

# Assessment of the exchange-hole dipole moment dispersion correction for the energy ranking stage of the seventh crystal structure prediction blind test

R. Alex Mayo,<sup>a</sup> Alastair J. A. Price,<sup>a</sup> Alberto Otero-de-la-Roza<sup>b</sup> and Erin R. Johnson<sup>a\*</sup>

Received 17 January 2024

Accepted 27 March 2024

Edited by A. Nangia, CSIR–National Chemical Laboratory, India

This article is part of a collection of articles covering the seventh crystal structure prediction blind test.

**Keywords:** crystal structure prediction; density-functional theory; London dispersion; free-energy correction.

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

<sup>a</sup>Department of Chemistry, Dalhousie University, 6243 Alumni Crescent, Halifax, Nova Scotia, B3H 4R2, Canada, and

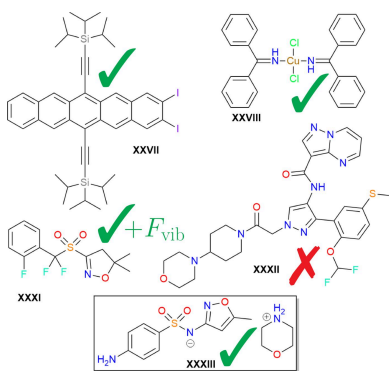
<sup>b</sup>Departamento de Química Física y Analítica and MALTA-Consolider Team, Facultad de Química, Universidad de Oviedo, 33006 Oviedo, Spain. \*Correspondence e-mail: erin.johnson@dal.ca

The seventh blind test of crystal structure prediction (CSP) methods substantially increased the level of complexity of the target compounds relative to the previous tests organized by the Cambridge Crystallographic Data Centre. In this work, the performance of density-functional methods is assessed using numerical atomic orbitals and the exchange-hole dipole moment dispersion correction (XDM) for the energy-ranking phase of the seventh blind test. Overall, excellent performance was seen for the two rigid molecules (XXVII, XXVIII) and for the organic salt (XXXIII). However, for the agrochemical (XXXI) and pharmaceutical (XXXII) targets, the experimental polymorphs were ranked fairly high in energy amongst the provided candidate structures and inclusion of thermal free-energy corrections from the lattice vibrations was found to be essential for compound XXXI. Based on these results, it is proposed that the importance of vibrational free-energy corrections increases with the number of rotatable bonds.

## 1. Introduction

First-principles molecular crystal structure prediction (CSP) is a grand challenge in the field of computational chemistry and is fruitful ground for ongoing method development. To assess the performance of various methods for both the structure generation and energy ranking stages of CSP, the Cambridge Crystallographic Data Centre (CCDC) regularly holds blind test competitions. Solved but unpublished crystal structures for a small set of compounds are held in reserve by the CCDC and participants have a set period of time to submit predicted crystal structures given only chemical diagrams of the compounds. There have been six previous blind tests to date (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005, 2009; Bardwell *et al.*, 2011; Reilly *et al.*, 2016), with the seventh held between 2020 and 2022 (Hunnisett *et al.*, 2024a; Hunnisett *et al.*, 2024b).

In the fourth (Day *et al.*, 2009), fifth (Bardwell *et al.*, 2011), and sixth (Reilly *et al.*, 2016) blind tests, dispersion-corrected density-functional theory (DFT) was shown to provide large improvements in energy ranking of candidate crystal structures compared to classical force-field methods. The Neumann group was the first to use DFT in the blind tests with their own empirical dispersion correction (Neumann & Perrin, 2005; Neumann *et al.*, 2008; Kendrick *et al.*, 2011) and applied their methodology to the earlier blind tests retroactively (Asmadi *et al.*, 2009). The more popular D3 (Grimme *et al.*, 2010, 2011) and MBD (Tkatchenko *et al.*, 2012; Ambrosetti *et al.*, 2014)



**Table 1**

Structure numbers for the experimental seed structures provided by the CCDC.

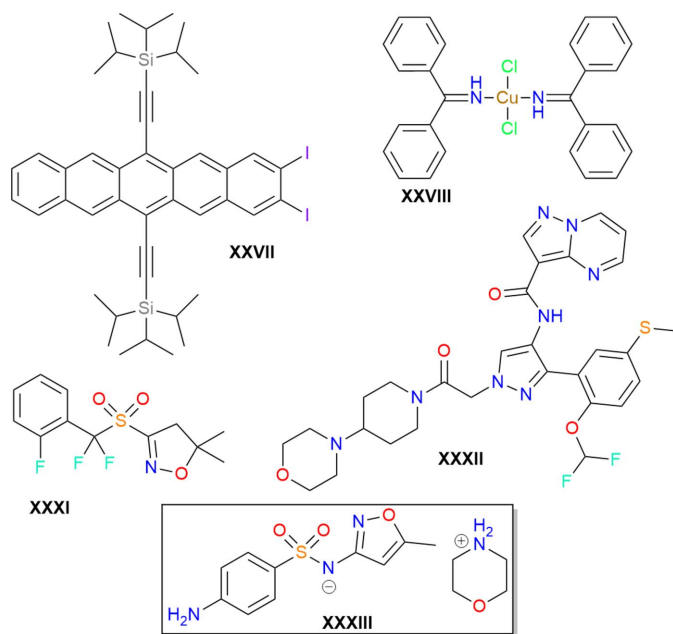
Also shown are three metrics to compare the provided and DFT-optimized crystal structures; these are the RMSD(20) obtained from the COMPACT algorithm (Motherwell & Chisholm, 2005) implemented in *Mercury* (Macrae *et al.*, 2020), as well as the powder difference (PWDF) and variable-cell powder difference (VC-PWDF) (Mayo *et al.*, 2022) evaluated using *critic2* (Otero-de-la-Roza *et al.*, 2014).

Compound	Form	Number	RMSD(20) (Å)	PWDF	VC-PWDF
XXVII	LT	28 = 61	0.483	0.049	0.004
XXVIII	–	144†	0.098	0.031	0.004
XXXI	A <sub>Maj</sub>	98	0.124	0.042	0.004
	A <sub>Min</sub>	1	0.241	0.098	0.012
	B	25	0.174	0.038	0.005
	C	89	0.106	0.032	0.004
XXXII	A <sub>Maj</sub>	317	0.137	0.023	0.002
	B <sub>RT</sub>	232	0.265	0.036	0.003
XXXIII	A	233	0.114	0.033	0.005
	B	452	0.159	0.061	0.006

† 145 and 207 were also equivalent to this structure after geometry optimization.

dispersion corrections have been used by other participants in submissions to the subsequent blind tests and also showed good performance (Brandenburg & Grimme, 2016; Reilly *et al.*, 2016; Hunnisett *et al.*, 2024a).

Dispersion corrections are routinely paired with generalized gradient approximation (GGA) functionals when modelling molecular crystals. However, GGAs suffer from delocalization error (Cohen *et al.*, 2008b; Kim *et al.*, 2013; Bryenton *et al.*, 2023), which can lead to overstabilization of organic salts, halogen bonds, cooperative hydrogen-bonding networks, and extended  $\pi$  conjugation. Delocalization error can be reduced through subsequent single-point energy



**Figure 1**

The five compounds considered for the energy-ranking phase of the seventh CSP blind test. XXXIII is an organic salt, with both ions shown within the box present in a 1:1 stoichiometric ratio.

evaluations using hybrid functionals, for example with PBE0-MBD, as has been used in the sixth blind test and several other CSP studies (Marom *et al.*, 2013; Shtukenberg *et al.*, 2017; Hoja & Tkatchenko, 2018; Hoja *et al.*, 2019; Mortazavi *et al.*, 2019). Quite recently, we assessed the performance of our exchange-hole dipole moment (XDM) dispersion correction (Johnson, 2017; Price *et al.*, 2023b) in combination with hybrid functionals for the compounds appearing in all six previous blind tests (Price *et al.*, 2023a). While delocalization error adversely affects the energy ranking provided with GGA functionals for organic salts in the fifth and sixth CSP blind tests (Whittleton *et al.*, 2017a), and leads to errors in the intramolecular conformational energies of many conjugated molecules (Whittleton *et al.*, 2017b; Greenwell & Beran, 2020; Greenwell *et al.*, 2020; Beran *et al.*, 2022), XDM-corrected hybrid functionals were shown to yield improved performance.

One of the key outcomes from discussion surrounding the sixth blind test was the idea of surpassing ‘zeroth-order’ CSP (Price, 2018), using only the electronic energies, by inclusion of thermal free-energy corrections (Hoja *et al.*, 2017). Such corrections have previously been used to convert experimental sublimation enthalpies to lattice energies in development of the C21 and X23 benchmarks of small-molecule crystals (Otero-de-la-Roza & Johnson, 2012; Reilly & Tkatchenko, 2013; Dolgonos *et al.*, 2019), as well as for studies of aspirin polytypes (Reilly & Tkatchenko, 2014). However, the importance of these corrections for CSP has been a matter of some debate in the literature, with (Nyman & Day, 2015) showing they are generally small – less than  $2 \text{ kJ mol}^{-1}$  in 94% of cases taken from a set of 508 polymorphic compounds – using a distributed multipole force field. For a set of 17 polymorph pairs, DFT calculations confirmed these corrections to be  $\sim 2 \text{ kJ mol}^{-1}$  or less (Weatherby *et al.*, 2022), although this study was limited to small, rigid molecules. DFT thermal free-energy corrections were also found to be of comparable magnitude for polymorphs of coumarin (Shtukenberg *et al.*, 2017) and rotigotine (Mortazavi *et al.*, 2019). Conversely, analogous calculations on the compounds from the sixth blind test occasionally showed significantly larger thermal free-energy corrections, up to  $8.4 \text{ kJ mol}^{-1}$  for the highly flexible compound XXIII (2-[[4-(3,4-dichlorophenethyl)phenyl]amino]benzoic acid; Hoja & Tkatchenko, 2018; Hoja *et al.*, 2019).

The seventh CSP blind test was separated into two phases, with the first focusing on structure generation and the second on energy ranking. In this second phase, sets of candidate crystal structures were provided by the CCDC for the five compounds shown in Fig. 1. These sets comprised 100 structures each for compounds XXVII and XXXI, and 500 structures each for compounds XXVIII, XXXII, and XXXIII. Within each set, one or more seed structure(s) (see Table 1) were provided that corresponded to the experimental crystal structure(s) after geometry optimization with a CSD-tailored force field (Hunnisett *et al.*, 2024a). For compound XXVII, two equivalent instances (see below) of the low-temperature experimental form were included as seed structures (rotation

of the isopropyl groups leads to a disordered form at high temperature). For the copper compound, XXVIII, only one polymorph was isolated experimentally and included as a seed structure. For the agrochemical compound XXXI, the provided set contained four experimental seed structures. One of these seeds corresponds to form B, which was found to be the most stable polymorph experimentally at low temperatures. At high temperature, a disordered form A became more stable; both the major and minor components of this form were included as seed structures. The final experimental seed structure for XXXI corresponds to the desolvate form C. For the large, flexible, drug-like compound XXXII, two experimental seed structures were provided, corresponding to the more stable form B and the major component of the less stable, disordered form A. Lastly, for the organic salt, XXXIII, two experimental seed structures were provided, with form B again more stable than form A, which was found to be a disappearing polymorph.

The present work reports on the performance of several XDM-corrected density functionals for the energy ranking phase of the seventh CSP blind test. Within this phase, our group was one of ten that employed dispersion-corrected DFT, with four of these using hybrid functionals; four groups also included thermal free-energy corrections (Hunnisett *et al.*, 2024a). Due to time constraints, we were not able to include free-energy corrections in our blind test submission, but do include them here for selected low-energy structures of compounds XXXI and XXXII. In this article, we will focus only on presentation of our own group's results. Interested readers are directed to Hunnisett *et al.* (2024a) for a comparison of the results across groups. Our methodology was able to rank the most stable experimentally isolated polymorph as the lowest in energy for two of the five compounds considered (XXVIII and XXXIII), and second lowest for one additional compound (XXVII). However, for the other two compounds (XXXI and XXXII) the isolated polymorphs were ranked substantially higher energetically. For compound XXXI, this result is unsurprising for form C since it possesses large crystal voids, while form B is correctly predicted to be the most stable polymorph upon inclusion of thermal free-energy corrections arising from the lattice vibrations. Including thermal free-energy corrections also yields an improved ranking for compound XXXII, and we conjecture that thermal and kinetic effects become increasingly important as the number of rotatable bonds increases.

## 2. Computational methods

### 2.1. Crystal structure comparison

A key concern in assessing CSP methods is how to quantify whether a candidate structure is the same or different when compared to the experimental polymorph. By far the most common method for crystal structure comparison is the COMPACK algorithm (Motherwell & Chisholm, 2005), implemented in the CCDC's *Mercury* program (Macrae *et al.*, 2020). COMPACK matches molecules within a given cluster

size (commonly 20) by generating an optimal overlay that minimizes the root-mean-square deviation (RMSD) in the atomic positions. However, we have recently illustrated a problem with COMPACK that precludes its routine application to molecules with several highly branched functional groups (Mayo *et al.*, 2022), such as compound XXVII in the seventh blind test. This problem likely arises from the application of a graph-matching algorithm (Ullmann, 1976) to identify matching molecules within the structures and align them.

COMPACK can also be quite sensitive to choices of cutoff values (Mayo & Johnson, 2021; Mayo *et al.*, 2022) and cluster sizes (Mayo *et al.*, 2022) that make its common use as a black-box method concerning. As a result, we have argued that distance-based methods like COMPACK should be used in conjunction with comparison methods based on powder X-ray diffraction (PXRD) to provide more certainty in the comparison results than can be achieved with either type of method alone.

PXRD-based methods quantify structure similarity using the overlap of the powder diffractograms of a pair of crystal structures. Among other ways, the overlaps can be computed using de Gelder's cross-correlation function (de Gelder *et al.*, 2001). The dissimilarity metric is then the powder pattern difference (PWDF), defined as one minus the overlap, such that exactly matching structures have a PWDF score of zero. However, PXRD-based comparison methods are only effective for pairs of crystal structures obtained under the same experimental conditions, or optimized with the same level of theory for *in silico* generated structures. This is because the PXRD peak locations are highly sensitive to the unit-cell dimensions, which change depending on temperature, pressure, and the particular choice of computational method. To overcome this limitation, we developed the variable-cell powder difference (VC-PWDF) method (Mayo *et al.*, 2022). VC-PWDF works by exploring a range of possible unit-cell definitions and distorting the cell vectors before evaluating the powder pattern difference. Overall, VC-PWDF was shown to agree with COMPACK in over 97% of almost 45 000 individual crystal-structure comparisons in the CSD (Mayo *et al.*, 2022), although it can provide incorrect classification of polytypes and conformational phases because of the similarity in their unit cells.

In this work, we employ a mix of RMSD(20), PWDF, and VC-PWDF methods for structure comparison. The RMSD(20) values were obtained using the Python API interface to *Mercury* (Macrae *et al.*, 2020), while PWDF and VC-PWDF scores were obtained using the *critic2* program (Otero-de-la-Roza *et al.*, 2014).

### 2.2. DFT calculations

The geometries of all provided candidate crystal structures were fully optimized using an in-house modified version of the *FHI-aims* program (version 210513; Blum *et al.*, 2009). Geometry optimizations were performed using the B86bPBE functional (Becke, 1986; Perdew *et al.*, 1996) with the XDM

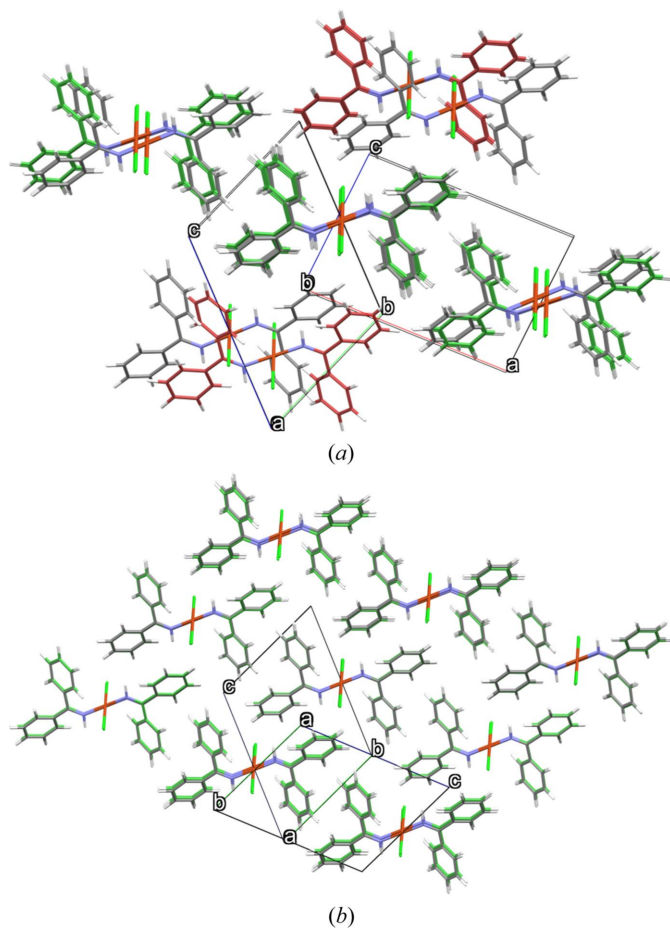
dispersion correction (Price *et al.*, 2023b), with the ‘lightdenser’ basis set and integration grid, and a relaxation convergence threshold of  $0.005 \text{ eV \AA}^{-1}$ . Subsequent single-point energy calculations were then performed on all optimized structures using either the hybrid B86bPBE-25X or B86bPBE-50X functionals with the same lightdenser basis set and grid settings, as well as with the B86bPBE functional with the tight basis settings, all with XDM dispersion. Price *et al.* (2023b) give the relevant damping parameters. Based on previous tests (Blum *et al.*, 2009), and excellent agreement with planewave results (Lejaeghere *et al.*, 2016; Price *et al.*, 2023b), there should be negligible basis-set superposition error. For compound XXVIII, all calculations were run for a ferromagnetic configuration with one unpaired electron per copper atom. The single unpaired *d* electron is evident since  $\text{Cu}^{\text{II}}$  will have an electronic configuration of  $[\text{Ar}]3d^9$ , and a ferromagnetic configuration is required in order to have a proper spin eigenstate that can be well described by Kohn–Sham DFT (Cohen *et al.*, 2008a).

We note that there are differences in the relative energies and computational timings between this work and our blind test submission, due to the time constraints for the blind test

and the fact that we were still making changes to the code when its second phase began. In this work, all geometry relaxations started from the crystallographic information files (CIF) provided by the CCDC with the finalized XDM implementation reported by Price *et al.* (2023b) and relaxed with a convergence threshold of  $0.005 \text{ eV \AA}^{-1}$ . However, in our blind test submission, we had relaxed all structures with a much looser threshold of  $0.025 \text{ eV \AA}^{-1}$ , and only completed relaxations with the  $0.005 \text{ eV \AA}^{-1}$  threshold for candidates lying within  $1.5 \text{ kcal mol}^{-1}$  ( $6.28 \text{ kJ mol}^{-1}$ ) of the minimum obtained with the looser convergence criterion. For the most part, the tighter convergence threshold did not substantially alter the relative energies. However, it did result in large geometry and energy changes for a handful of structures. One key example is structures 145 and 207 for compound XXVIII. Here, the initial structures differed from the experimental polymorph by a translation of some of the molecules in the unit cell (see Fig. 2) and they were (effectively degenerate) high-energy structures, lying  $>30 \text{ kJ mol}^{-1}$  above the minimum in our blind test submission. However, after full optimization with the tighter convergence threshold, the two structures became duplicate matches to the experimental form. Presumably, with the looser criteria, the optimization was deemed converged when the structure was occupying some nearly flat shoulder of a high-dimensional potential energy surface. With tighter criteria, the optimization continued and proceeded to follow a path with small negative gradients until it could move away from the flat shoulder and find a much more stable local minimum. This should serve as a warning to ensure sufficiently tight convergence criteria during geometry optimizations for CSP, despite the appeal of using looser criteria to reduce computational cost.

### 2.3. Harmonic vibrational calculations

To apply thermal free-energy corrections in the harmonic approximation, vibrational frequency (phonon) calculations were performed for the 16 lowest-energy candidates of compound XXXI and the 26 lowest-energy candidates of compound XXXII, according to the GGA/light ranking. These numbers correspond to the ranks of the least stable of the experimental seed structures with the GGA/light level of theory (excluding the high-energy solvate for compound XXXI). The phonon calculations used the finite-difference approach as implemented in the *phonopy* package (Togo & Tanaka, 2015). The energies for all supercells were again evaluated with B86bPBE-XDM, lightdenser setting, using *FHI-aims* (Blum *et al.*, 2009). For compound XXXI, the thermal free-energy corrections were converged to within  $0.5 \text{ kJ mol}^{-1}$  per molecule with respect to supercell size. For compound XXXII, the thermal free-energy corrections were only converged to within  $1.0 \text{ kJ mol}^{-1}$  per molecule due to the large unit-cell sizes, which made calculations on larger supercells intractable. Lists of the candidate structures carried forward to phonon calculations, and the corresponding supercells used, are given in the supporting information.



**Figure 2**  
COMPACT overlays of structure 145 of compound XXVIII and the reference experimental polymorph. Results are shown for the initial structure provided by the CCDC (top) and after full geometry optimization (bottom) with a convergence threshold of  $0.005 \text{ eV \AA}^{-1}$ .



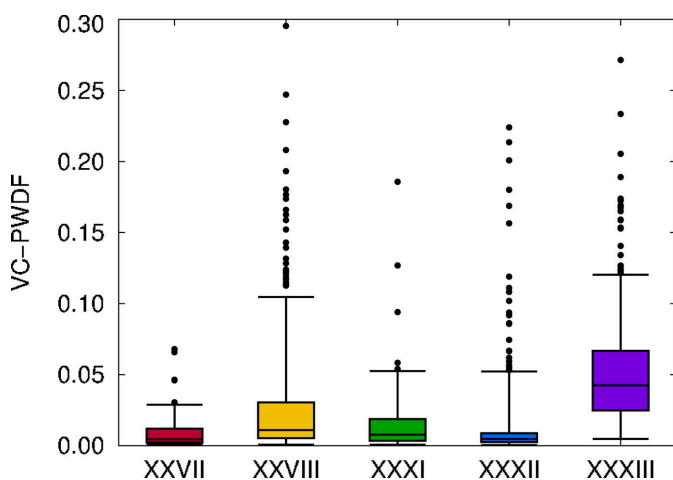
### 3. Results and discussion

#### 3.1. Structure changes

Because the candidate crystal structures provided by the CCDC were all optimized with a CSD-tailored force field, they may not correspond to stable minima with DFT. To assess how much structural change occurred during the DFT geometry optimizations, we evaluated the VC-PWDF score between the initial and final geometries for each candidate structure of each compound. The results are summarized by the box plot in Fig. 3, where the boxes span the interquartile range of the VC-PWDF values. The whiskers span 95% of the data, while outliers are shown as individual points. For context, a VC-PWDF score of  $< 0.03$  is usually a good indicator of significant structural similarity for rigid molecules and  $< 0.05$  for more flexible molecules (Mayo & Johnson, 2021; Mayo *et al.*, 2022).

From Fig. 3, compound XXXIII shows the largest structural changes during optimization, followed by compound XXVIII, despite both systems comprising fairly rigid molecules. It is reasonable that the force field would be less accurate for the organic salt (XXXIII) due to the importance of electrostatic interactions. Similarly, the organometallic copper complex (XXVIII) may be problematic due to difficulties in parameterization for copper and the unpaired electron in the complex. Instances of large geometric changes with VC-PWDF  $> 0.1$  also occurred for some candidate structures of compounds XXXI and XXXII, although the ranges spanned by 95% of the points are much narrower compared to those for XXVIII and XXXIII.

As it is fairly common for force fields to predict more energy minima than DFT (Price, 2013; Neumann & van de Streek, 2018; Francia *et al.*, 2020; Butler & Day, 2023), we also assess how many of the candidate structures were duplicates (*i.e.* corresponding to the same polymorph) before and after geometry optimization. This provides insight into whether different force-field local minima correspond to a single DFT



**Figure 3** Box and whisker plot showing the extent of structural change resulting from DFT geometry optimization, as quantified by the VC-PWDF score between the provided and optimized structures for each candidate. The boxes show the interquartile ranges, the whiskers encompass 95% of the data, and remaining outliers are shown as individual points.

**Table 2**

Number of duplicate structures for each compound, as quantified by a PWDF score of 0.01 or less between pairs of candidates.

Results are shown for both the initial structures provided by the CCDC and those after DFT geometry optimization.

Compound	No. of duplicates	
	CCDC	DFT
XXVII	2	2
XXVIII	0	40
XXXI	0	4
XXXII	2	8
XXXIII	0	27

minimum. Since we are comparing structures that were all optimized with the same method (*i.e.* either the CSD-tailored force field or DFT), we do not need to consider variable cells and we can use the PWDF scores as our similarity metric. Here, two structures are deemed equivalent if their PWDF value is  $< 0.01$ .

The results in Table 2 show the number of duplicate structures before and after optimization. For the CCDC provided structures, there were only two sets of duplicates each for XXVII (structure IDs 28/61 and 63/77) and XXXII (IDs 58/66 and 169/488). These may have initially been deemed different by COMPACK due to problems with the graph-matching algorithm for highly branched, pseudosymmetric molecules (Mayo *et al.*, 2022), such as XXVII, or due to the choice of distance and angle tolerances. After optimization, at least two pairs of duplicate structures are seen for all the compounds, with many duplicate structures occurring for XXVIII and XXXIII in particular. This finding is consistent with the larger structural changes seen for these two compounds in Fig. 3. An example was already discussed in Section 2.2 for compound XXVIII, where two structures (IDs 145 and 207) optimize to become equivalent to the seed structure for the experimental form (ID 144, see Fig. 2).

#### 3.2. BT7-predicted GGA landscapes

We consider now the computed crystal energy landscapes obtained from the geometry relaxations with B86bPBE-XDM and the light NAO basis set; these results are shown in Fig. 4. For compound XXVII, two identical seed structures (IDs 28 and 61), chosen to correspond to the low-temperature experimental form, were included in the set of candidates provided by the CCDC. This seed structure was ranked second lowest on the computed crystal-energy landscape in Fig. 4(a). As it lies  $< 1.5 \text{ kJ mol}^{-1}$  above the global minimum structure (ID 58), which is well within the accepted energy window for the stability of experimentally observable polymorphs (Nyman & Day, 2015), we consider this to be a successful prediction. Additionally, this energy difference is small enough that thermal free-energy corrections may reverse the ranking (Nyman & Day, 2016, 2015; Weatherby *et al.*, 2022). Rotations of the isopropyl groups, which become disordered at room temperature, can also significantly affect the energy ranking (Hunnisett *et al.*, 2024a).

## crystal structure prediction

For the copper compound, XXVIII, only a single polymorph was identified experimentally. The crystal-energy landscape in Fig. 4(b) shows that the experimental seed structure was ranked as the most stable candidate, again indicating a successful crystal structure prediction. As noted above, two additional candidates became equivalent to the seed structure after geometry optimization.

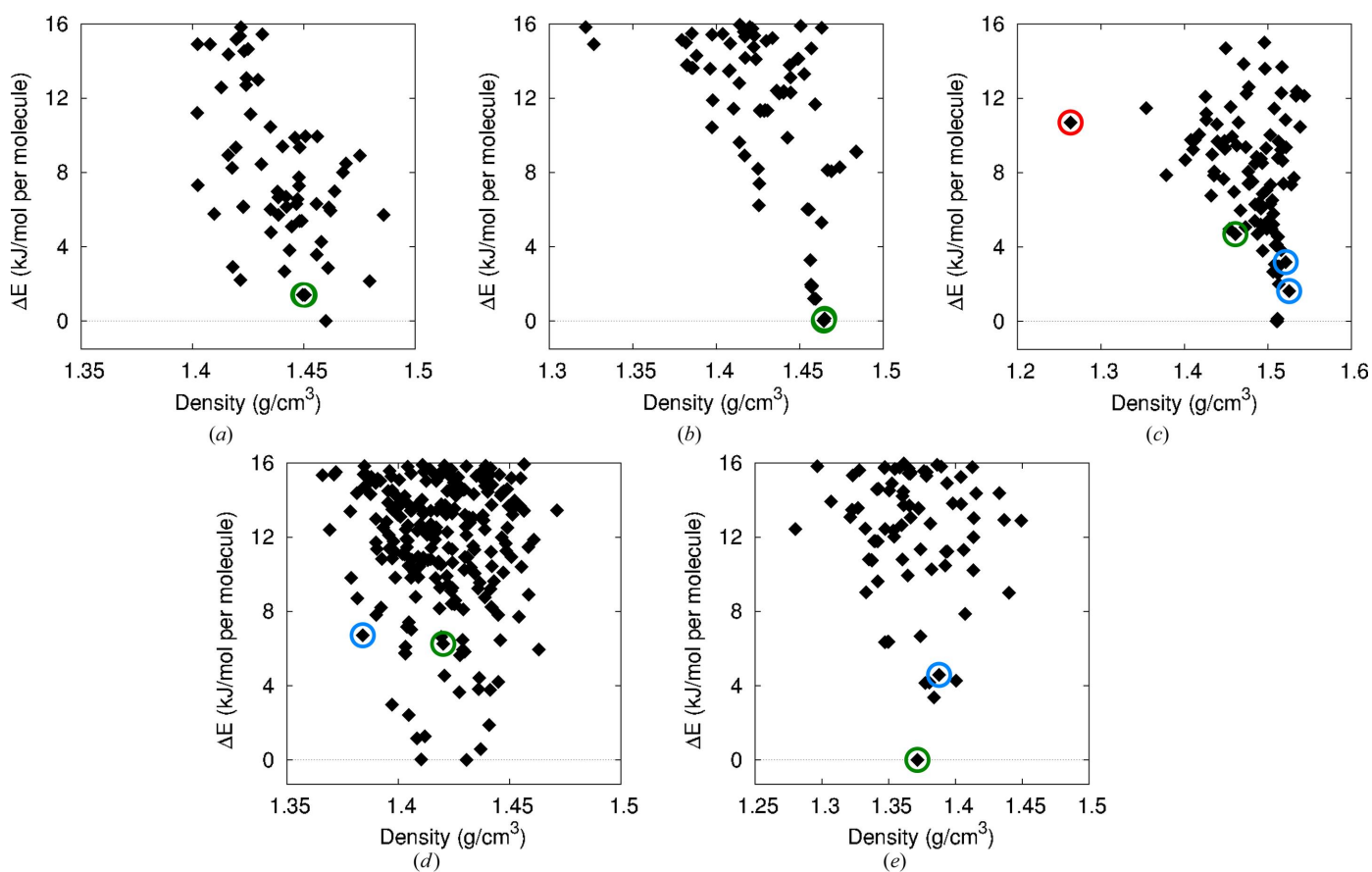
For compound XXXI, Fig. 4(c) shows that the DFT calculations predict form C to lie quite high in energy, as expected due to the large voids in this crystal structure that formed upon solvent evaporation. While this form lies  $10.7 \text{ kJ mol}^{-1}$  above the global minimum, significantly outside the usual 'polymorph window' ( $< 7.2 \text{ kJ mol}^{-1}$  above the global minimum in 95% of cases) (Nyman & Day, 2015), it has been shown that similar high-energy, low-density structures can be isolated from solvates (Yang & Day, 2021). Turning to the more stable experimental forms, our calculations predict the seed structures for form B and the major/minor components of form A to all lie within  $5 \text{ kJ mol}^{-1}$  of the global minimum. However, the predicted stability ordering is incorrect, with form B ranked as the least stable, lying  $4.7 \text{ kJ mol}^{-1}$  above the global minimum and  $3.1 \text{ kJ mol}^{-1}$  above the major component of form A.

As shown in Fig. 4(d), our calculations were quite unsuccessful in identifying the experimental forms of compound XXXII as being low in energy amongst the candidates. Here, both forms were found to lie  $6\text{--}7 \text{ kJ mol}^{-1}$  in energy above the global minimum. The more stable form B was ranked 22nd in energy, while form A was ranked 26th, recovering the correct experimental ordering, but with a predicted energy difference of only  $0.4 \text{ kJ mol}^{-1}$  separating the two forms.

Finally, for compound XXXIII, Fig. 4(e) shows that form B is correctly ranked as the most stable of the candidate structures, constituting another successful prediction. Form A is also found to be a fairly stable structure, ranked sixth in energy and lying  $< 5 \text{ kJ mol}^{-1}$  above form A, which is still within the accepted polymorph window (Nyman & Day, 2015).

### 3.3. Dependence on functional and basis set

In previous benchmarking of XDM-corrected density functionals for the compounds comprising the first six blind tests, we found that hybrid functionals outperformed GGAs in some cases (Price *et al.*, 2023a), specifically those with significant delocalization error (Bryenton *et al.*, 2023). In particular, delocalization error was found to affect the ranking of the



**Figure 4**

Computed crystal energy landscapes obtained using B86bPBE-XDM with the light basis set: (a) XXVII, (b) XXVIII, (c) XXXI, (d) XXXII, (e) XXXIII. The green circles indicate the structure that is the best match to the most-stable experimental polymorph, while the blue circles indicate the structure of a second, less-stable polymorph, including major and minor components for compound XXXI. For this same compound, an additional low-density form with large crystal voids, indicated by the red circle, was identified after removal of solvent from a solvate structure.

**Table 3**

Summary of computational results.

For each compound, the energy rank of all isolated polymorphs and their relative energies ( $\Delta E$  in  $\text{kJ mol}^{-1}$  per molecule, or per molecule pair in the case of the salt) above the minimum on the corresponding crystal energy landscape are shown for each of the methods used.

Compound	Form	Light					Tight						
		GGA		25X		50X	GGA		25X		50X		
		Rank	$\Delta E$	Rank	$\Delta E$	Rank	$\Delta E$	Rank	$\Delta E$	Rank	$\Delta E$	Rank	$\Delta E$
XXVII	LT	2	1.4	2	1.0	2	1.0	3	1.4	2	1.2	3	1.0
XXVIII	–	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0
XXXI	A <sub>Maj</sub>	3	1.6	5	2.7	7	3.9	3	0.5	3	1.6	7	2.8
	A <sub>Min</sub>	9	3.2	10	4.4	15	5.8	4	1.6	8	2.9	9	4.2
	B	16	4.7	16	5.3	13	5.7	21	4.6	18	5.3	15	5.6
	C	77	10.7	73	11.5	66	11.9	80	10.2	76	10.9	66	12.2
XXXII	A <sub>Maj</sub>	26	6.7	24	8.8	30	11.2	34	9.6	31	10.2	32	10.8
	B <sub>RT</sub>	22	6.3	28	9.3	41	12.6	20	8.6	30	10.2	39	11.6
XXXIII	A	6	4.6	5	4.5	5	5.2	6	5.7	5	5.6	6	6.4
	B	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0

organic salts (compounds XIX and XXIV) (Price *et al.*, 2023a; Whittleton *et al.*, 2017a), as well as for compound X where the intramolecular conformational energy has a substantial contribution to the ranking (Price *et al.*, 2023a; Greenwell & Beran, 2020; Whittleton *et al.*, 2017b). Similarly, significant errors in GGA energy rankings have been found for other flexible molecules due to errors in the conformational energy (Greenwell & Beran, 2020; Greenwell *et al.*, 2020; Beran *et al.*, 2022). As a result, we performed single-point energy calculations with the B86bPBE-25X-XDM and B86bPBE-50X-XDM functionals, both with the light basis set, which was expected to lead to more accurate ranking relative to the B86bPBE-XDM results alone. We also performed single-point energy calculations with B86bPBE-XDM and the tight basis set to identify if the errors seen for compounds XXXI and XXXII were a result of basis-set incompleteness. This allowed us to approximate hybrid functional results with the tight basis using an additive scheme (Hoja & Tkatchenko, 2018; Hoja *et al.*, 2019; Price *et al.*, 2023b):

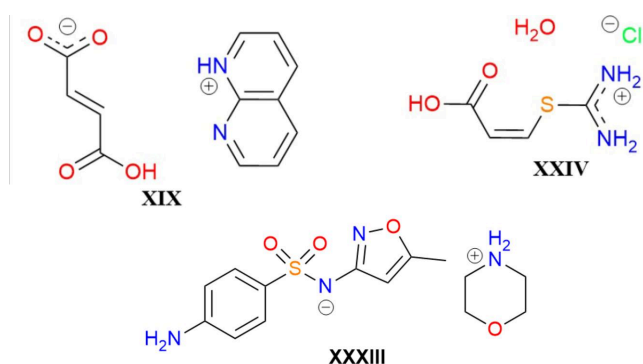
$$E(\text{hybrid/tight}) \approx E(\text{GGA/tight}) + E(\text{hybrid/light}) - E(\text{GGA/light}).$$

The identities of the minimum-energy structures obtained with each combination of functional and basis set are given in the supporting information. A comparison of the overall rankings and relative energies of the experimental seed structures, relative to the respective global minimums, are shown in Table 3. Note that the rankings were determined by taking some pairs of structures as equivalent: the experimental matches, structures 28 and 61, for compound XXVII (POWDIFF = 0.0004); the experimental matches, structures 144, 145 and 207, for compound XXVIII (POWDIFF  $\leq$  0.0005); and the DFT energy minima, structures 17 and 59, for compound XXXI (POWDIFF = 0.0039). Full crystal-energy landscapes are provided in the supporting information.

For the blind test submission, we selected results from the B86bPBE-25X functional for compounds XXVII and XXVIII, and results from B86bPBE-50X for the remaining three compounds, as we expected them to potentially be more

susceptible to delocalization error due to being able to adopt a range of stable conformations (XXXI and XXXII) or being an organic salt (XXXIII). However, the results in Table 3 show that these were not the optimal choices. Overall, all three functionals give very similar ranking for compounds XXVII, XXVIII, and XXXIII. It would seem that the crystal energy landscapes of these three compounds are relatively easy to model accurately with dispersion-corrected DFT, since the ranking does not depend on the choice of functional in any consequential way. This consistence was expected for the first two compounds, as they are relatively rigid, non-polar molecules, and their crystals are therefore unlikely to have much inherent delocalization error.

In light of our previous results for the salt compounds XIX and XXIV, the lack of functional dependence was somewhat surprising for compound XXXIII. Therefore, we conclude that not all organic salts are equally affected by delocalization error. One likely explanation for the difference is that delocalization error frequently manifests in cases with extended conjugation (Greenwell & Beran, 2020; Greenwell *et al.*, 2020; Beran *et al.*, 2022; Whittleton *et al.*, 2017b) and cooperative hydrogen bonding (Whittleton *et al.*, 2017a), such as for COOH groups where the OH can act as a hydrogen-bond donor and the =O can act as a hydrogen-bond acceptor. As seen from their chemical structures, shown in Fig. 5, moieties

**Figure 5**

Chemical structures of the three organic salts considered in the CSP blind tests: XIX from BT5, XXIV from BT6, and XXXIII from BT7.

in salts XIX and XXIV both exhibit extended conjugation, with all species being planar, and one ion in each possesses COOH groups that form cooperative hydrogen-bonds in the crystal (Whittleton *et al.*, 2017a; Price *et al.*, 2023a). Conversely, for XXXIII, the cation is not planar, the conjugation in the anion is broken by the SO<sub>2</sub> group, and neither can participate in cooperative hydrogen bonding.

We now move to the two flexible molecules, compounds XXXI and XXXII. In sharp contrast to the other three members of this blind test, these compounds proved extremely challenging for DFT and also exhibit more functional dependence. From Table 3, the experimental forms of compounds XXXI and XXXII are ranked increasingly higher in energy (with respect to the global minimum) with increasing fractions of exact exchange. Increasing exact exchange does lower the energy of form B relative to the major component of form A for compound XXXI, but not enough to recover the correct ordering. However, for compound XXXII, form B was correctly predicted to be more stable than the major component of form A with the GGA functional and this ordering is reversed with both hybrids. Therefore, our choice of 50% exact exchange for the blind test submission was unfortunate here and this demonstrates how difficult it can be to know, a priori, whether delocalization error will affect a CSP landscape. Finally, we find that the tight basis set does not improve the rankings, widening the energy gap between forms A and B for compound XXXI and raising the experimental forms of compound XXXII to substantially higher energies. This allows us to rule out basis-set incompleteness for the very electronegative F atoms appearing in these two compounds as the source of the poor rankings.

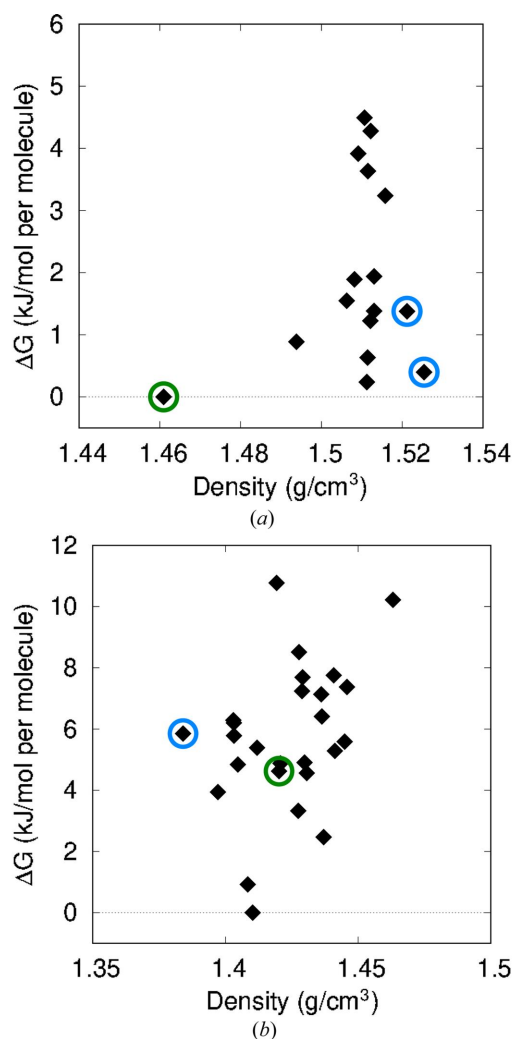
### 3.4. Vibrational free-energy corrections

To summarize our results so far, the DFT energy rankings were quite successful for three of the blind test compounds (XXVII, XXVIII and XXXIII), but performed much less well for the landscapes of XXXI and XXXII. Notably, these are both fairly large, flexible molecules with several rotatable bonds. This is reminiscent of compound XX from the fifth CSP blind test, another large, flexible molecule for which our DFT energy rankings showed considerable errors (Price *et al.*, 2023a). We previously conjectured that thermal free-energy corrections, neglected in zeroth-order CSP (Price, 2018), may be responsible for the poor energy ranking here, and this may also be the case for compounds XXXI and XXXII. To test this hypothesis, phonon calculations were performed for the 16 lowest-energy candidates for compound XXXI and the 26 lowest-energy candidates for compound XXXII at the B86bPBE-XDM/light level. These ranges of structures were selected to include all of the experimental seed structures, except for the high-energy solvate (see Table 3).

Fig. 6 shows the free-energy landscapes for compounds XXXI and XXXII, where the thermal free-energy corrections for a temperature of 298 K have been added to the B86bPBE-XDM/light electronic energies. Free-energy landscapes obtained using different base functionals and/or basis sets for

the electronic energies are shown in the supporting information. Fig. 7 shows the change in ranking of the considered structures upon inclusion of the thermal free-energy corrections.

The results in Figs. 6(a) and 7(a) show that inclusion of thermal free-energy terms corrects the energy ranking for compound XXXI. Using the electronic energies alone, form B was found to lie 4.7 kJ mol<sup>-1</sup> above the DFT minimum. However, using free energies, form B is successfully predicted to be the global minimum structure, while the major component of form A is ranked third, <0.5 kJ mol<sup>-1</sup> above form B. Experimentally, form B is the more stable form up to a temperature of 48–55°C (Hunnissett *et al.*, 2024b), after which the disordered form A is more stable. It should be noted, however, that this ranking only applies to the subset of

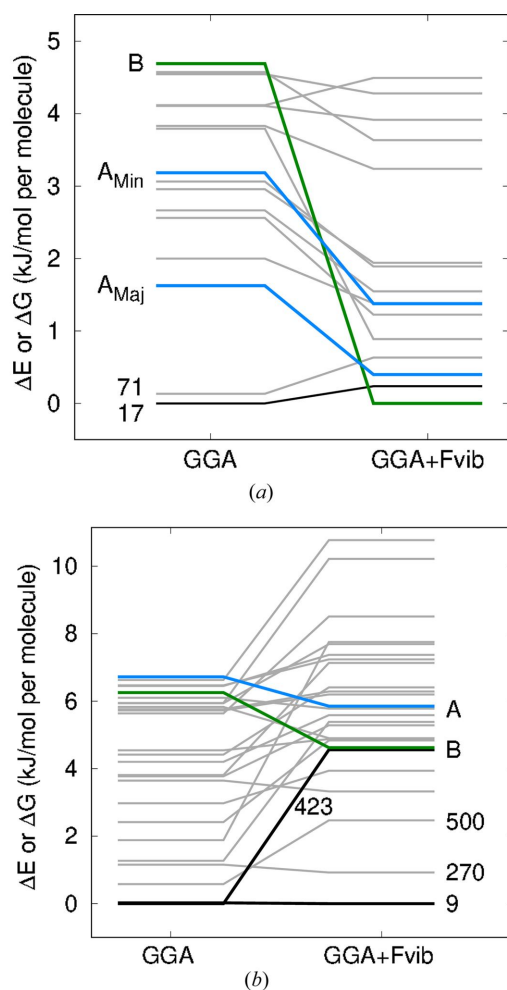


**Figure 6** Computed crystal free-energy landscapes for a temperature of 298 K obtained using B86bPBE-XDM with the light basis set: (a) XXXI and (b) XXXII. Phonon calculations to evaluate the thermal free-energy corrections were only performed for the 16 lowest-energy structures of compound XXXI and the 26 lowest-energy structures of compound XXXII. The green circles indicate the structure that is the best match to the most-stable experimental polymorph, while the blue circles indicate the structure of a second, less-stable polymorph, including major and minor components for compound XXXI.



structures for which free-energy calculations were performed. It is possible that some structures with even higher lattice energies would also be stabilized sufficiently by inclusion of free-energy corrections to affect these rankings.

Figs. 6(b) and 7(b) show that inclusion of thermal free-energy corrections also improves the ranking for compound XXXII. Here, form B is now ranked seventh with a  $\Delta G$  of  $4.6 \text{ kJ mol}^{-1}$ , while form A is ranked fifth with a  $\Delta G$  of  $5.9 \text{ kJ mol}^{-1}$ , in contrast to the rankings of 22nd and 26th in Table 3 obtained with electronic energies alone. However, the two experimental forms do remain fairly high in free energy above the global minimum. This finding suggests that (i) there are errors in our DFT treatment, (ii) anharmonicity and/or thermal expansion is important in determining the vibrational free-energy correction and the quasi-harmonic approximation is needed, (iii) kinetic and/or surface-energy effects prevent formation of the lower-energy candidates identified with DFT, or (iv) some combination of these factors (Price, 2013).



**Figure 7**  
Changes in polymorph ranking with B86bPBE-XDM/light upon inclusion of thermal free-energy energy corrections: (a) XXXI and (b) XXXII. The GGA results correspond to relative electronic energies, while GGA+Fvib corresponds to relative free energies for a temperature of 298 K.

#### 4. Conclusions

In this work, we applied XDM-corrected DFT methods to the energy ranking phase of the seventh CSP blind test. Large changes in some of the provided candidate structures were observed after DFT geometry optimization, particularly for compounds XXVIII and XXXIII. Additionally, many candidates collapsed to the same local minimum, giving duplicate structures after optimization. This points to the importance of developing improved force fields for the initial structure generation and geometry optimization steps of CSP. In terms of the DFT energy ranking itself, our methodology gave successful predictions of the most stable experimentally isolated polymorphs as being the lowest or second-lowest energy structures for compounds XXVII, XXVIII and XXXIII. However, the most stable experimental polymorphs were ranked relatively high in energy for compounds XXXI and XXXII.

We conclude that our DFT methodology provides good reliability for CSP studies of fairly rigid molecules, but there remain challenges for large and highly flexible pharmaceutical targets. We suggest that thermal free-energy corrections may be much more important for such structures than previously demonstrated for smaller, more rigid molecules (Nyman & Day, 2016; Nyman & Day, 2015; Weatherby *et al.*, 2022). While the thermal free-energy corrections are typically  $\leq 2 \text{ kJ mol}^{-1}$ , they amounted to as much as  $4.7$  and  $5.9 \text{ kJ mol}^{-1}$  for compounds XXXI and XXXII, respectively. Previous work found even larger thermal free-energy corrections, up to  $8.4 \text{ kJ mol}^{-1}$ , for compound XXIII in the sixth blind test (Hoja & Tkatchenko, 2018; Hoja *et al.*, 2019). Here, the most stable isolated polymorph for compound XXXI was successfully predicted as the global minimum with DFT upon inclusion of a vibrational free-energy correction. This correction improved the ranking for compound XXXII as well, although the experimental structures were still ranked sixth and 15th in energy. It is possible that corrections beyond the harmonic approximation are required, or that kinetic and/or surface energy effects prevent crystallization of the lower-energy candidates identified with DFT in this case.

#### Acknowledgements

ERJ thanks the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support of this research. RAM thanks the International Centre for Diffraction Data for a Ludo Frevel Crystallography Scholarship. All authors are grateful to the Atlantic Computing Excellence Network (ACENET) and Compute Canada for providing computational resources.

#### Funding information

AOR thanks the Spanish Ministerio de Ciencia e Innovación and the Agencia Estatal de Investigación, projects PGC2021-125518NB-I00 and RED2022-134388-T cofinanced by EU FEDER funds; the Principality of Asturias (FICYT), project

AYUD/2021/51036 cofinanced by EU FEDER; and the Spanish MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR for grant TED2021-129457B-I00.

## References

- Ambrosetti, A., Reilly, A. M., DiStasio, R. A. Jr & Tkatchenko, A. (2014). *J. Chem. Phys.* **140**, 18A508.
- Asmadi, A., Neumann, M. A., Kendrick, J., Girard, P., Perrin, M.-A. & Leusen, F. J. J. (2009). *J. Phys. Chem. B*, **113**, 16303–16313.
- 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.* **B67**, 535–551.
- Becke, A. D. (1986). *J. Chem. Phys.* **85**, 7184–7187.
- Beran, G. J. O., Sugden, I. J., Greenwell, C., Bowskill, D. H., Pantelides, C. C. & Adjiman, C. S. (2022). *Chem. Sci.* **13**, 1288–1297.
- Blum, V., Gehrke, R., Hanke, F., Havu, P., Havu, V., Ren, X., Reuter, K. & Scheffler, M. (2009). *Comput. Phys. Commun.* **180**, 2175–2196.
- Brandenburg, J. G. & Grimme, S. (2016). *Acta Cryst.* **B72**, 502–513.
- Bryantson, K. R., Adeleke, A. A., Dale, S. G. & Johnson, E. R. (2023). *WIREs Comput. Mol. Sci.* **13**, e1631.
- Butler, P. W. V. & Day, G. M. (2023). *Proc. Natl Acad. Sci. USA*, **120**, e2300516120.
- Cohen, A. J., Mori-Sánchez, P. & Yang, W. (2008a). *J. Chem. Phys.* **129**, 121104.
- Cohen, A. J., Mori-Sánchez, P. & Yang, W. (2008b). *Science*, **321**, 792–794.
- 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.* **B65**, 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.* **B61**, 511–527.
- Dolgonos, G. A., Hoja, J. & Boese, A. D. (2019). *Phys. Chem. Chem. Phys.* **21**, 24333–24344.
- Francia, N. F., Price, L. S., Nyman, J., Price, S. L. & Salvalaglio, M. (2020). *Cryst. Growth Des.* **20**, 6847–6862.
- Gelder, R. de, Wehrens, R. & Hageman, J. A. (2001). *J. Comput. Chem.* **22**, 273–289.
- Greenwell, C. & Beran, G. J. O. (2020). *Cryst. Growth Des.* **20**, 4875–4881.
- Greenwell, C., McKinley, J. L., Zhang, P., Zeng, Q., Sun, G., Li, B., Wen, S. & Beran, G. J. O. (2020). *Chem. Sci.* **11**, 2200–2214.
- Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. (2010). *J. Chem. Phys.* **132**, 154104.
- Grimme, S., Ehrlich, S. & Goerigk, L. (2011). *J. Comput. Chem.* **32**, 1456–1465.
- Hoja, J., Ko, H. Y., Neumann, M. A., Car, R., Distasio, R. A. & Tkatchenko, A. (2019). *Sci. Adv.* **5**, eaau3338.
- Hoja, J., Reilly, A. M. & Tkatchenko, A. (2017). *WIREs Comput. Mol. Sci.* **7**, e1294.
- Hoja, J. & Tkatchenko, A. (2018). *Faraday Discuss.* **211**, 253–274.
- Hunnisett, L. M., Francia, N., Nyman, J., Abraham, N. S., Aitipamula, S., Alkhalid, T., Almehairbi, M., Anelli, A., Anstine, D. M., Anthony, J. E., Arnold, J. E., Bahrami, F., Bellucci, M. A., Beran, G. J. O., Bhardwaj, R. M., Bianco, R., Bis, J. A., Boese, A. D., Bramley, J., Braun, D. E., Butler, P. W. V., Cadden, J., Carino, S., Červinka, C., Chan, E. J., Chang, C., Clarke, S. M., Coles, S. J., Cook, C. J., Cooper, R. I., Darden, T., Day, G. M., Deng, W., Dietrich, H., DiPasquale, A., Dhokale, B., van Eijck, B. P., Elsegood, M. R. J., Firaha, D., Fu, W., Fukuzawa, K., Galanakis, N., Goto, H., Greenwell, C., Guo, R., Harter, J., Helfferich, J., Hoja, J., Hone, J., Hong, R., Hušák, M., Ikabata, Y., Isayev, O., Ishaque, O., Jain, V., Jin, Y., Jing, A., Johnson, E. R., Jones, I., Jose, K. V. J., Kabova, E. A., Keates, A., Kelly, P. F., Klimeš, J., Kostková, V., Li, H., Lin, X., List, A., Liu, C., Liu, Y. M., Liu, Z., Lončarić, I., Lubach, J. W., Ludík, J., Maryewski, A. A., Marom, N., Matsui, H., Mattei, A., Mayo, R. A., Melkumov, J. W., Mladineo, B., Mohamed, S., Momenzadeh Abardeh, Z., Muddana, H. S., Nakayama, N., Nayal, K. S., Neumann, M. A., Nikhar, R., Obata, S., O'Connor, D., Oganov, A. R., Okuwaki, K., Otero-de-la-Roza, A., Parkin, S., Parunov, A., Podeszwa, R., Price, A. J. A., Price, L. S., Price, S. L., Probert, M. R., Pulido, A., Ramteke, G. R., Rehman, A. U., Reutzel-Edens, S. M., Rogal, J., Ross, M. J., Rumson, A. F., Sadiq, G., Saeed, Z. M., Salimi, A., Sasikumar, K., Sekharan, S., Shankland, K., Shi, B., Shi, X., Shinohara, K., Skillman, A. G., Song, H., Strasser, N., van de Streek, J., Sugden, I. J., Sun, G., Szalewicz, K., Tan, L., Tang, K., Tarczynski, F., Taylor, C. R., Tkatchenko, A., Touš, P., Tuckerman, M. E., Unzueta, P. A., Utsumi, Y., Vogt-Maranto, L., Weatherston, J., Wilkinson, L. J., Willacy, R. D., Wojtas, L., Woollam, G. R., Yang, Y., Yang, Z., Yonemochi, E., Yue, X., Zeng, Q., Zhou, T., Zhou, Y., Zubatyuk, R. & Cole, J. C. (2024). *Acta Cryst.* **B80**, <https://doi.org/10.1107/S2052520624008679>.
- Hunnisett, L. M., Nyman, J., Francia, N., Abraham, N. S., Adjiman, C. S., Aitipamula, S., Alkhalid, T., Almehairbi, M., Anelli, A., Anstine, D. M., Anthony, J. E., Arnold, J. E., Bahrami, F., Bellucci, M. A., Bhardwaj, R. M., Bier, I., Bis, J. A., Boese, A. D., Bowskill, D. H., Bramley, J., Brandenburg, J. G., Braun, D. E., Butler, P. W. V., Cadden, J., Carino, S., Chan, E. J., Chang, C., Cheng, B., Clarke, S. M., Coles, S. J., Cooper, R. I., Couch, R., Cuadrado, R., Darden, T., Day, G. M., Dietrich, H., Ding, Y., DiPasquale, A., Dhokale, B., van Eijck, B. P., Elsegood, M. R. J., Firaha, D., Fu, W., Fukuzawa, K., Glover, J., Goto, H., Greenwell, C., Guo, R., Harter, J., Helfferich, J., Hofmann, D. W. M., Hoja, J., Hone, J., Hong, R., Hutchison, G., Ikabata, Y., Isayev, O., Ishaque, O., Jain, V., Jin, Y., Jing, A., Johnson, E. R., Jones, I., Jose, K. V. J., Kabova, E. A., Keates, A., Kelly, P. F., Khakimov, D., Konstantinopoulos, S., Kuleshova, L. N., Li, H., Lin, X., List, A., Liu, C., Liu, Y. M., Liu, Z., Liu, Z.-P., Lubach, J. W., Marom, N., Maryewski, A. A., Matsui, H., Mattei, A., Mayo, R. A., Melkumov, J. W., Mohamed, S., Momenzadeh Abardeh, Z., Muddana, H. S., Nakayama, N., Nayal, K. S., Neumann, M. A., Nikhar, R., Obata, S., O'Connor, D., Oganov, A. R., Okuwaki, K., Otero-de-la-Roza, A., Pantelides, C. C., Parkin, S., Pickard, C. J., Pilia, L., Pivina, T., Podeszwa, R., Price, A. J. A., Price, L. S., Price, S. L., Probert, M. R., Pulido, A., Ramteke, G. R., Rehman, A. U., Reutzel-Edens, S. M., Rogal, J., Ross, M. J., Rumson, A. F., Sadiq, G., Saeed, Z. M., Salimi, A., Salvalaglio, M., Sanders de Almada, L., Sasikumar, K., Sekharan, S., Shang, C., Shankland, K., Shinohara, K., Shi, B., Shi, X., Skillman, A. G., Song, H., Strasser, N., van de Streek, J., Sugden, I. J., Sun, G., Szalewicz, K., Tan, B. I., Tan, L., Tarczynski, F., Taylor, C. R., Tkatchenko, A., Tom, R., Tuckerman, M. E., Utsumi, Y., Vogt-Maranto, L., Weatherston, J., Wilkinson, L. J., Willacy, R. D., Wojtas, L., Woollam, G. R., Yang, Z., Yonemochi, E., Yue, X., Zeng, Q., Zhang,

- Y., Zhou, T., Zhou, Y., Zubatyuk, R. & Cole, J. C. (2024). *Acta Cryst.* **B80**, <https://doi.org/10.1107/S2052520624007492>.
- Johnson, E. R. (2017). In *Non-covalent Interactions in Quantum Chemistry and Physics*, edited by A. Otero-de-la-Roza & G. A. DiLabio, ch. 5, pp. 169–194. Elsevier.
- Kendrick, J., Leusen, F. J. J., Neumann, M. A. & van de Streek, J. (2011). *Chem. A Eur. J.* **17**, 10736–10744.
- Kim, M.-C., Sim, E. & Burke, K. (2013). *Phys. Rev. Lett.* **111**, 073003.
- Lejaeghere, K., Bihlmayer, G., Björkman, T., Blaha, P., Blügel, S., Blum, V., Caliste, D., Castelli, I. E., Clark, S. J., Dal Corso, A., de Gironcoli, S., Deutsch, T., Dewhurst, J. K., di Marco, I., Draxl, C., Dulak, M., Eriksson, O., Flores-Livas, J. A., Garrity, K. F., Genovese, L., Giannozzi, P., Giantomassi, M., Goedecker, S., Gonze, X., Grånäs, O., Gross, E. K. U., Gulans, A., Gygi, F., Hamann, D. R., Hasnip, P. J., Holzwarth, N. A. W., Iuşan, D., Jochym, D. B., Jollet, F., Jones, D., Kresse, G., Koepf, K., Küçükbenli, E., Kvashnin, Y. O., Loch, I. L. M., Lubeck, S., Marsman, M., Marzari, N., Nitzsche, U., Nordström, L., Ozaki, T., Paulatto, L., Pickard, C. J., Poelmans, W., Probert, M. I. J., Refson, K., Richter, M., Rignanese, G.-M., Saha, S., Scheffler, M., Schlipf, M., Schwarz, K., Sharma, S., Tavazza, F., Thunström, P., Tkatchenko, A., Torrent, M., Vanderbilt, D., van Setten, M. J., Speybroeck, V. V., Wills, J. M., Yates, J. R., Zhang, G.-X. & Cottenier, S. (2016). *Science*, **351**, aad3000.
- 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.* **B56**, 697–714.
- Macrae, C. F., Sovago, I., Cottrell, S. J., Galek, P. T. A., McCabe, P., Pidcock, E., Platings, M., Shields, G. P., Stevens, J. S., Towler, M. & Wood, P. A. (2020). *J. Appl. Cryst.* **53**, 226–235.
- Marom, N., DiStasio, R. A. Jr, Atalla, V., Levchenko, S., Reilly, A. M., Chelikowsky, J. R., Leiserowitz, L. & Tkatchenko, A. (2013). *Angew. Chem. Int. Ed.* **52**, 6629–6632.
- Mayo, R. A. & Johnson, E. R. (2021). *CrystEngComm*, **23**, 7118–7131.
- Mayo, R. A., Otero-de-la-Roza, A. & Johnson, E. R. (2022). *CrystEngComm*, **24**, 8326–8338.
- Mortazavi, M., Hoja, J., Aerts, L., Quéré, L., van de Streek, J., Neumann, M. A. & Tkatchenko, A. (2019). *Commun. Chem.* **2**, 1–7.
- Motherwell, S. & Chisholm, J. A. (2005). *J. Appl. Cryst.* **38**, 228–231.
- 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.* **B58**, 647–661.
- Neumann, M. A., Leusen, F. J. J. & Kendrick, J. (2008). *Angew. Chem. Int. Ed.* **47**, 2427–2430.
- Neumann, M. A. & Perrin, M.-A. (2005). *J. Phys. Chem. B*, **109**, 15531–15541.
- Neumann, M. A. & van de Streek, J. (2018). *Faraday Discuss.* **211**, 441–458.
- Nyman, J. & Day, G. M. (2015). *CrystEngComm*, **17**, 5154–5165.
- Nyman, J. & Day, G. M. (2016). *Phys. Chem. Chem. Phys.* **18**, 31132–31143.
- Otero-de-la-Roza, A. & Johnson, E. R. (2012). *J. Chem. Phys.* **137**, 054103.
- Otero-de-la-Roza, A., Johnson, E. R. & Luaña, V. (2014). *Comput. Phys. Commun.* **185**, 1007–1018.
- Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865–3868.
- Price, A. J. A., Mayo, R. A., Otero-de-la-Roza, A. & Johnson, E. R. (2023a). *CrystEngComm*, **25**, 953–960.
- Price, A. J. A., Otero-de-la-Roza, A. & Johnson, E. R. (2023b). *Chem. Sci.* **14**, 1252–1262.
- Price, S. L. (2013). *Acta Cryst.* **B69**, 313–328.
- Price, S. L. (2018). *Faraday Discuss.* **211**, 9–30.
- 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, R. A. Jr, 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., 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., Vazquez-Mayagoitia, A., Vogt, L., Wang, Y., Watson, R. E., de Wijs, G. A., Yang, J., Zhu, Q. & Groom, C. R. (2016). *Acta Cryst.* **B72**, 439–459.
- Reilly, A. M. & Tkatchenko, A. (2013). *J. Chem. Phys.* **139**, 024705.
- Reilly, A. M. & Tkatchenko, A. (2014). *Phys. Rev. Lett.* **113**, 055701.
- Shtukenberg, A. G., Zhu, Q., Carter, D. J., Vogt, L., Hoja, J., Schneider, E., Song, H., Pokroy, B., Polishchuk, I., Tkatchenko, A., Oganov, A. R., Rohl, A. L., Tuckerman, M. E. & Kahr, B. (2017). *Chem. Sci.* **8**, 4926–4940.
- Tkatchenko, A., DiStasio, R. A., Car, R. & Scheffler, M. (2012). *Phys. Rev. Lett.* **108**, 236402.
- Togo, A. & Tanaka, I. (2015). *Scr. Mater.* **108**, 1–5.
- Ullmann, J. R. (1976). *JACM*, **23**, 31–42.
- Weatherby, J. A., Rumson, A. F., Price, A. J. A., Otero de la Roza, A. & Johnson, E. R. (2022). *J. Chem. Phys.* **156**, 114108.
- Whittleton, S. R., Otero-de-la-Roza, A. & Johnson, E. R. (2017a). *J. Chem. Theory Comput.* **13**, 441–450.
- Whittleton, S. R., Otero-de-la-Roza, A. & Johnson, E. R. (2017b). *J. Chem. Theory Comput.* **13**, 5332–5342.
- Yang, S. & Day, G. M. (2021). *J. Chem. Theory Comput.* **17**, 1988–1999.