$	4	1.49
22-3	$P2_1/c$	4	1.79	24-7	$P2_1/c$	4	1.49
22-4	$P2_1/n$	4	1.75	24-8	$Pna2_1$	4	1.56
22-5	$P\bar{1}$	2	1.78	24-9	$P2_12_12_1$	4	1.49
22-6	$P\bar{1}$	2	1.74	24-10	$P2_1/c$	4	1.57
22-7	$P2_1$	2	1.70	25-0	$P2_1/c$	4	1.42
22-8	$P2_1/n$	4	1.77	25-1	$Pbca$	8	1.43
22-9	$Pna2_1$	4	1.81	25-2	$P\bar{1}$	2	1.45
22-10	$P2_1/c$	4	1.76	25-3	$P2_1/c$	4	1.41
23-0α	$P2_1/c$	4	1.41	25-4	$P2_1/c$	4	1.39
23-0β	$P\bar{1}$	2	1.46	25-5	$P\bar{1}$	2	1.39
23-0γ	$P\bar{1}$	4	1.46	25-6	$P\bar{1}$	2	1.42
23-0δ	$P2_1/n$	4	1.40	25-7	$P\bar{1}$	2	1.41
23-0ϵ	$P\bar{1}$	4	1.44	25-8	$P\bar{1}$	2	1.43
23-1	$P\bar{1}$	2	1.44	25-9	$P2_1/c$	4	1.38
23-2	$P\bar{1}$	2	1.47	25-10	$P2_1/c$	4	1.41
23-3	$P\bar{1}$	2	1.46	26-0	$P\bar{1}$	2	1.40
23-4	$P\bar{1}$	2	1.47	26-1	$P2_1/n$	4	1.43
23-5	$P\bar{1}$	2	1.46	26-2	$P2_1/n$	4	1.44
23-6	$P\bar{1}$	2	1.49	26-3	$P2_1/n$	4	1.42
23-7	$P\bar{1}$	2	1.42	26-4	$P2_1/n$	4	1.43
23-8	$P\bar{1}$	2	1.46	26-5	$P2_1/c$	4	1.44
23-9	$P2_1/c$	4	1.42	26-6	$P\bar{1}$	2	1.43
23-10	$P2_1/n$	4	1.45	26-7	$C2/c$	8	1.46
24-0	$P2_1/c$	4	1.60	26-8	$P2_1/n$	4	1.37
24-1	$P2_1/c$	4	1.55	26-9	$P\bar{1}$	4	1.46
24-2	$P2_1/n$	4	1.50	26-10	$P\bar{1}$	4	1.41
24-3	$P2_1/c$	4	1.51				

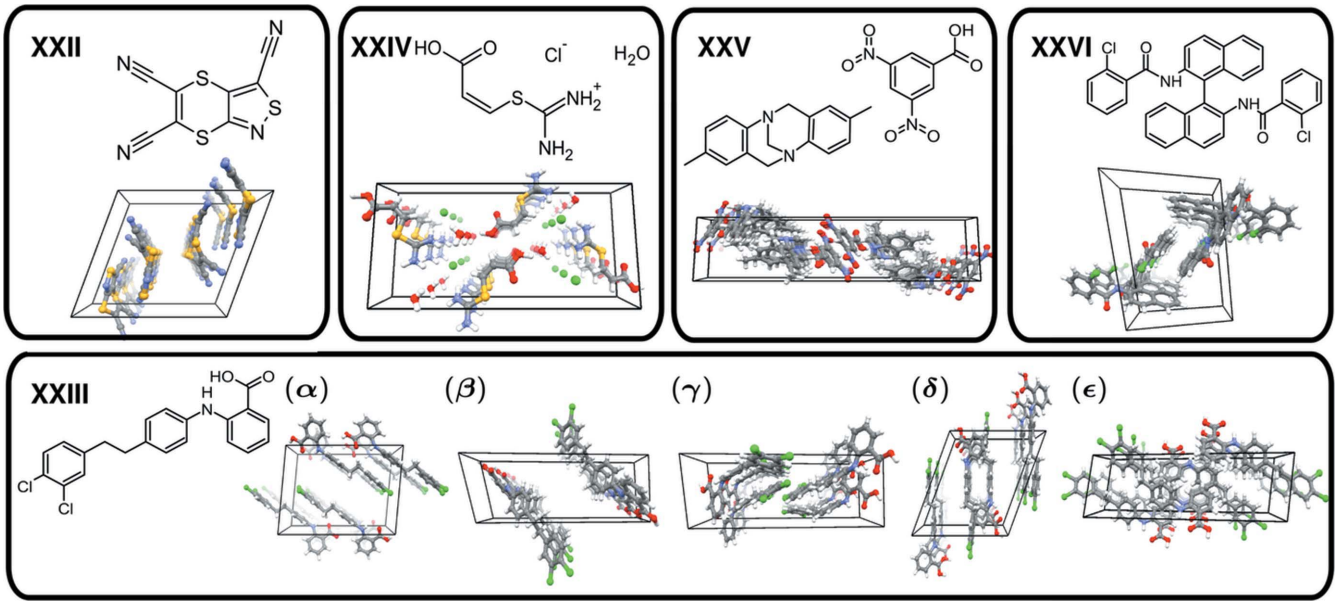


Figure 2
Target systems of the sixth crystal structure prediction blind test depicted as Lewis structures and with a perspective *ac* projection of the experimentally realised polymorphs.

Table 4
Benchmark geometries of the experimentally realised blind test polymorphs isotropically corrected for ZPVE and thermal effects.

	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Vol. (Å^3)
22-0							
Ref.	11.906	6.673	12.555	90	108.6	90	945.50
TPSS-D3	11.909	6.646	12.511	89.9	108.5	90	938.77
PBE-MBD	12.103	6.756	12.865	90	109.7	90	990.57
HF-3c	11.449	6.553	12.312	89.9	105.7	90	889.43
23-0α							
Ref.	10.980	10.356	15.968	90	95.7	90	1806.63
TPSS-D3	10.657	10.436	16.357	90	93.7	90	1815.23
PBE-MBD	10.803	10.489	16.335	90	94.4	90	1845.55
HF-3c	10.462	10.521	16.342	90	92.0	90	1797.67
23-0β							
Ref.	6.901	7.688	18.619	85.3	80.8	65.8	888.49
TPSS-D3	7.214	7.692	17.631	88.5	80.8	64.8	873.05
PBE-MBD	7.221	7.770	17.943	87.6	80.5	64.5	895.57
HF-3c	7.185	7.734	17.317	89.5	80.8	65.4	862.05
23-0γ							
Ref.	7.482	11.793	20.026	84.8	85.4	80.1	1729.58
TPSS-D3	7.430	11.995	19.802	86.4	89.1	83.0	1748.24
PBE-MBD	7.492	12.011	19.977	86.1	88.4	82.9	1779.30
HF-3c	7.380	11.860	19.716	87.4	90.4	83.2	1711.73
23-δ							
Ref.	13.648	10.544	13.837	90	113.6	90	1824.28
TPSS-D3	13.739	10.591	13.460	89.1	111.4	90.7	1823.51
PBE-MBD	13.789	10.632	13.553	89.2	111.8	90.6	1845.08
HF-3c	14.638	10.367	13.049	90	114.2	90	1806.80
23-0ϵ							
Ref.	6.575	11.976	23.982	102.9	96.7	97.2	1805.39
TPSS-D3	6.689	11.539	23.949	99.2	98.2	98.0	1781.38
PBE-MBD	6.705	11.652	23.968	99.0	98.2	97.7	1806.96
HF-3c	6.566	11.341	24.520	100.3	98.0	98.9	1748.47
24-0							
Ref.	3.979	21.058	10.008	90	97.8	90	830.90
TPSS-D3	4.024	20.761	10.001	90	98.4	90	826.65
PBE-MBD	4.093	21.008	10.009	90	98.1	90	852.19
HF-3c	4.058	21.016	10.003	90	98.3	90	844.29
25-0							
Ref.	10.243	27.109	7.983	90	109.6	90	2088.51
TPSS-D3	10.376	27.278	8.104	90	109.8	90	2157.53
PBE-MBD	10.346	27.217	8.050	90.1	110.1	90	2128.64
HF-3c	10.318	27.274	7.773	90	109.5	90	2061.81
26-0							
Ref.	10.247	10.866	13.968	76.8	73.3	63.5	1323.94
TPSS-D3	10.238	10.872	13.888	78.1	73.7	63.8	1325.12
PBE-MBD	10.273	10.960	13.914	78.2	73.8	64.0	1345.81
HF-3c	10.087	10.740	13.650	80.5	75.4	65.6	1300.43

4. Results and discussion

4.1. Equilibrium geometries

Although the accuracy of the TPSS-D3 geometries has been tested for crystals on the two benchmark sets X23 (Moellmann & Grimme, 2014; organic molecules) and ICE10 (Brandenburg *et al.*, 2015; ice polymorphs), we first analyze the performance for the blind test targets. Furthermore, analytical gradients for MBD dispersion correction have been implemented recently (Blood-Forsythe *et al.*, 2016; Bučko *et al.*, 2016) and the blind test solids provide an unbiased set of realistic test cases. In order to compare the measured X-ray structures with the computed ones on the electronic energy surface, one has to estimate the effects of ZPVE and thermal (entropic) effects on the structure. As the intermolecular potentials of organic molecules typically exhibit a slight asymmetry with smaller curvature at longer intermolecular distances, most organic crystals expand under heating. In the first approximation, one can describe the effect isotropically distributed in all spatial directions and analyze the potential energy surface (PES) with respect to the unit-cell volumes. Here we follow the procedure described in Brandenburg *et al.* (2015; Grimme *et al.*, 2015). The electronic PES for all nine systems is computed at the DFTB3-D3 level. Harmonic frequencies are calculated at all relaxed (constrained volume) geometries and the corresponding electronic energy and free energy minima are extracted from a Murnaghan equation of state fit. The ratio $\rho_e^{\rm calc}/\rho_T^{\rm calc}$ is used to correct the measured mass densities ρ_T to a ZPVE exclusive equilibrium mass density at $T = 0$ K ρ_e (*cf.* Table 3). The error in the final benchmark values ρ_e should be about 1% as estimated previously (Brandenburg *et al.*, 2015). In the same way, one can isotropically correct the cell parameters leading to the benchmark geometries summarized in Table 4.

Table 5
Statistical performance of dispersion-corrected DFAs and low-cost alternatives to compute the polymorph geometries.

MD: mean deviation; MAD: mean absolute deviation; SD: standard deviation; MAX: maximum absolute deviation.

	TPSS-D3	PBE-MBD	HF-3c	DFTB3-D3
Cell length (Å)				
MD	−0.04	0.04	−0.10	−0.34
MAD	0.17	0.16	0.31	0.50
SD	0.27	0.21	0.43	0.62
MAX	0.99	0.68	1.30	1.82
Cell angles (°)				
MD	0.3	0.3	0.6	1.5
MAD	1.0	1.0	1.4	2.4
SD	1.6	1.4	2.0	4.3
MAX	3.7	3.9	5.0	19.1
Cell volume (%)				
MD	0.1	2.0	−1.8	−9.3
MAD	1.0	2.0	2.1	9.3
SD	1.5	1.4	2.1	4.0
MAX	3.3	4.8	5.9	18.1

The statistical performances of TPSS-D3, PBE-MBD and the two low-cost alternatives on the geometry benchmark are given in Table 5. The cell parameters of both dispersion-corrected DFAs are very accurate with a MAD below 0.2 Å. HF-3c is slightly worse with a MAD of 0.3 Å and a small systematic shift of 0.1 Å towards cell parameters that are too small, while the tight-binding based scheme yields substantially underestimated cell lengths by 0.3 Å. The trend of too small intermolecular (noncovalent) distances from DFTB3-D3 is known from the above-mentioned X23 and ICE10 benchmark sets and seems to be a general property of the method. Concerning the cell angle, *i.e.* the shape of the unit cell, all methods are reasonably accurate and produce no clear

outlier. Interestingly, the largest errors of the two DFAs for both cell parameters occur for the same systems, specifically 23-0β (parameter c) and 23-0ϵ (parameter α). Typically, the most sensitive geometry property is the unit-cell volume or the related mass density. Here, TPSS-D3 performs excellently with the smallest MAD of 1%, which is within our estimated reference error. While PBE-MBD has slightly too large unit-cell volumes and MAD of 2%, HF-3c underestimates the unit-cell volumes leading to the same MAD.

The unit-cell volumes are directly translated into crystal mass densities. A graphical representation of the methods performance is shown in Fig. 3. Overall, TPSS-D3 seems to provide very good geometries of the molecular crystals within the blind test challenge. Note that a direct comparison with the X-ray data without proper treatment of the ZPVE and thermal effects substantially increases the errors of all shown methods. For instance, the MADs of the unit-cell volumes increases to 4 and 3% for TPSS-D3 and PBE-MBD, respectively. The structure match conducted in the blind-test evaluation is of this type, which is reasonable as in principle free-energy structures should be submitted. The averaged RMSD (crystal packing similarity tool, see §3, full table in the supporting information) on all nine structures is 0.27, 0.24, 0.41 and 0.59 Å for TPSS-D3, PBE-MBD, HF-3c and DFTB3-D3, respectively. The values for the dispersion-corrected DFAs are very reasonable, considering that no thermal expansion had been taken into account here. As PBE-MBD systematically overestimates the unit-cell volumes while HF-3c underestimates it, the PBE-MBD errors are slightly smaller.

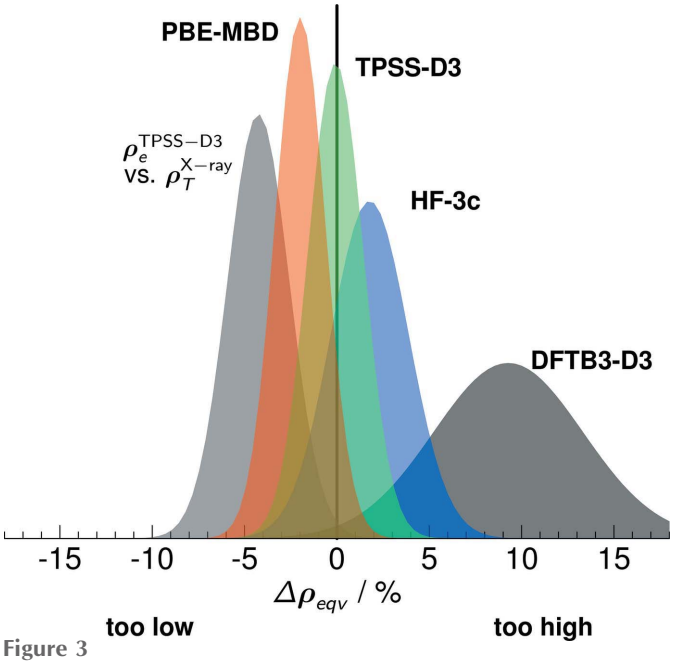


Figure 3
Deviations of crystal mass densities computed by several electronic structure methods compared with back-corrected experimental references. The statistical measures are represented as normal error distributions although their actual distribution might be slightly asymmetric. We highlight the systematic error, when the TPSS-D3 equilibrium (at T = 0 K, ZPVE exclusive) densities $\rho_e^{\mathrm{TPSS-D3}}$ are directly compared with the uncorrected finite temperature X-ray densities $\rho_T^{\mathrm{X-ray}}$.

Interestingly, the deviations for system 23 are largest. For this system, the estimated thermal expansion is most pronounced (*cf.* Table 2). As known from previous studies (Brandenburg *et al.*, 2015), the tight-binding method has substantially too close intermolecular distances leading to RMSDs of $\sim$0.5 Å compared with the original X-ray structure.

4.2. Energy ranking

The TPSS-D3 geometries are of high quality, as shown in the previous section. Therefore, we use the TPSS-D3 structures for the POLY59 set (as computed in the sixth crystal structure prediction blind test) as fixed benchmark geometries to test other electronic structure methods.

We use the methods summarized in §2, *i.e.* focusing on dispersion-corrected DFAs. For each experimental polymorph, we compare its lattice energy with those of the competing structures and sort them energetically. The complete lists are given in the supporting information. We give the corresponding rank of the experimental polymorphs and the relative energy ΔE with respect to the theoretically most stable structure. Ideally, the rank is one and therefore ΔE zero. For the target (XXIII) with five polymorphs, the analysis is less clear as their stability depends on the experimental condition. However, we assume that these five should be the lowest-energy polymorphs without defining their order. In this way we apply the same ranking system and the overall best summed and mean rank over all nine polymophs is 19 and 2.1, respectively. The worst possible ranking within our restricted POLY59 set is 109 and 12 for the summed and mean ranks. All individual values are given in the supporting information and a statistical summary is given in Table 6.

In Fig. 4 we show the relative lattice energies of the POLY59 set computed with TPSS-D3. For the five targets 22, 23b, 24, 25 and 26, the experimental polymorph is indeed computed to be the energetically most stable one. Although there is a close-lying crystal structure 22-1, target 22 seems to be rather easy for the dispersion-corrected DFAs as all find

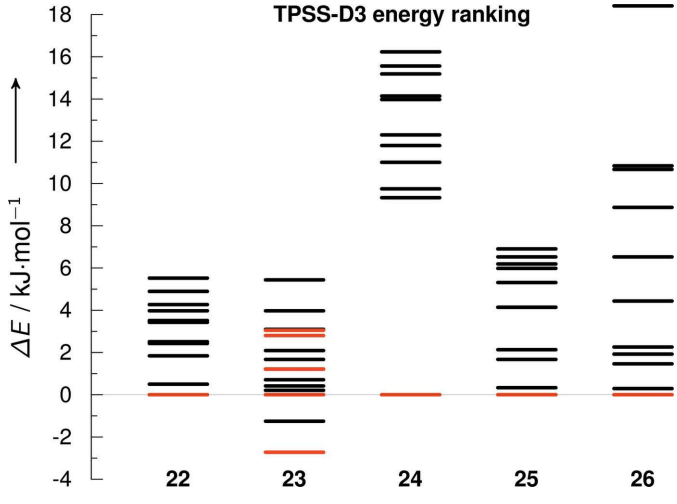


Figure 4
Relative crystal energies of the POLY59 set as computed with TPSS-D3. The experimental polymorphs are shown as red bars.

the experimental structure as the most stable one. The results of the less successful HF-3c and DFTB3-D3 methods indicate that target 22 requires an accurate treatment of induction and exchange repulsion. Target 23 is the most problematic system of the whole blind test. While the five experimental polymorphs are within an energy window of ~ 6 kJ mol^{-1}, there are several competing structures. Further, there is a large spread of the results of all applied DFAs. By just considering the Lewis structure of target 24 with partially conjugated functional groups and significant partial charges, it seems to be a substantial challenge. While this is certainly true for the search algorithms, the final energy evaluation is comparably easy. This can be seen from the significant energy gap (~ 8 kJ mol^{-1}) between the experimental polymorph 24 and the next stable structure. As a result, all methods that account for dispersion interactions in some way are successful. This may also explain why the tailor-made force field of Neumann and co-workers (in their blind test submission; Reilly *et al.*, 2016) is able to describe this system without considering explicitly induction effects. The co-crystal (25) has one close lying competing structure that is estimated as the most stable structure by PBE-D2 and PBE-MBD. According to the successful methods TPSS-D3, PBE-TS and PBE-D3 its energy difference is only 0.5–1.5 kJ mol^{-1}. Together this indicates that the orthorhombic structure 25-1 may be a competing polymorph to the experimentally found 25-0. A similar situation can be seen for target 26. Here, two alternative structures (monoclinic 26-1 and 26-2) compete energetically with the experimentally found triclinic structure (26-0). Again, several methods compute them to be energetically more stable (M06L, vdW-DF2, PBE-TS and PBE-dDsC), while others predict them to be less stable compared with the experimental polymorph by less than 4 kJ mol^{-1} (TPSS-D3, PBE-D3 and PBE-MBD).

Overall, TPSS-D3 and PBE-D3 have the highest success rate of all the electronic structure methods tested. PBE-MBD also performs well and seems to be a substantial improvement over PBE-TS although the number of hits is only four for both. However, with PBE-MBD the mean rank of the experimental structures and their mean energy difference to the most stable polymorph is with 11.3 and -0.9 kJ mol^{-1}, respectively, very good. This is similar to the PBE-D3 rankings of 10.3 and -1.2 kJ mol^{-1}. While the C_6 only dispersion-corrected methods (D2, TS, dDsC) perform reasonably, they are clearly outperformed in all the analyzed statistics by the D3- and MBD-corrected DFAs. The accuracy of M06L and vdW-DF2 seems to be closer to the simple pairwise dispersion corrections. The results of the plain PBE functional have to be judged carefully. While the mean computed rank is similar to vdW-DF2, this is only possible by using the good TPSS-D3 geometries. The uncorrected PBE lattice energies are positive for several crystals, which would result in unbound crystals. Interestingly, all pairwise dispersion corrections seem to overestimate the dispersion interaction as indicated by an over-stabilization of closed-packed structures. Only upon inclusion of higher-order terms (C_8, C_9 *etc.*) a clear improvement over plain PBE is apparent.

Table 6
Performance of electronic structure methods for the POLY59 set.

We give the mean rank of the experimental polymorph, the mean energy difference to the most stable structure (kJ mol^{-1}), the mean energy interval of all structures (kJ mol^{-1}), and the number of hits, *i.e.* how often the experimental polymorph is computed as most stable structure. For each quality measure we highlight the two best performing methods.

	Rank	ΔE	$\Delta E^{\rm min}_{\rm max}$	# hits	X23†
TPSS-D3	3.4	−2.0	11.4	5	4.6
TPSS-D3//HF-3c	4.9	−2.2	9.7	2	–
HF-3c	7.2	−9.3	20.4	1	9.2‡
DFTB3-D3	8.2	−14.7	29.2	1	10.0§
M06L	5.7	−3.9	14.5	3	9.6
vdW-DF2	4.1	−2.7	18.4	2	6.3
PBE	4.0	−4.6	19.5	2	48.5
PBE-D2	5.4	−2.9	12.0	3	7.5
PBE-TS	4.7	−3.7	14.9	4	10.0
PBE-dDsC	4.2	−14.6	19.3	3	–
PBE-D3	2.8	−1.2	10.3	5	5.0
PBE-MBD	2.7	−0.9	11.3	4	6.3

† Mean absolute deviation on lattice energy benchmark set (X23) in kJ mol^{-1} as summarized in Grimme *et al.* (2016). ‡ From Brandenburg *et al.* (2014). § From Brandenburg & Grimme (2014*a*).

Note that although optimized structures at the DFTB3-D3 and HF-3c levels exist for the full POLY59 set (*cf.* with the previous section), we evaluate them on the TPSS-D3 geometries for consistent comparison with the other DFAs. The low-cost alternatives DFTB3-D3 and HF-3c perform as expected and the mean ranks increase to 7.2 and 8.2, respectively, and mean relative energies to -9.3 and -14.7 kJ mol^{-1}, respectively. While this is worse than most tested DFAs, one has to consider the substantial computational speed-up of up to four orders of magnitudes. A rather cost-efficient improvement of the HF-3c method is a TPSS-D3 single-point energy calculation on the HF-3c structure. The accuracy of the HF-3c structures have been shown in the previous section and this seems to translate into more or less parallel potential energy surfaces with respect to TPSS-D3. The mostly parallel potential energy surfaces of PBE-TS and PBE-MBD are probably also the reason why the approach of Tkatchenko and co-workers to compute PBE-MBD energies on PBE-TS structures can be successfully applied (Reilly *et al.*, 2016).

As a comparison and cross check, Table 6 additionally gives the MADs on the lattice energy benchmark set X23 as summarized in Grimme *et al.* (2016). One can see a clear correspondence between a good performance on the X23 set and a successful energy ranking of the POLY59 structures. For instance, the Pearson correlation coefficient between the mean energy difference to the most stable structure ($\overline{\Delta E}$) and the X23 MAD for all dispersion accounting methods is 0.67 and for the PBE based methods it even increases to 0.91. Apparently, one can generally improve the relative crystal stabilities by computing better absolute lattice energies. While this is reasonable, it is still an important finding as the reference values of both sets have certain limitations.

In order to analyze if certain similar structures are more consistently described by the various applied methods, we tried to correlate the deviation of different methods (*e.g.* the

TPSS-D3 stabilities with the PBE-MBD ones) with the mass density. However, for none of the density functionals is a clear correlation apparent. Only for DFTB3-D3, a mild correlation (overestimation of densely packed structures) is found, which can be explained by its general problems with the Pauli repulsion regime (Gaus *et al.*, 2011). Additionally, we compared the mean absolute difference of all structures between the TPSS-D3 and PBE-MBD stabilities with a subset containing only the same space group as the experimental structure. Although we expect this subset to exhibit similar packing motifs, the energy difference is only slightly reduced by 5% on average.

In Fig. 5 we compare the energy rankings of several methods for the co-crystal 25. We show the D3-based hierarchical methods, specifically DFTB3-D3, HF-3c, TPSS-D3//HF-3c (computed on HF-3c geometries) and TPSS-D3. As indicated by the statistical performance, the ranking improves in the order DFTB3-D3 (6), HF-3c (5), TPSS-D3//HF-3c (3) and TPSS-D3 (1). A general trend is that the less accurate methods (with worse mean ranking) also have a larger energy spread, the Pearson correlation coefficient is 0.7. This can also be seen for target 25, *i.e.* the energy spread decreases in the above given order (*cf.* Fig. 5). Additionally, we compare the TPSS-D3 ranking with PBE-D3 and PBE-MBD. The rankings are rather similar, but even here we see some re-ordering, denoting that more accurate electronic structure methods are needed for energetically very close lying polymorphs. This example also highlights that the effect of the density functional (TPSS *versus* PBE) is more pronounced compared with the effect of the dispersion correction (D3 *versus* MBD). Therefore, the next reasonable step to gain higher accuracies is the application of a hybrid DFA, for instance PBE0-D3 energies on the TPSS-D3 structures.

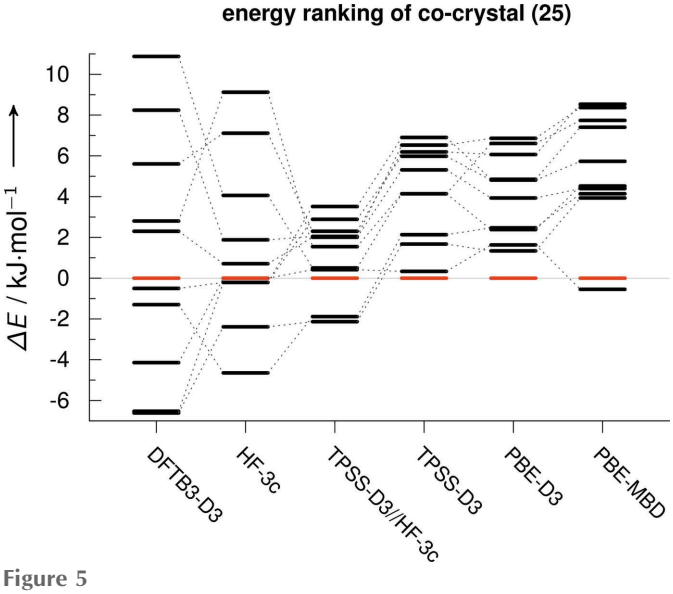


Figure 5
Relative crystal energies of the co-crystal (target 25) computed with the D3-based hierarchical methods and PBE-MBD as comparison. The experimental polymorphs are shown as red bars and the reordering is denoted by dotted lines.

Table 7
Dependence of the crystal stabilities on the TPSS-D3 geometry.

The energy difference ΔE between the full relaxation (equilibrium energy surface) and the scaled unit cell (free energy surface) is given per mole.

No.	Scaling (%)	ΔE (kJ mol^{-1})
23-0α	5.11	0.97
23-0β	4.64	0.64
23-0γ	6.38	2.19
23-0δ	5.32	0.70
23-0ϵ	5.19	0.48

4.3. Stability analysis

In order to help and guide future applications of the provided test set, we want to analyze the importance of the neglected effects here. The first is our basic assumption that the TPSS-D3 equilibrium geometries are accurate and consistent enough for the other electronic structure methods. We showed in §4.1 that the predicted mass densities (and other lattice parameters) agree excellently with the X-ray measurements when the ZPVE and thermal effects are back-corrected from the measured properties. Still, the use of the equilibrium geometries throughout the energy-ranking analysis might introduce a mild bias. Therefore, we recomputed the stability of all measured polymorph structures of 23 with appropriately scaled unit-cell volumes (*cf.* §4.1) and subsequent relaxation of the atomic positions. The modified stabilities are summarized in Table 7.

All energies rise slightly as they should. Although some unit cells are scaled significantly by 5%, the resulting energy differences are comparably small. This is due to the apparently flat potential energy landscape. The largest energy difference is ~ 1 kJ mol^{-1}. While for the relative energies this is still 0.5 kJ mol^{-1}, which can in principle be important for the stabilities.

The free-energy contributions to the relative stability are the second neglected effect. In our second blind test submission (Reilly *et al.*, 2016) we computed ZPVE and thermal contributions *via* harmonic HF-3c frequencies, which did not change the rank of the energetically lowest predicted structures (within the TPSS-D3 list) and the effect on relative energies was always below 2 kJ mol^{-1}. We therefore suggest computing this contribution only for the very close lying polymorph candidates. For energy landscapes with a significant gap between the structures (*e.g.* target 24), a pure electronic energy ranking will presumably be sufficient. In order to demonstrate typical relative contributions of the different thermodynamic contributions, we summarized them in Table 8 for the five polymorphs of target 23. For the entropic contribution, the low lying modes are crucial. Previously, we computed Γ-point frequencies at the HF-3c level for the primitive unit cells. We estimate the effect of low-lying modes at the DFTB3-D3 level by generating supercells with a minimum unit-cell parameter of 12.5 Å, *i.e.* the given thermodynamic contributions are

$$\Phi^{\mathrm{HF-3c}}_{\mathrm{supercell}} = \Phi^{\mathrm{HF-3c}}_{\mathrm{unit\,cell}} + \left(\Phi^{\mathrm{DFTB3-D3}}_{\mathrm{supercell}} - \Phi^{\mathrm{DFTB3-D3}}_{\mathrm{unit\,cell}}\right).$$

Table 8
Thermodynamic functions for the five polymorphs of target 23 under normal conditions.

We give the lattice energy E_{latt} at the TPSS-D3 level computed on scaled TPSS-D3 geometries (see above). Additionally, zero-point and thermal vibrational energies ΔE_{ZPVE} and ΔE_t, volume work ΔPV, entropy ΔTS, enthalpy ΔH and free enthalpy, a.k.a. Gibbs free energy ΔG, are computed with harmonic HF-3c frequencies including corrections for low lying modes estimated at the DFTB3-D3 level. All energies are given per molecule, relative to the gas phase, and in kJ mol^{-1}.

No.	E_{latt}	ΔE_{ZPVE}	ΔE_t	ΔPV	ΔTS	ΔH	ΔG
23-0α	−190.1	7.5	6.3	−2.5	−65.5	−178.7	−113.2
23-0β	−193.2	8.7	5.0	−2.5	−71.9	−181.9	−110.0
23-0γ	−188.0	8.9	5.7	−2.5	−70.6	−175.4	−105.3
23-0δ	−189.0	8.2	5.8	−2.5	−69.0	−177.4	−108.4
23-0ϵ	−189.4	8.7	5.6	−2.5	−69.7	−177.6	−107.9

The lattice energy is given as the above estimated TPSS-D3 energy on the free energy surface.

The volume work PV is identical for all structures, which is due to the similar volume per molecule of the different polymorphs, and the normal pressure condition. Larger differences can be archived by increasing the external pressure. The zero-point vibrational energy and the thermal vibrational energy have a range of ~ 1.5 kJ mol^{-1}. The additive quantity $\Delta E_{ZPVE} + \Delta E_t + \Delta PV$ has an energy spread below 1 kJ mol^{-1}. This results in an unmodified enthalpy ranking of the five polymorphs compared with the plain electronic lattice energy. The entropy term deviates more pronounced, yielding an energy difference ΔTS of up to 6 kJ mol^{-1}. Interestingly, this leads to a destabilization of polymorph 23β in accordance with the measurements, where form α and δ should be the most stable structures around room temperature. This is in reasonable agreement with our free energy estimates. Still, it is not 100% consistent as α should be most stable at 257 K and δ at 293 K. Within our estimates, the temperature has to be significantly higher to destabilize form β relative to form δ. Further, the relative rank of α and δ is reversed, which could only agree with the experimental measurements when the entropy contributions are modified.

Thus, some inconsistencies remain, but the most pronounced effect of destabilizing form β seems to be correct. Especially when investigating polymorphism at different external conditions, explicit temperature and pressure investigations seems necessary.

5. Conclusions

We used the crystal energy landscapes of the sixth crystal structure prediction blind test to compile a benchmark set POLY59. As we are testing relative crystal stabilities, the set reflects the ability of a method to compute polymorph stabilities. The set consists of nine experimental structures (five polymorphs) and is complemented with the ten energetically lowest structures from our TPSS-D3 blind test submission. First, we back-corrected the measured X-ray geometries for thermal and zero-point vibrational effects. This enables a direct comparison with optimizations on the electronic energy surface. We highlighted the excellent geometries of PBE-MBD and TPSS-D3, where the latter one yields more accurate mass densities with an MAD of only 1%. Second, the TPSS-D3 geometries are used to benchmark several dispersion-corrected DFAs concerning their energetic polymorph ranking. We use the number of hits (the lowest energy structure is the experimental one) as well as the mean computed rank and corresponding energy difference to the most stable structure as quality measures. Again, TPSS-D3 is one of the most successful methods together with PBE-D3 and PBE-MBD providing 5, 5 and 4 hits, respectively. The mean energy difference of the experimental structure with respect to the most stable computed structure is below 2 kJ mol^{-1} for all three methods, *i.e.* we gain about a factor of two in accuracy for the computation of relative lattice energies compared with absolute sublimation enthalpies. We showed that the D3- and MBD-corrected DFAs outperform other dispersion corrections (D2, TS, dDsC) as well as the Minnesota functional M06L and the van der Waals density functional vdW-DF2. Some low-cost methods are shown to provide reasonable results useful for screening applications.

For higher accuracies (*e.g.* a correct energy ranking within 1 kJ mol^{-1}), one has to consider several additional steps:

(*a*) Improve the electronic energy contribution by using a hybrid density functional, *e.g.* PBE0-D3 or HSE06-D3 on the TPSS-D3 geometries.

(*b*) Compute the free energy corrections to the electronic energy.

(*c*) Use a (isotropic) free-energy correction on the geometries to compute (*a*) and (*b*) on the correct free energy surface. Concerning point (*a*) high-level quantum chemical methods could in principle be applied as demonstrated by Chan coworkers for benzene (Yang *et al.*, 2014; Beran, 2015). However, the computational costs are substantial and rarely applicable in routine crystal structure predictions. More cost-efficient strategies such as a related fragmented MP2c approach by Beran could be useful to confirm PBE0-D3 results (Beran *et al.*, 2014). The ZPVE and thermal (entropic) contributions to the free energy could possibly be computed at a lower computational level and may have an impact on the polymorph ranking (Nyman & Day, 2015; Heit *et al.*, 2016). In our TPSS-D3 blind test rankings we computed harmonic HF-3c frequencies, which did not change the rank of the experimental polymorph and the effect on relative energies was always below 2 kJ mol^{-1}. Therefore, we think that this has to be combined with improvements (*a*) and (*c*) for conclusive results. We showed how the isotropic geometry corrections to free energies can be performed at the tight-binding level, and this can be done for all predicted structures with subsequent constrained optimizations at the DFA level.

Overall we can conclude that a robust routine scheme to high-accuracy crystal energy landscapes has been developed at a full quantum chemical, first-principles level. A prerequisite is that efficient and accurate prefiltering at a lower theory level has been conducted so that a limited number of structures (in the thousands) are finally evaluated. Hence, the two main future challenges are the initial sampling of the huge poly-

morph space and an improved relative free-energy calculation of the lowest polymorph structures within 1 kJ mol^{-1}.

APPENDIX *A*
Technical details of the calculations

The periodic structure relaxations were conducted with a developer version of the *CRYSTAL*14 program (Dovesi *et al.*, 2014). It is the ideal choice for cost-effective HF-3c calculations in small basis sets as it can exploit full point and space group symmetry (Orlando *et al.*, 2014). The tight-binding DFTB3 is used in its third-order variant with self-consistent charge redistribution as implemented in `dftb+` (Seifert & Joswig, 2012; Aradi *et al.*, 2007; Elstner, 2007). We use the 3OB Slater–Koster files generated by Elstner and coworkers and damp all hydrogen-containing pair potentials with an exponent of 4.2 (Gaus *et al.*, 2013). The DFA energy and gradient contributions are computed with the *VASP*5.4 program package (Kresse & Furthmüller, 1996*a*,*b*). We use the projector augmented plane wave (PAW) basis sets with upper energy cutoff of 800 eV (Blöchl, 1994; Kresse & Joubert, 1999). The Brillouin zone is sampled with a Γ-centered *k*-mesh with grid density of at least 0.025 Å^{-1}. Standard integral thresholds and large DFA integration grids were used. The calculation of the D2 (Grimme, 2006) and D3 (Grimme *et al.*, 2010, 2011) dispersion correction is carried out within the *CRYSTAL*14 code, which includes efficient analytical derivatives of the three-body dispersion. They are accessible in our freely available `dftd3` from our website (http://www.thch.uni-bonn.de/tc/dftd3). The TS (Tkatchenko & Scheffler, 2009), MBD (Tkatchenko *et al.*, 2012), dDsC (Steinmann *et al.*, 2009), M06L (Zhao & Truhlar, 2006) and vdW-DF2 (Lee *et al.*, 2010) energy contributions are calculated within *VASP* with default thresholds. The efficient reciprocal space implementation of the MBD gradient was recently introduced by Bučko and coworkers (Bučko *et al.*, 2016) and we used it with the identical Brillouin zone sampling as given above. Note that due to slightly tighter numerical thresholds compared with the blind test submission, small energy differences of $\sim$ 0.4 kJ mol^{-1} can occur.

Acknowledgements

We thank Sally Price for providing optimized crystal structures (1000 per blind test target) for further refinement, helpful discussions concerning the interpretation of crystal energy landscapes and proof-reading of the manuscript. JGB acknowledges support by the Alexander von Humboldt foundation within the Feodor–Lynen program.

References

Aradi, B., Hourahine, B. & Frauenheim, T. (2007). *J. Phys. Chem. A*, **111**, 5678–5684.
Axilrod, B. M. & Teller, E. (1943). *J. Chem. Phys.* **11**, 299–300.
Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., Kazantsev, A. V., Kendrick, J., Kuleshova, L. N., Leusen, F. J. J., Maleev, A. V., Misquitta, A. J., Mohamed, S., Needs, R. J., Neumann, M. A., Nikylov, D., Orendt, A. M., Pal, R., Pantelides, C. C., Pickard, C. J., Price, L. S., Price, S. L., Scheraga, H. A., van de Streek, J., Thakur, T. S., Tiwari, S., Venuti, E. & Zhitkov, I. K. (2011). *Acta Cryst.* B**67**, 535–551.
Becke, A. D. & Johnson, E. R. (2005*a*). *J. Chem. Phys.* **123**, 154101.
Becke, A. D. & Johnson, E. R. (2005*b*). *J. Chem. Phys.* **122**, 154104.
Becke, A. D. & Johnson, E. R. (2006). *J. Chem. Phys.* **124**, 014104.
Becke, A. D. & Johnson, E. R. (2007). *J. Chem. Phys.* **127**, 124108.
Beran, G. J. O. (2015). *Angew. Chem. Int. Ed.* **54**, 396–398.
Beran, G. J. O. (2016). *Chem. Rev.* **116**, 5567–5613.
Beran, G. J. O., Wen, S., Nanda, K., Huang, Y. & Heit, Y. (2014). *Top. Curr. Chem.* **345**, 59–93.
Blöchl, P. E. (1994). *Phys. Rev. B*, **50**, 17953–17979.
Blood-Forsythe, M. A., Markovich, T., DiStasio, R. A., Car, R. & Aspuru-Guzik, A. (2016). *Chem. Sci.* **7**, 1712–1728.
Brandenburg, J. G. & Grimme, S. (2014*a*). *J. Phys. Chem. Lett.* **5**, 1785–1789.
Brandenburg, J. G. & Grimme, S. (2014*b*). *Top. Curr. Chem.* **345**, 1–23.
Brandenburg, J. G., Hochheim, M., Bredow, T. & Grimme, S. (2014). *J. Phys. Chem. Lett.* **5**, 4275–4284.
Brandenburg, J. G., Maas, T. & Grimme, S. (2015). *J. Chem. Phys.* **142**, 124104.
Bryantsev, V. S., Diallo, M. S., van Duin, A. C. T. & Goddard, W. A. (2009). *J. Chem. Theory Comput.* **5**, 1016–1026.
Bučko, T., Lebègue, S., Gould, T. & Ángyán, J. G. (2016). *J. Phys. Condens. Matter*, **28**, 045201.
Burns, L. A., Vázquez-Mayagoitia, A., Sumpter, B. G. & Sherrill, C. D. (2011). *J. Chem. Phys.* **134**, 084107.
Buřvar, D.-K., Lancaster, R. W. & Bernstein, J. (2015). *Angew. Chem. Int. Ed.* **54**, 6972–6993.
Carter, D. J. & Rohl, A. L. (2014). *J. Chem. Theory Comput.* **10**, 3423–3437.
Casimir, H. B. G. & Polder, D. (1948). *Phys. Rev.* **73**, 360–372.
Day, G. M., Cooper, T. G., Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., Della Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., Leusen, F. J. J., Kendrick, J., Price, S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G. W. A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., van de Streek, J., Wolf, A. K. & Schweizer, B. (2009). *Acta Cryst.* B**65**, 107–125.
Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Leusen, F. J. J., Liang, C., Pantelides, C. C., Karamertzanis, P. G., Price, S. L., Lewis, T. C., Nowell, H., Torrisi, A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U. & Verwer, P. (2005). *Acta Cryst.* B**61**, 511–527.
DiStasio, R. A., Gobre, V. V. & Tkatchenko, A. (2014). *J. Phys. Condens. Matter*, **26**, 213202.
Dobson, J. F. (2014). *Int. J. Quantum Chem.* **114**, 1157–1161.
Dovesi, R., Orlando, R., Erba, A., Zicovich-Wilson, C. M., Civalleri, B., Casassa, S., Maschio, L., Ferrabone, M., De La Pierre, M., D'Arco, P., Noël, Y., Causà, M., Rérat, M. & Kirtman, B. (2014). *Int. J. Quantum Chem.* **114**, 1287–1317.
Elstner, M. (2007). *J. Phys. Chem. A*, **111**, 5614–5621.
Gaus, M., Cui, Q. & Elstner, M. (2011). *J. Chem. Theory Comput.* **7**, 931–948.
Gaus, M., Goez, A. & Elstner, M. (2013). *J. Chem. Theory Comput.* **9**, 338–354.
Goerigk, L. & Grimme, S. (2011). *Phys. Chem. Chem. Phys.* **13**, 6670–6688.
Grimme, S. (2006). *J. Comput. Chem.* **27**, 1787–1799.

Grimme, S. (2011). *WIREs Comput. Mol. Sci.* **1**, 211–228.
Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. (2010). *J. Chem. Phys.* **132**, 154104.
Grimme, S., Brandenburg, G., Bannwarth, C. & Hansen, A. (2015). *J. Chem. Phys.* **143**, 054107.
Grimme, S., Ehrlich, S. & Goerigk, L. (2011). *J. Comput. Chem.* **32**, 1456–1465.
Grimme, S., Hansen, A., Brandenburg, J. G. & Bannwarth, C. (2016). *Chem. Rev.* **116**, 5105–5154.
Heit, Y. N., Nanda, K. D. & Beran, G. J. O. (2016). *Chem. Sci.* **7**, 246–255.
Karamertzanis, P. G., Day, G. M., Welch, G. W. A., Kendrick, J., Leusen, F. J. J., Neumann, M. A. & Price, S. L. (2008). *J. Chem. Phys.* **128**, 244708.
Karamertzanis, P. G. & Pantelides, C. C. (2007). *Mol. Phys.* **105**, 273–291.
Kazantsev, A. V., Karamertzanis, P. G., Adjiman, C. S. & Pantelides, C. C. (2010). *CrystalOptimizer. An Efficient Algorithm for Lattice Energy Minimisation of Organic Crystal using Isolated-Molecule Quantum Mechanical Calculations*, Vol. 6. Weinheim: Wiley.
Kazantsev, A. V., Karamertzanis, P. G., Adjiman, C. S., Pantelides, C. C., Price, S. L., Galek, P. T., Day, G. M. & Cruz-Cabeza, A. J. (2011). *Int. J. Pharm.* **418**, 168–178.
Kennedy, M. R., McDonald, A. R., DePrince, A. E., Marshall, M. S., Podeszwa, R. & Sherrill, C. D. (2014). *J. Chem. Phys.* **140**, 121104.
Klimeš, J. & Michaelides, A. (2012). *J. Chem. Phys.* **137**, 120901.
Kresse, G. & Furthmüller, J. (1996*a*). *Comput. Mater. Sci.* **6**, 15–50.
Kresse, G. & Furthmüller, J. (1996*b*). *Phys. Rev. B*, **54**, 11169–11186.
Kresse, G. & Joubert, J. (1999). *Phys. Rev. B*, **59**, 1758–1775.
Kruse, H. & Grimme, S. (2012). *J. Chem. Phys.* **136**, 154101.
Lee, K., Murray, E. D., Éamonn, D., Kong, L., Lundqvist, B. I. & Langreth, D. C. (2010). *Phys. Rev. B*, **82**, 081101.
Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* B**56**, 697–714.
Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Marom, N., DiStasio, R. A., Atalla, V., Levchenko, S., Reilly, A. M., Chelikowsky, J. R., Leiserowitz, L. & Tkatchenko, A. (2013). *Angew. Chem. Int. Ed.* **52**, 6629–6632.
Moellmann, J. & Grimme, S. (2014). *J. Phys. Chem. C*, **118**, 7615–7621.
Mooij, W. T. M., van Eijck, B. P. & Kroon, J. (2000). *J. Am. Chem. Soc.* **122**, 3500–3505.
Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2002). *Acta Cryst.* B**58**, 647–661.
Muto, Y. (1943). *Proc. Phys. Soc. Jpn*, **17**, 629.
Nyman, J. & Day, G. M. (2015). *CrystEngComm*, **17**, 5154–5165.
Orlando, R., De La Pierre, M., Zicovich-Wilson, C. M., Erba, A. & Dovesi, R. (2014). *J. Chem. Phys.* **141**, 104108.
Otero-de-la-Roza, A. & Johnson, E. R. (2012). *J. Chem. Phys.* **137**, 054103.
Otero-de-la-Roza, A. & Johnson, E. R. (2015). *J. Chem. Theory Comput.* **11**, 4033–4040.
Parr, R. G. & Yang, W. (1989). *Density-Functional Theory of Atoms and Molecules*. Oxford University Press.
Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865–3868.
Peverati, R. & Truhlar, D. G. (2014). *Philos. Trans. R. Soc. A*, **372**, 20120476.
Price, S. L. (2014). *Chem. Soc. Rev.* **43**, 2098–2111.
Price, S. L., Leslie, M., Welch, G. W. A., Habgood, M., Price, L. S., Karamertzanis, P. G. & Day, G. M. (2010). *Phys. Chem. Chem. Phys.* **12**, 8478–8490.
Price, S. L. (2009). *Acc. Chem. Res.* **42**, 117–126.
Reilly, A. M., Cooper, R. I., Adjiman, C. S., Bhattacharya, S., Boese, A. D., Brandenburg, J. G., Bygrave, P. J., Bylsma, R., Campbell, J. E., Car, R., Case, D. H., Chadha, R., Cole, J. C., Cosburn, K., Cuppen, H. M., Curtis, F., Day, G. M., DiStasio Jr, R. A., Dzyabchenko, A., van Eijck, B. P., Elking, D. M., van den Ende, J. A., Facelli, J. C., Ferraro, M. B., Fusti-Molnar, L., Gatsiou, C.-A., Gee, T. S., de Gelder, R., Ghiringhelli, L. M., Goto, H., Grimme, S., Guo, R., Hofmann, D. W. M., Hoja, J., Hylton, R. K., Iuzzolino, L., Jankiewicz, W., de Jong, D. T., Kendrick, J., de Klerk, N. J. J., Ko, H.-Y., Kuleshova, L. N., Li, X., Lohani, S., Leusen, F. J. J., Lund, A. M., Lv, J., Ma, Y., Marom, N., Masunov, A. E., McCabe, P., McMahon, D. P., Meekes, H., Metz, M. P., Misquitta, A. J., Mohamed, S., Monserrat, B., Needs, R. J., Neumann, M. A., Nyman, J., Obata, S., Oberhofer, H., Oganov, A. R., Orendt, A. M., Pagola, G. I., Pantelides, C. C., Pickard, C. J., Podeszwa, R. l., Price, L. S., Price, S. L., Pulido, A., Read, M. G., Reuter, K., Schneider, E., Schober, C., Shields, G. P., Singh, P., Sugden, I. J., Szalewicz, K., Taylor, C. R., Tkatchenko, A., Tuckerman, M. E., Vacarro, F., Vasileiadis, M., Vázquez-Mayagoitia, Á., Vogt, L., Wang, Y., Watson, R. E., de Wijs, G. A., Yang, J., Zhu, Q. & Groom, C. R. (2016). *Acta Cryst.* B**72**, 439–459.
Reilly, A. M. & Tkatchenko, A. (2013). *J. Chem. Phys.* **139**, 024705.
Reilly, A. M. & Tkatchenko, A. (2014). *Phys. Rev. Lett.* **113**, 055701.
Řezáč, J., Riley, K. E. & Hobza, P. (2011). *J. Chem. Theory Comput.* **7**, 2427–2438.
Seifert, G. & Joswig, J.-O. (2012). *WIREs Comput. Mol. Sci.* **2**, 456–465.
Spek, A. L. (2009). *Acta Cryst.* D**65**, 148–155.
Steinmann, S. N., Csonka, G. & Corminboeuf, C. (2009). *J. Chem. Theory Comput.* **5**, 2950–2958.
Stone, A. J. (1997). *The Theory of Intermolecular Forces.* Oxford University Press.
Sure, R. & Grimme, S. (2013). *J. Comput. Chem.* **34**, 1672–1685.
Sure, R. & Grimme, S. (2015). *J. Chem. Theory Comput.* **11**, 3785–3801.
Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. (2003). *Phys. Rev. Lett.* **91**, 146401.
Thanthiriwatte, K. S., Hohenstein, E. G., Burns, L. A. & Sherrill, C. D. (2011). *J. Chem. Theory Comput.* **7**, 88–96.
Tkatchenko, A., DiStasio, R. A., Car, R. & Scheffler, M. (2012). *Phys. Rev. Lett.* **108**, 236402.
Tkatchenko, A. & Scheffler, M. (2009). *Phys. Rev. Lett.* **102**, 073005.
Wen, S. & Beran, G. J. O. (2012). *Cryst. Growth Des.* **12**, 2169–2172.
Woods, L. M., Dalvit, D. A. R., Tkatchenko, A., Rodriguez-Lopez, P., Rodriguez, A. W. & Podgornik, R. (2016). *Rev. Mod. Phys.* Submitted. arXiv: 1509.03338.
Yang, J., Hu, W., Usvyat, D., Matthews, D., Schütz, M. & Chan, G. K.-L. (2014). *Science*, **345**, 640–643.
Zhao, Y. & Truhlar, D. G. (2006). *J. Chem. Phys.* **125**, 194101.