

Cite this: DOI: 00.0000/xxxxxxxxxx

## Accurate and Efficient Polymorph Energy Ranking with XDM-Corrected Hybrid DFT †

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

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Accurate and efficient computation of relative energies of molecular crystal polymorphs is of central importance for solid-state pharmaceuticals and in other technologically relevant fields. In recent years, dispersion-corrected density-functional theory (DFT) has emerged as the pre-eminent energy ranking method for crystal structure prediction (CSP). However, planewave implementations of these methods are hindered by poor scaling for large unit cells and are limited to semi-local functionals that suffer from delocalisation error. In this work, we demonstrate that a recent implementation of the exchange-hole dipole moment (XDM) dispersion correction in the Fritz Haber Institute *ab initio* materials simulation (FHI-aims) package provides excellent performance for the energy ranking step of CSP. Thanks to its use of highly optimized numerical atom-centred orbitals, FHI-aims provides effectively linear scaling with system size and allows efficient use of hybrid density functionals with minimal basis-set incompleteness errors. We assess the performance of this methodology for the 26 compounds that formed the first 6 CSP blind tests. The hybrid results show significant improvements for 4/26 compounds, where delocalisation error affects the quality of predicted crystal energy landscapes.

## 1 Introduction

Molecular crystals are of central importance as pharmaceuticals,<sup>1–3</sup> energetic materials,<sup>4,5</sup> and in the emerging field of organic electronics.<sup>6,7</sup> Due to the sensitivity of solid-state properties such as solubility and charge transport on crystal packing, it is important to identify all likely polymorphs when developing compounds for these applications.<sup>8–13</sup> The problem of theoretical identification of isolable polymorphs is termed first-principles crystal structure prediction (CSP).

Periodically, the Cambridge Crystallographic Data Centre (CCDC) organises blind tests of CSP methods in which crystal structures of small sets of compounds are determined by X-ray diffraction, but are not released to the community until researchers have attempted to predict the structure(s) of the isolated polymorph(s).<sup>14–19</sup> There are two challenges at the core of CSP – that of exhaustive structure generation and that of accurate energy ranking of the resulting structure candidates. Ideally, the experimentally observed polymorph(s) should be among the most energetically stable of the putative crystal structures.

The focus of the present work is on the energy ranking step of CSP, where computational methods that provide well-balanced

descriptions of electrostatics, charge transfer, polarisation, non-bonded repulsion, and London dispersion are required. The advent of dispersion corrections has opened the door for the use of periodic-boundary density-functional theory (DFT) for CSP, as illustrated by Neumann and coworkers<sup>20–23</sup> in the 5th CSP blind test<sup>18</sup> and further demonstrated in many subsequent studies.<sup>13,24–37</sup> As dispersion-corrected DFT outperforms alternative energy-ranking methods for CSP, it is essential to have a DFT-based dispersion method that is as accurate and efficient as possible for molecular crystals.

Many successful applications of DFT to molecular crystals have used planewave basis sets and the projector augmented-wave method.<sup>13,20–25,28,29,38,39</sup> However, such an approach has two key limitations. The first is that planewave methods have unfavourable computational scaling for large unit cells. This drastically limits their applicability to CSP since many polymorphs of fairly complex active pharmaceutical compounds, and other materials of interest, can contain many molecules per unit cell. The second is that planewave calculations are limited to generalised gradient approximation (GGA) methods and do not allow routine use of hybrid density functionals. This is important due to delocalisation error,<sup>40–42</sup> which is prevalent in GGAs. Delocalisation error can affect polymorph ranking for flexible molecules where there is a competition between intramolecular conjugation and intermolecular interactions,<sup>29,43,44</sup> as well as for organic salts<sup>28</sup> and acid-base co-crystal systems.<sup>45</sup> Hybrid functionals typically provide better performance than GGAs for non-covalent interactions in finite molecular systems,<sup>42,46–49</sup> particularly in cases

<sup>a</sup> Department of Chemistry, Dalhousie University, 6274 Coburg Rd, Halifax, Nova Scotia, B3H 4R2, Canada. E-mail: erin.johnson@dal.ca

<sup>b</sup> Departamento de Química Física y Analítica, Facultad de Química, Universidad de Oviedo, 33006 Oviedo, Spain. E-mail: oteroalberto@uniovi.es

† Electronic Supplementary Information (ESI) available: Tables of relative energies. See DOI: 10.1039/cXCP00000x/

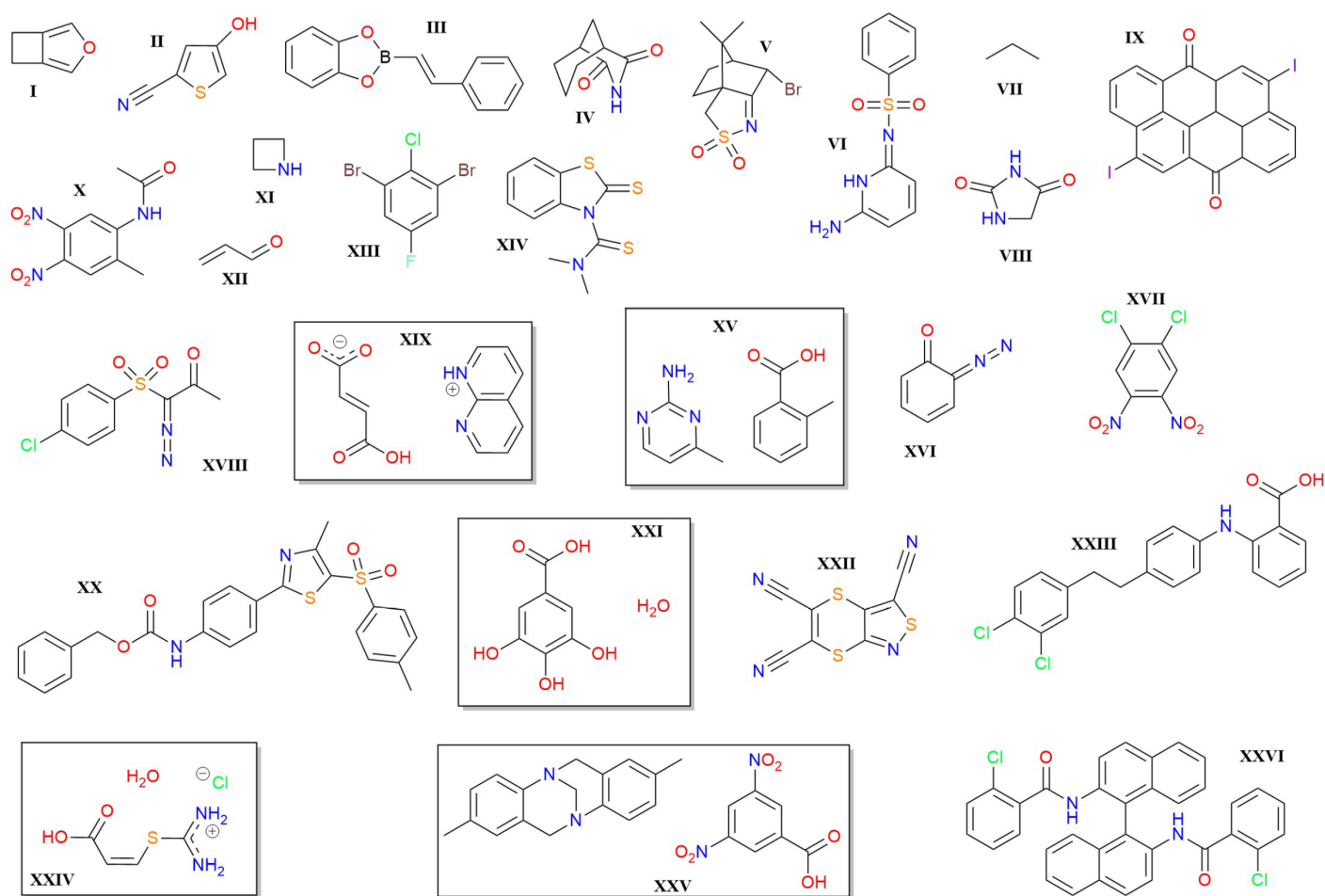


Fig. 1 Structures of the compounds provided by the CCDC for the first 6 CSP blind tests. Rectangles are used for molecule grouping in co-crystals. I: 3-oxabicyclo[3.2.0]hepta-1,4-diene, II: 4-hydroxy-2-thiophenecarbonitrile, III: 2-(2-phenylethenyl)-1,3,2-benzodioxaborole, IV: 3-azabicyclo[3.3.1]nonane-2,4-dione, V: 7-endo-(bromocamphoroylsulfonyl)imine, VI: 6-amino-2-phenylsulfonylimino-1,2-dihydropyridine, VII: propane, VIII: hydantoin, IX: 2,9-diiodoanthanthrone, X: 2-acetamido-4,5,-dinitrotoluene, XI: azetidine, XII: 2-propenal, XIII: 1,3-dibromo-2-chloro-5-fluorobenzene, XIV: N-(dimethylthiocarbamoyl)benzothiazole-2-thione, XV: 2-amino-4-methylpyrimidine:2-methylbenzoic acid, XVI: 2-diazo-3,5-cyclohexadiene-2-one, XVII: 1,2-dichloro-4,5-dinitrobenzene, XVIII: (1-((4-chlorophenyl)sulfonyl)-2-oxopropylidene)diazonium, XIX: 1,8-naphthyridinium fumarate, XX: benzyl-(4-(4-methyl-5-(p-tolyl-sulfonyl)-1,3-thiazol-2-yl)phenyl)carbamate, XXI: gallic acid monohydrate, XXII: tricyano-1,4-dithiaino[c]-isothiazole, XXIII: 2-((4-(3,4-dichlorophenethyl)phenyl)amino)benzoic acid, XXIV: chloride salt hydrate of (Z)-3-((diaminomethyl)thio)acrylic acid, XXV: 2,8-dimethyl-6H,12H-5,11-methanodibenzo[b,f][1,5]diazocine:3,5-dinitrobenzoic acid, XXVI: N,N'-([1,1'-Binaphthalene]-2,2'-diyl)bis(2-chlorobenzamide).

with significant delocalisation error, and we see similar improvement for solids based on lattice energies of small-molecule crystals.<sup>50,51</sup>

To avoid the limitations of planewave basis sets, we turn to numerical atom-centered orbitals (NAOs). NAOs of finite extent are a highly promising alternative since they allow elimination of integrals involving distant atomic centers in DFT calculations, resulting in formal linear scaling for large systems.<sup>52</sup> In particular, the Fritz Haber Institute *ab initio* materials simulation (FHI-aims) package<sup>53–55</sup> is a very robust NAO code that, thanks to its design, has minimal basis-set incompleteness. Implementation in FHI-aims has allowed application of the many-body dispersion (MBD) method,<sup>56,57</sup> paired with hybrid DFT, to molecular crystal benchmarks and CSP studies.<sup>32–37</sup> Moreover, we recently implemented our exchange-hole dipole moment (XDM) dispersion model in FHI-aims and showed that, when paired with selected hybrid functionals, it provides unprecedented accuracy for evalu-

ation of molecular-crystal lattice energies.<sup>51</sup>

In this article we assess the performance of XDM paired with hybrid functionals and the light NAO basis set for the energy ranking step of CSP. Specifically, we consider the 26 compounds that formed the first 6 CSP blind tests, shown in Figure 1.<sup>14–19</sup> Of the submitted putative crystal structures, the experimental polymorph is consistently ranked among the 10 most-stable unique candidates. Use of a hybrid functional with 50% exact-exchange mixing is shown to significantly improve on GGA results for challenging crystal energy landscapes where delocalisation error in GGA functionals is known to adversely impact the ranking.

## 2 Theory and Computational Methods

### 2.1 The XDM Dispersion Correction

XDM<sup>58,59</sup> is a post-self-consistent correction to the energy computed with some base density functional approximation (DFA):

$$E = E_{\text{DFA}} + E_{\text{XDM}}. \quad (1)$$

The XDM dispersion energy itself is expressed as a sum over all pairs of atoms,  $i$  and  $j$ , in the system

$$E_{\text{XDM}} = -\frac{1}{2} \sum_{n=6,8,10} \sum_{ij} \frac{C_{n,ij}}{R_{ij}^n + R_{\text{vdW},ij}^n}, \quad (2)$$

where  $R_{ij}$  is the internuclear distance. In a solid, the sum runs over all surrounding unit cells. Here,  $C_6$ ,  $C_8$ , and  $C_{10}$  are termed the atomic dispersion coefficients, which are functions of the self-consistent electron density, its gradient and Laplacian, the kinetic energy density, and Hirshfeld partitioning weights. As a result of their density and derivative dependence, the dispersion coefficients are highly responsive to changes in the chemical environment of an atom due to charge transfer, coordination, hydrogen bonding, and even weaker van der Waals interactions with distant atoms.<sup>60</sup>

The  $R_{\text{vdW},ij}$  in Eqn. 2 is the sum of effective van der Waals radii of atoms  $i$  and  $j$ , which is defined using an average of three possible ratios of the dispersion coefficients and involves two empirical fit parameters,  $a_1$  and  $a_2$ :

$$R_{\text{vdW},ij} = \frac{a_1}{3} \left[ \left( \frac{C_{8,ij}}{C_{6,ij}} \right)^{\frac{1}{2}} + \left( \frac{C_{10,ij}}{C_{8,ij}} \right)^{\frac{1}{2}} + \left( \frac{C_{10,ij}}{C_{6,ij}} \right)^{\frac{1}{4}} \right] + a_2. \quad (3)$$

The  $a_1$  and  $a_2$  parameters are fit once for each choice of density functional and basis set by minimising the root-mean-square percent error in computed binding energies of 49 molecular dimers.<sup>46,61</sup> After this, the parameters are kept fixed and are transferable to all elements of the periodic table, as well as between finite (molecular) and periodic-boundary (solid-state) calculations.<sup>51</sup>

### 2.2 Exchange-Correlation Functionals

All density-functional dispersion corrections must be paired with a base DFA. In this work, we consider three DFA exchange-correlation functionals of the general form

$$E_{\text{XC}} = (1 - a_{\text{X}})E_{\text{X}}^{\text{B86b}} + a_{\text{X}}E_{\text{X}}^{\text{HF}} + E_{\text{C}}^{\text{PBE}}, \quad (4)$$

which combine the B86b<sup>62</sup> exchange functional with PBE<sup>63</sup> correlation. B86b is our exchange functional of choice for molecular crystals due to its high accuracy for describing non-bonded repulsion.<sup>61,64-67</sup> The parameter  $a_{\text{X}}$  controls the extent of exact-exchange mixing, and we will consider values of 0, 0.25, and 0.50, which correspond to the B86bPBE GGA, and the B86bPBE-25X and B86bPBE-50X hybrid functionals, respectively.<sup>50,51</sup>

### 2.3 Computational Methods

All calculations in this work were carried out with the FHI-aims program (version 210513). The XDM method and the B86bPBE, B86bPBE-25X, and B86bPBE-50X functionals were all implemented in a copy of the code,<sup>51</sup> and these methods are now available in the distribution version of the software. For compounds from the first to fifth blind tests (I to XXI), the starting structures were the equilibrium B86bPBE-XDM structures obtained in our previous studies,<sup>28,29</sup> which in turn were obtained from the Supplementary Information of the corresponding blind test articles, including the reference experimental structures.<sup>14-18</sup> For the sixth blind test compounds (XXII to XXVI), in addition to the top three structures submitted by the various participant groups,<sup>19</sup> we considered the 100 structures submitted by Neumann's group<sup>19</sup> as well as the all structures in the POLY59 set compiled by Brandenburg and Grimme.<sup>25</sup> The numbers of candidate structures considered for each compound is given in the ESI. All atomic positions and lattice vectors in all molecular crystals were fully optimised with B86bPBE-XDM. Subsequent single-point energy calculations were performed with the B86bPBE-25X and B86bPBE-50X hybrids, again paired with XDM. While geometry optimizations are possible with the hybrid functionals, we do not expect this to lead to improved accuracy based on our previous work,<sup>51</sup> so we consider only GGA geometries to reduce the computational expense.

Calculations used the "light" (double- $\zeta$ ) NAO basis sets for computational efficiency, along with the dense integration grids (including Lebedev meshes up to 434 angular grid points) that are used by default in conjunction with the tight settings. The  $\mathbf{k}$ -point grids were chosen to have dimensions  $n_1 \times n_2 \times n_3$  such that

$$n_i = \text{int}[\max(1, R_k |b_i| + 0.5)], \quad (5)$$

where  $|b_i|$  is the magnitude of the corresponding reciprocal lattice vector and the length parameter,  $R_k$ , was set to 50 Bohr. The XDM damping parameters were taken from Ref. 51 and are  $a_1 = 0.8219$ ,  $a_2 = 1.2069 \text{ \AA}$  for B86bPBE,  $a_1 = 0.5235$ ,  $a_2 = 2.1995 \text{ \AA}$  for B86bPBE-25X, and  $a_1 = 0.0831$ ,  $a_2 = 3.7362 \text{ \AA}$  for B86bPBE-50X.

## 3 Results and Discussion

### 3.1 Summary of Polymorph Ranking

Results of the energy ranking of putative crystal structures for the 26 compounds constituting the first 6 CSP blind tests are summarized in Table 1. The B86bPBE-XDM results are similar to those from previous planewave calculations performed with the same methodology,<sup>28,29</sup> confirming the minimal effects of basis-set incompleteness in these NAO calculations. Overall, the GGA and B86bPBE-25X-XDM hybrid identify an experimental polymorph as the minimum-energy structure in 17/26 cases, and this fraction increases to 18/26 cases with the B86bPBE-50X-XDM hybrid functional.

Notably, with the 50% hybrid, an experimental polymorph is ranked second in another 6/26 cases. In 3 of these, the experimental form is nearly degenerate with the minimum-energy candidate, while it lies 1.2 kJ/mol or less above the minimum otherwise. For compound V, the experimental structure is ranked 4th in

Table 1 Ranking of the (most stable) experimental polymorph on crystal energy landscapes computed with the B86bPBE-XDM (GGA), B86bPBE-25X-XDM (25X), and B86bPBE-50X-XDM (50X) functionals. Also shown is the energy ( $\Delta E$ , in kJ/mol) of that experimental polymorph relative to the most stable candidate structure on each landscape.

Compound	GGA		25X		50X	
	Rank	$\Delta E$	Rank	$\Delta E$	Rank	$\Delta E$
I	1 <sup>a</sup>	0.0	1 <sup>b</sup>	0.0	1 <sup>b</sup>	0.0
II	4	1.1	2	0.7	2	0.5
III	1	0.0	1	0.0	1	0.0
IV	1	0.0	1	0.0	1	0.0
V	4	1.9	4	1.9	4	1.5
VI	1	0.0	1	0.0	1	0.0
VII	1	0.0	1	0.0	1	0.0
VIII	1	0.0	1	0.0	1	0.0
IX	1	0.0	1	0.0	1	0.0
X	2	0.9	2	0.4	1	0.0
XI	1	0.0	1	0.0	1	0.0
XII	1	0.0	1	0.0	1	0.0
XIII	1	0.0	1	0.0	1	0.0
XIV	1	0.0	1	0.0	1	0.0
XV	2	0.3	2	0.8	2	1.2
XVI	1	0.0	1	0.0	1	0.0
XVII	1	0.0	1	0.0	1	0.0
XVIII	1	0.0	1	0.0	1	0.0
XIX	6	4.3	4	2.2	2	0.1
XX	4	6.7	4	6.5	6	6.7
XXI	1	0.0	1	0.0	1	0.0
XXII	9	2.7	3	0.4	2	0.1
XXIII	3 <sup>c</sup>	0.8	2 <sup>c</sup>	0.0	2 <sup>c</sup>	0.0
XXIV	8	2.0	3	0.7	2	0.4
XXV	1	0.0	1	0.0	1	0.0
XXVI	1	0.0	1	0.0	1	0.0

<sup>a</sup>polymorph 1, <sup>b</sup>polymorph 2, <sup>c</sup>polymorph c

energy, but still lies within 1.5 kJ/mol of the minimum. These are all sufficiently small energy differences that thermal free-energy contributions (which are generally  $<2$  kJ/mol)<sup>68,69</sup> may be sufficient to reverse the ordering. However, evaluating the thermal corrections requires very computationally expensive phonon calculations, which is beyond the scope of the present work.

### 3.2 Compound XX

All three functionals considered perform consistently poorly for the large, flexible compound XX (benzyl-(4-(4-methyl-5-(*p*-tolylsulfonyl)-1,3-thiazol-2-yl)phenyl)carbamate) from the 5th blind test, the structure of which is shown in Figure 1. Here, the minimum-energy structure is predicted to be more stable than the isolated experimental form by 6.5-6.7 kJ/mol.

One potential source of error for application of DFT to flexible molecules is the conformational energy<sup>43,44</sup> and examination of the structures shows that the DFA-predicted minimum has a different molecular conformation than the experimental form. However, we expect conformational changes to be much better described by hybrid functionals than with GGAs. Additionally, since the conjugated central portion of the molecule must remain planar, the conformational differences primarily involve a 180° degree twist of the amide group. This conformational change

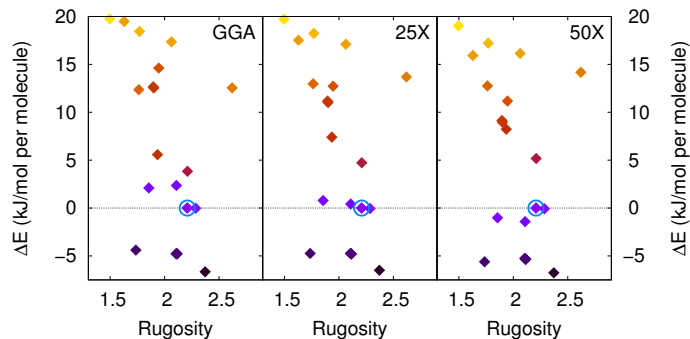


Fig. 2 Computed crystal rugosity-energy landscapes for compound XX, using the rugosity as the ordinate. Results are shown for full geometry relaxations using B86bPBE-XDM (left), as well as single-point energy calculations with B86bPBE-25X-XDM (middle) and B86bPBE-50X-XDM (right). The energy range is truncated to focus on the low-energy candidates and the experimental form is circled.

should be well described by DFT and not responsible for a large energy reordering of the putative crystal structures.

An alternative explanation for the relatively high energy of the experimental structure may be that the crystallisation is controlled by the rugosity.<sup>70</sup> This refers to the roughness of a cleaved crystal surface – crystals with a smooth cleavage plane should be more likely to form than those with a rough surface. We examined the rugosity of the putative crystal structures for compound XX using the CSD-Particle suite of tools within the Mercury program<sup>71</sup> (2022.2.0) and the resulting crystal rugosity-energy landscape is shown in Figure 2. While the minimum-energy crystal structure is found to have a higher rugosity than the experimental form, several other low-energy candidates also have lower rugosities, so this does not reconcile theory and experiment.

A final potential source of error in the DFT ranking is the neglect of thermal free-energy corrections. It is notable that the experimental crystal structure of compound XX was successfully predicted by two participating groups<sup>72</sup> in the 5th CSP blind test,<sup>18</sup> both of whom used the DMACRYs force field<sup>73</sup> with an empirical Lennard-Jones dispersion term. It is possible that the force field partly captures effects of thermal expansion on the crystal structure through its empirical parameterization. The 4.4-6.7 kJ/mol energy differences between the low-energy and experimental forms obtained with DFT are much larger than the contributions from thermal effects seen for small, rigid molecules ( $<2$  kJ/mol).<sup>68,69</sup> However, it would be expected that the magnitude of the thermal free-energy corrections would increase with the number of rotatable bonds, leading to much larger contributions for compound XX (and XXIII), relative to the other members of the blind test set. Indeed, for compound XXIII, Hoja and Tkatchenko<sup>35</sup> noted that inclusion of the vibrational free energy correction could alter the relative stabilities by as much as 8.4 kJ/mol. Thus, one potential reason that compound XX is an outlier for DFT is a much larger contribution from thermal effects than what is seen for rigid molecules.

### 3.3 Effects of Delocalisation Error on Landscapes

To highlight the improved performance of hybrid functionals based on B86bPBE-XDM over the parent GGA for CSP, we focus on four challenging crystal energy landscapes. The first two are for compounds X (2-acetamido-4,5-dinitrotoluene)<sup>16</sup> and XIX (1,8-naphthyridinium fumarate),<sup>18</sup> where the delocalisation error inherent in GGA functionals has been previously shown to adversely affect the predicted landscapes.<sup>28,29,43</sup> This results in one or more candidate structures being predicted to lie lower in energy than the known polymorph, as shown in Figure 3.

The crystal energy landscape of compound X has been extensively investigated by both ourselves<sup>29</sup> and by Beran and coworkers.<sup>43</sup> Here, there are two competing low-energy structures. The experimental form has the acetamide group twisted somewhat out of plane to maximise intermolecular hydrogen bonding. Conversely, GGA functionals favour a competing structure in which the acetamide group lies in plane to maximise conjugation. The energy ordering can be corrected by evaluating the conformational energy change of the isolated molecule using a correlated-wavefunction theory, such as MP2D.<sup>43</sup> Conversely, compound XIX is an organic salt with rigid molecular components. Salts are problematic for GGA functionals, which are known to give large errors for polarization and charge transfer.<sup>42,45,74–77</sup> Here, GGA functionals favour the structure with the most cooperative H-bonding network.<sup>28</sup>

The other two landscapes highlighted in Fig. 3 are for compounds XXII (tricyano-1,4-dithiino[c]-isothiazole) and XXIV (chloride salt hydrate of (Z)-3-((diaminomethyl)thio)acrylic acid) from the 6th CSP blind test.<sup>19</sup> Like compound XIX, compound XXIV is also an organic salt, so we again expect significant delocalisation error. Examination of the minimum-energy GGA structure shows an unusual arrangement in which the H atom of the carboxylic acid moiety points away from the  $sp^2$  oxygen to form a hydrogen bond to the chloride (see Fig. 4). It is likely that the strength of this ionic H-bond is overestimated by the GGA functional. Finally, it is not immediately obvious how delocalisation error would affect the crystal energy landscape of compound XXII. Examination of the minimum-energy GGA and experimental crystal structures reveals that, while the experimental structure has the molecules stacked in dimers, the GGA minimum has all the molecules aligned along the  $c$  axis (see Fig. 5). It is possible that this results in excessive polarisation at the GGA level, overestimating the electrostatic stabilization.

Hybrid functionals, particular those with near 50% exact-exchange mixing, are known to reduce delocalisation error. As such we expect B86bPBE-50-XDM to significantly improve the crystal energy landscapes in cases where this error plays a role. The hybrid results shown in Figure 3 confirm this to be the case. B86bPBE-50-XDM predicts the experimental form of compound X to be lowest in energy. Moreover, it almost entirely corrects the energy ranking for compound XXII and both organic salts, with the most-stable candidate structures now nearly degenerate with the experimental forms.

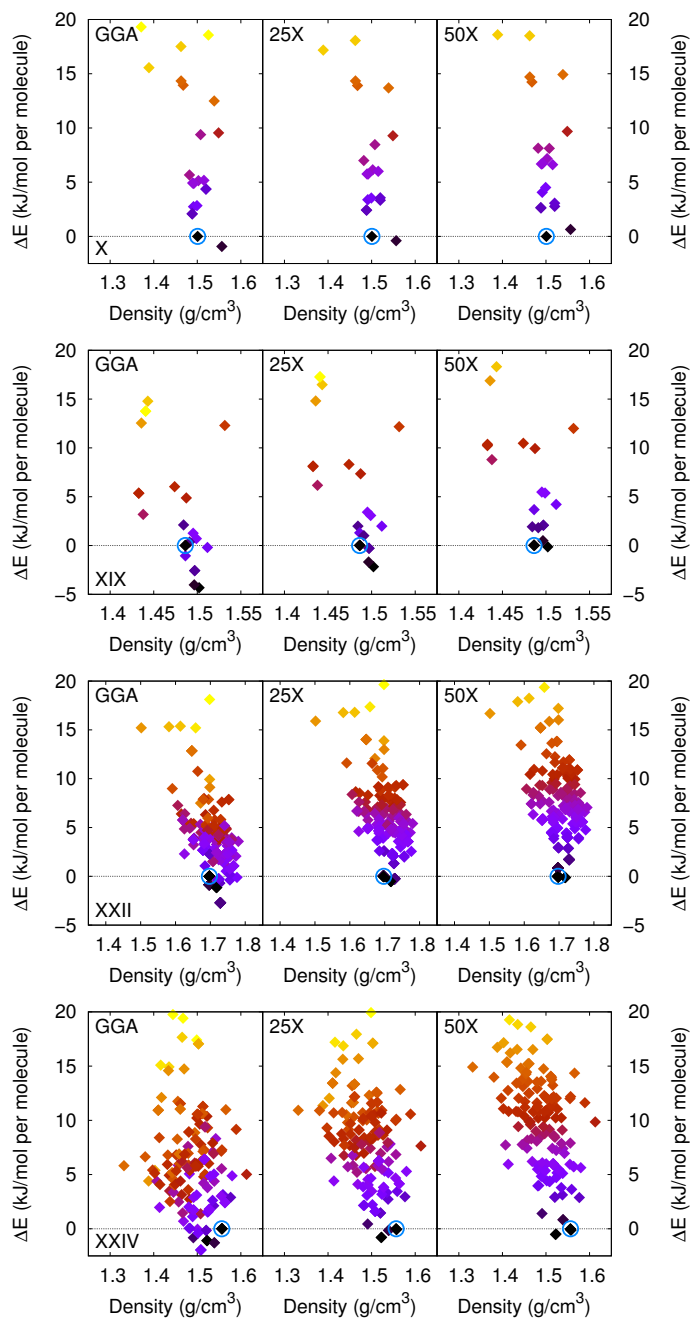


Fig. 3 Computed crystal energy landscapes for compounds X, XIX, XXII, and XXIV (top to bottom). Results are shown for full geometry relaxations using B86bPBE-XDM (left), as well as single-point energy calculations with B86bPBE-25X-XDM (middle) and B86bPBE-50X-XDM (right). The energy ranges are truncated to focus on the low-energy candidates and the experimental forms are circled.

## 4 Summary

In this work, we applied a recent NAO implementation of the XDM dispersion model to rank the submitted structures for the first 6 CSP blind tests. Unlike planewave basis sets, NAOs allow efficient use of hybrid DFT, so the B86bPBE GGA and two related hybrid functionals with 25% and 50% exact exchange mixing were considered. Pairing XDM with the B86bPBE-50X hybrid was found to rank an isolated experimental polymorph

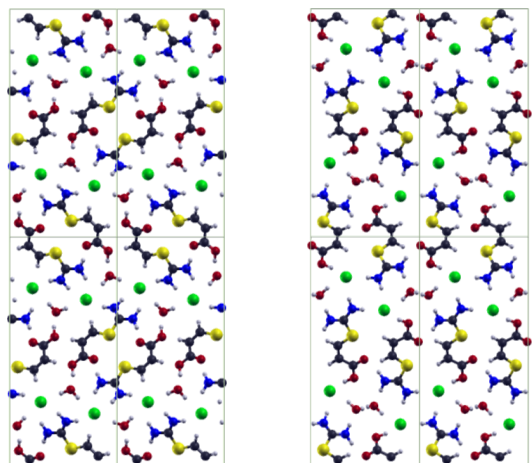


Fig. 4 The experimental (left) and the DFT-predicted minimum-energy (right) crystal structures for compound XXIV, chloride salt hydrate of (Z)-3-((diaminomethyl)thio)acrylic acid, viewed in the  $yz$  plane. C: grey, H: white, N: blue, O: red, S: yellow, Cl: green.

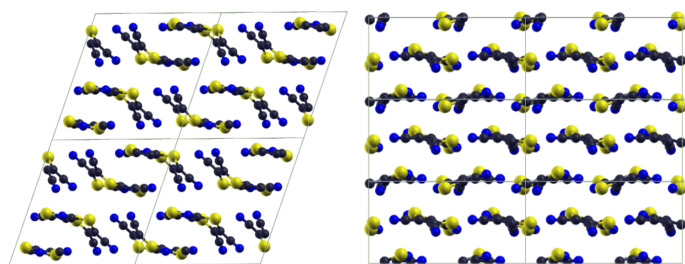


Fig. 5 The experimental (left) and the DFT-predicted minimum-energy (right) crystal structures for compound XXII, tricyano1,4-dithiino[c]-isothiazole, viewed in the  $ac$  plane. C: grey, H: white, N: blue, S: yellow.

as the most stable structure for 18/26 compounds considered, and as the second most-stable structure for another 6/26. In 7 cases, the DFT minimum and the experimental form were separated by 1.5 kJ/mol or less, meaning that thermal free-energy corrections from the phonons may be sufficient to reverse the ranking. The only clear failure of hybrid DFT occurred for compound XX, where the experimental structure lies 6.7 kJ/mol above the DFT minimum, where the larger number of rotatable bonds may magnify the importance of thermal effects. Finally, B86bPBE-50X-XDM was found to provide excellent performance for several challenging crystal energy landscapes where GGA calculations suffer from delocalisation error. This includes both cases where the error affects intramolecular conformational energies, and intermolecular charge transfer, making it a more general approach than monomer energy corrections. Overall, the FHI-aims implementation of XDM-corrected hybrid functionals provides an unprecedented combination of accuracy and efficiency for DFT-based modelling of molecular crystals that should facilitate high-throughput use of first-principles CSP.

## Conflicts of interest

There are no conflicts to declare

## Acknowledgements

This research was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada. AOR thanks: the Spanish Ministerio de Ciencia e Innovación and the Agencia Estatal de Investigación, project PGC2021-125518NB-I00 cofinanced by EU FEDER; the Principality of Asturias (FICYT), project AYUD/2021/51036 cofinanced by EU FEDER; the Spanish MINECO for a Ramón y Cajal fellowship (RyC-2016-20301); and the Spanish MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR for grant TED2021-129457B-I00. The authors are grateful to ACENET, Compute Canada, and the MALTA Consolider supercomputing centre for computational resources.

## Notes and references

- 1 S. L. Price and S. M. Reutzel-Edens, *Drug Discovery Today*, 2016, **21**, 912–923.
- 2 S. L. Price, D. E. Braun and S. M. Reutzel-Edens, *Chem. Commun.*, 2016, **52**, 7065–7077.
- 3 J. Nyman and S. M. Reutzel-Edens, *Faraday Discuss.*, 2018, **211**, 459–476.
- 4 O. Bolton and A. J. Matzger, *Angew. Chem.*, 2011, **50**, 8960–8963.
- 5 I. Bier, D. O'Connor, Y.-T. Hsieh, W. Wen, A. M. Hiszpanski, T. Y.-J. Han and N. Marom, *CrystEngComm*, 2021, **23**, 6023–6038.
- 6 A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper and G. M. Day, *Nature*, 2017, **543**, 657–664.
- 7 J. Yang, S. De, J. E. Campbell, S. Li, M. Ceriotti and G. M. Day, *Chem. Mater.*, 2018, **30**, 4361–4371.
- 8 J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter and J. Morris, *Pharm. Res.*, 2001, **18**, 859–866.
- 9 N. Blagden, M. de Matas, P. T. Gavan and P. York, *Adv. Drug Deliv. Rev.*, 2007, **59**, 617–630.
- 10 D. Singhai and W. Curatolo, *Adv. Drug Deliv. Rev.*, 2004, **56**, 335–347.
- 11 D. Bučar, R. W. Lancaster and J. Bernstein, *Angew. Chem. Int. Ed.*, 2015, **54**, 6972–6993.
- 12 D. I. Millar, I. D. H. Oswald, D. J. Francis, W. G. Marshall, C. R. Pulham and C. A. S, *Chem. Commun.*, 2009, 562–564.
- 13 Y. Yang, B. Rice, X. Shi, J. R. Brandt, R. Correa da Costa, G. J. Hedley, D.-M. Smilgies, J. M. Frost, I. D. W. Samuel, A. Otero-de-la-Roza, E. R. Johnson, K. E. Jelfs, J. Nelson, A. J. Campbell and M. J. Fuchter, *ACS Nano*, 2017, **11**, 8329–8338.
- 14 J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti and et al., *Acta Cryst.*, 2000, **B58**, 647–661.
- 15 W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk and et al., *Acta Cryst.*, 2002, **B58**, 647–661.
- 16 G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerigter, R. G. Della Valle and et al., *Acta Cryst.*, 2005, **B61**,

- 511–527.
- 17 G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon and et al., *Acta Cryst.*, 2009, **B65**, 107–125.
  - 18 D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter and et al., *Acta Cryst.*, 2011, **B67**, 535–551.
  - 19 A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylisma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn et al., *Acta Cryst. B*, 2016, **72**, 439–459.
  - 20 M. A. Neumann and M.-A. Perrin, *J. Phys. Chem. B*, 2005, **109**, 15531–15541.
  - 21 M. A. Neumann, F. J. J. Leusen and J. Kendrick, *Angew. Chem. Int. Ed.*, 2008, **47**, 2427–2430.
  - 22 A. Asmadi, M. A. Neumann, J. Kendrick, P. Girard, M.-A. Perrin and F. J. J. Leusen, *J. Phys. Chem. B*, 2009, **113**, 16303–16313.
  - 23 J. Kendrick, F. J. J. Leusen, M. A. Neumann and J. van de Streek, *Chem. Eur. J.*, 2011, **17**, 10736–10744.
  - 24 C. W. Lehmann, *Angew. Chem. Int. Ed.*, 2011, **50**, 5616–5617.
  - 25 J. G. Brandenburg and S. Grimme, *Acta Crystallogr. Sect. B: Struct. Sci.*, 2016, **72**, 502–513.
  - 26 J. G. Brandenburg and S. Grimme, *Top. Curr. Chem.*, 2014, **345**, 1–24.
  - 27 S. L. Price and J. G. Brandenburg, *Non-Covalent Interactions in Quantum Chemistry and Physics*, Elsevier, 2017, ch. 11, pp. 333–363.
  - 28 S. R. Whittleton, A. Otero-de-la-Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 441–450.
  - 29 S. R. Whittleton, A. Otero-de-la-Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 5332–5342.
  - 30 L. M. LeBlanc, and E. R. Johnson, *CrystEngComm.*, 2019, **21**, 5995–6009.
  - 31 L. Kronik and A. Tkatchenko, *Acc. Chem. Res.*, 2014, **47**, 3208–3216.
  - 32 N. Marom, R. A. DiStasio, Jr., V. Atalla, S. Levchenko, A. M. Reilly, J. R. Chelikowsky, L. Leiserowitz and A. Tkatchenko, *Angew. Chem. Int. Ed.*, 2013, **52**, 6629–6632.
  - 33 J. Hoja, A. M. Reilly and A. Tkatchenko, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2017, **7**, e1294.
  - 34 A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, J. Hoja, E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman and B. Kahr, *Chem. Sci.*, 2017, **8**, 4926–4940.
  - 35 J. Hoja and A. Tkatchenko, *Faraday Discuss.*, 2018, **211**, 253–274.
  - 36 J. Hoja, H. Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio and A. Tkatchenko, *Sci. Adv.*, 2019, **5**.
  - 37 M. Mortazavi, J. Hoja, L. Aerts, L. Quéré, J. van de Streek, M. A. Neumann and A. Tkatchenko, *Commun. Chem.*, 2019, **2**, 1–7.
  - 38 B. Rice, L. M. LeBlanc, A. Otero-de-la-Roza, M. J. Fuchter, E. R. Johnson, J. Nelson and K. E. Jelfs, *Nanoscale*, 2018, **10**, 1865–1876.
  - 39 A. Otero-de-la Roza, B. H. Cao, I. K. Price, J. E. Hein and E. R. Johnson, *Angew. Chem. Int. Ed.*, 2014, **53**, 7879–7882.
  - 40 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Science*, 2008, **321**, 792.
  - 41 M.-C. Kim, E. Sim and K. Burke, *Phys. Rev. Lett.*, 2013, **111**, 073003.
  - 42 K. R. Bryenton, A. A. Adeleke, S. G. Dale and E. R. Johnson, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2022, e1631.
  - 43 C. Greenwell and G. J. O. Beran, *Cryst. Growth Des.*, 2020, **20**, 4875–2881.
  - 44 C. Greenwell, J. L. McKinley, P. Zhang, Q. Zeng, G. Sun, B. Li, S. Wen and G. J. O. Beran, *Chem. Sci.*, 2020, **11**, 2200–2214.
  - 45 L. M. LeBlanc, S. G. Dale, C. R. Taylor, A. D. Becke, G. M. Day and E. R. Johnson, *Angew. Chem. Int. Ed.*, 2018, **57**, 14906–14910.
  - 46 A. Otero-de-la-Roza and E. R. Johnson, *J. Chem. Phys.*, 2013, **138**, 204109.
  - 47 L. Goerigk, H. Kruse and S. Grimme, *ChemPhysChem*, 2011, **12**, 3421–3433.
  - 48 K. Kříž and J. Řezáč, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14794–14804.
  - 49 J. Řezáč, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14780–14793.
  - 50 A. Otero-de-la-Roza, L. M. LeBlanc and E. R. Johnson, *J. Chem. Theory Comput.*, 2019, **15**, 4933–4944.
  - 51 A. J. A. Price, A. Otero de la Roza and E. R. Johnson, *Chem. Sci.*, 2023.
  - 52 V. Havu, V. Blum, P. Havu and M. Scheffler, *J. Computat. Phys.*, 2009, **228**, 8367–8379.
  - 53 V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, *Comp. Phys. Comm.*, 2009, **180**, 2175–2196.
  - 54 S. Levchenko, X. Ren, J. Wieferink, R. Johanni, P. Rinke, V. Blum and M. Scheffler, *Compu. Phys. Comm.*, 2015, **192**, 60–69.
  - 55 F. Knuth, C. Carbogno, V. Atalla, V. Blum and M. Scheffler, *Comp. Phys. Comm.*, 2015, **190**, 33–50.
  - 56 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.
  - 57 A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr. and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.
  - 58 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2007, **127**, 154108.
  - 59 E. R. Johnson, *Non-covalent Interactions in Quantum Chemistry and Physics*, Elsevier, 2017, ch. 5, pp. 169–194.
  - 60 A. Otero-de-la-Roza, L. M. LeBlanc and E. R. Johnson, *Phys. Chem. Chem. Phys.*, 2020, **22**, 8266–8276.
  - 61 A. Otero-de-la-Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **136**, 174109.
  - 62 A. D. Becke, *J. Chem. Phys.*, 1986, **85**, 7184.
  - 63 J. P. Perdew, K. Burke and M. Ernzerhof, *Physical review letters*, 1996, **77**, 3865.
  - 64 D. J. Lacks and R. G. Gordon, *Phys. Rev. A*, 1993, **47**, 4681.
  - 65 Y. Zhang, W. Pan and W. Yang, *J. Chem. Phys.*, 1997, **107**,

- 7921–7925.
- 66 F. O. Kannemann and A. D. Becke, *J. Chem. Theory Comput.*, 2009, **5**, 719–727.
- 67 A. J. A. Price, K. R. Bryenton and E. R. Johnson, *J. Chem. Phys.*, 2021, **154**, 230902.
- 68 J. Nyman and G. M. Day, *Phys. Chem. Chem. Phys.*, 2016, **18**, 31132–31143.
- 69 J. A. Weatherby, A. F. Rumson, A. J. A. Price, A. Otero de la Roza and E. R. Johnson, *J. Chem. Phys.*, 2022, **156**, 114108.
- 70 R. Montis, R. J. Davey, S. E. Wright, G. R. Woollam and A. J. Cruz-Cabeza, *Angew. Chem. Int. Ed.*, 2020, **59**, 20357–20360.
- 71 C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, *J. Appl. Cryst.*, 2020, **53**, 226–235.
- 72 A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman, C. C. Pantelides, S. L. Price, P. T. A. Galek, G. M. Day and A. J. Cruz-Cabeza, *Int. J. Pharm.*, 2011, **418**, 168–178.
- 73 S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis and G. M. Day, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8478–8490.
- 74 E. Ruiz, D. R. Salahub and A. Vela, *J. Chem. Phys.*, 1996, **100**, 12265–12276.
- 75 G. Sini, J. S. Sears and J. L. Bredas, *J. Chem. Theory Comput.*, 2011, **7**, 602–609.
- 76 M.-C. Kim, E. Sim and K. Burke, *J. Chem. Phys.*, 2014, **140**, 18A528.
- 77 A. Otero-de-la-Roza, G. A. DiLabio and E. R. Johnson, *J. Chem. Theory Comput.*, 2016, **12**, 3160–3175.