Contrasting Conformational Behaviors of Molecules XXXI and XXXII in the Seventh Blind Test of Crystal Structure Prediction

Gregory J. O. Beran,* Cameron C. Cook and Pablo A. Unzueta

Department of Chemistry, University of California, Riverside, California 92521 USA. E-mail: gregory.beran@ucr.edu

Abstract

Accurate modeling of conformational energies is key to the crystal structure prediction of conformational polymorphs. Focusing on molecules XXXI and XXXII from the 7th Blind Test of Crystal Structure Prediction, this study employs various electronic structure methods up to the level of domain-local pair natural orbital coupled cluster singles and doubles with perturbative triples (DLPNO-CCSD(T1)) to benchmark the conformational energies and to assess their impact on the crystal energy landscapes. Molecule XXXI proves to be a relatively straightforward case, with the conformational energies from generalized gradient approximation (GGA) functional B86bPBE-XDM changing only modestly when using more advanced density functionals such as PBE0-D4, ωB97M-V, and revDSD-PBEP86-D4, dispersion-corrected second-order Møller-Plesset perturbation theory (SCS-MP2D), or DLPNO-CCSD(T1). In contrast, the conformational energies of molecule XXXII prove difficult to determine reliably, and variations in the computed conformational energies appreciably impact the crystal

energy landscape. Even high-level methods such as revDSD-PBEP86-D4 and SCS-MP2D exhibit significant disagreements with the DLPNO-CCSD(T1) benchmarks, highlighting the difficulty of predicting conformational energies for complex, drug-like molecules. The best-converged predicted crystal energy landscape obtained here for molecule XXXII here disagrees significantly with what has been inferred about the solid-form landscape experimentally. The identified limitations of the calculations are probably insufficient to account for the discrepancies between theory and experiment on molecule XXXII, and further investigation of the experimental solid-form landscape would be valuable. Finally, assessment of several semi-empirical methods finds r^2SCAN-3c to be the most promising, with conformational energy accuracy intermediate between the GGA and hybrid functionals and a low computational cost.

1. Introduction

Over the past few decades, organic molecular crystal structure prediction (CSP) has transformed from a seemingly intractable problem to one where successful prediction examples are common. CSP is increasingly being used to understand pharmaceutical solid form landscapes, to help solve challenging crystal structures, and to help design new organic materials (Beran, 2023). The seven Blind Tests of Crystal Structure Prediction between 1999–2022 have played a central role in highlighting the current capabilities of the models being used and the future improvements needed (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005; Day *et al.*, 2009; Bardwell *et al.*, 2011; Reilly *et al.*, 2016; Hunnisett *et al.*, 2024).

Results from the blind tests and other studies demonstrate that reliable CSP for rigid organic molecules is quickly becoming routine (Whittleton *et al.*, 2017*a*; Price *et al.*, 2023; Nikhar & Szalewicz, 2022). While crystal structure predictions for flexible molecules are frequently successful as well (Neumann *et al.*, 2015; Whittleton *et al.*,

2017*b*; Braun *et al.*, 2017; Mortazavi *et al.*, 2019; Hoja *et al.*, 2019; Braun *et al.*, 2019; Bhardwaj *et al.*, 2019; Firaha *et al.*, 2023), the prediction of conformational polymorphs remains more challenging than rigid-molecule CSP for multiple reasons. First, having numerous degrees of intramolecular conformational freedom dramatically increases the search space of potential crystal structures (Oganov, 2018). Second, the selection of conformations to consider for solid-state packing is complicated by the fact that gas-phase conformational energies can be a poor proxy for crystalline stability (Thompson & Day, 2014). For example, whereas an "extended" conformation of a molecule might be considerably less stable than a "folded" one in the gas phase, intermolecular interactions in the solid state may preferentially stabilize the extended conformation. Thus, a CSP search may need to consider alternate criteria beyond gas-phase conformational energies when generating crystal structures.

Third, capturing the interplay of intra- and intermolecular interactions properly in the crystal structure optimization and energy ranking has been a long-standing challenge in CSP. Many years ago, it was recognized that typical classical mechanical force fields frequently lack the requisite accuracy for intramolecular conformational energies (Karamertzanis *et al.*, 2008; Price, 2008). This motivated the development of models which combined quantum mechanical descriptions of the intramolecular energy (either computed directly or via custom-fitted potentials) with force field intermolecular energies (Bowskill *et al.*, 2021).

The widespread successes of van der Waals-inclusive density functional theory (DFT) methods in the 4th Blind Test (Neumann *et al.*, 2008) and many subsequent studies have shifted the many researchers in the field away from such hybrid intra-/intermolecular approaches, at least in the final stage(s) of CSP structure refinement and ranking. However, despite many successful structure predictions, increasing numbers of examples have been found in the past several years where widely-used generalized-gradient

approximation (GGA) and hybrid density functionals (particularly those with ∼20–25% exact exchange) perform surprisingly poorly. For some acid-base co-crystals, GGA functionals spuriously transfer protons and incorrectly predict a salt form to be more stable than the neutral co-crystal (LeBlanc *et al.*, 2018). In the molecule ROY, named for its red, orange, and yellow polymorphs, a variety of DFT methods erroneously predict form Y to be among the least stable polymorphs, when it is actually the most stable experimentally (Tan *et al.*, 2018; Nyman *et al.*, 2019; Greenwell *et al.*, 2020).

The common origin of these poor predictions lies in density-driven delocalization error in the approximate density functionals. In the acid-base co-crystals, GGAs over-stabilize the delocalized charges of the salt forms. In ROY, delocalization error causes GGAs and many hybrids to over-stabilize more planar molecular conformations with extended π-conjugation, such as those found in the red and orange polymorphs, over the nearly perpendicular conformations with more localized electron densities that occur in the yellow polymorphs (Thomas & Spackman, 2018; Nyman *et al.*, 2019; Beran *et al.*, 2022*a*). Beyond ROY, conformational energy errors arising from delocalization error affect crystal polymorph rankings in pharmaceuticals, organic semi-conductor materials, molecule X from the 3rd Blind Test, and other small molecules (Whittleton *et al.*, 2017*b*; Greenwell & Beran, 2020; Greenwell & Beran, 2021). They also impact the chemical reaction energies in photochromic materials (Beran, 2019; Gately *et al.*, 2021; Gately *et al.*, 2022; Cook *et al.*, 2022; Cook *et al.*, 2023; Perry & Beran, 2023) and the lattice energies in halogen bonded crystals (Otero-de-la Roza *et al.*, 2019).

Delocalization error issues can potentially be addressed using hybrid functionals that include a larger fraction (e.g. ∼50%) of exact exchange (Price *et al.*, 2022; Price *et al.*, 2023), though care is needed to ensure the density functional remains well-balanced overall despite the larger exact exchange contribution. For example, BHLYP-D3(BJ),

which includes 50% exchange, performs notably worse than B3LYP-D3(BJ) for hydrogen bonding interactions (Řezáč, 2020; Řezáč, 2020). For the cases where the delocalization error is primarily intramolecular in nature, another strategy returns to the idea of decoupling the intra- and intermolecular interactions by combining crystalline periodic DFT lattice energies with a molecular conformational energy correction that is evaluated for the isolated molecule(s) taken from the crystal (Greenwell & Beran, 2020). Because gas-phase calculations on an isolated molecule are relatively inexpensive, the intramolecular correction can be evaluated with more advanced/robust electronic structure models such as the range-separated hybrid functional ωB97M-V (Mardirossian & Head-Gordon, 2016), double-hybrid functionals such as revDSD-PBEP86-D4 (Santra *et al.*, 2019), spin-component-scaled dispersion-corrected second-order Møller-Plesset perturbation theory (SCS-MP2D) (Greenwell *et al.*, 2022), or domain-local pair-natural orbital coupled cluster singles, doubles, and perturbative triples (DLPNO-CCSD(T)) (Riplinger *et al.*, 2013; Riplinger *et al.*, 2016).

When applied to the crystal structure prediction for ROY, for example, the combination of dispersion-corrected periodic DFT lattice energies (B86bPBE-XDM) and intramolecular SCS-MP2D corrections produced the first crystal energy landscape that was generally consistent with experimentally-known polymorph stabilities (Beran *et al.*, 2022*a*). The same approach improves the polymorph rankings for systems such as molecule X or the pharmaceuticals axitinib and galunisertib (Greenwell & Beran, 2020), and it produces much more sensible reaction energetics for the solid-state photochromic reactions described above. Several examples where intramolecular corrections have proved useful have been reviewed recently (Beran *et al.*, 2023), and such intramolecular corrections have been incorporated in commercial CSP work flows (Firaha *et al.*, 2023).

Of course, these intramolecular corrections address only one manifestation of delo-

calization error, and intermolecular delocalization error can still be problematic in some systems (such as the acid-base co-crystals mentioned above). Cases have been found where there are significant errors in both the intra- and intermolecular interactions that fortuitously cancel each other somewhat in the periodic DFT treatment (Beran *et al.*, 2022*b*). In those systems, correcting the intramolecular error can have the unintended effect of exposing the intermolecular errors more fully, leading to worse agreement with experiment. On the other hand, relying on the cancellation of intra- and intermolecular errors can also be problematic, since the extent of error cancellation may differ across different polymorphs on the crystal energy landscape.

Given the clear challenges associated with modeling conformational polymorphs, the present study investigates the impact of conformational energies on the polymorph energy landscapes of molecules XXXI and XXXII (Figure 1) from the 7th Blind Test. Molecule XXXII exhibits many flexible dihedrals, including several that might impact the extent of π-conjugation and make its polymorph energy rankings susceptible to density-driven delocalization error with some DFT functionals. In contrast, while molecule XXXI exhibits considerable flexibility, that flexibility does not alter the extent of the π-conjugation, reducing the likely impact of intramolecular delocalization error. Thus, these two systems provide an interesting comparison of how DFT and other methods perform for conformational polymorphs when π-conjugation-driven delocalization error either is or is not prominent.

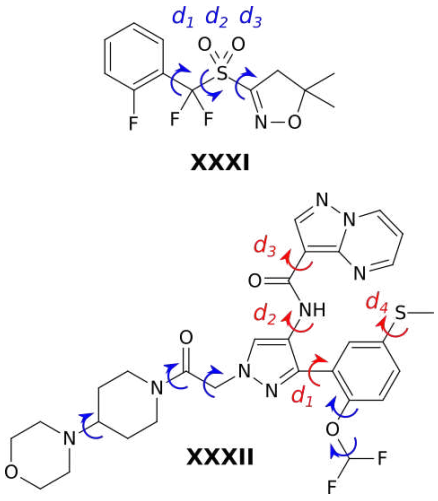

Fig. 1. Molecules XXXI and XXXII from the 7th Blind Test of CSP. Arrows indicate flexible dihedral angles. For Molecule XXXII, rotation about dihedral angles d_1–d_4 in red impacts the extent of π-conjugation.

This study originated with our participation in the structure ranking phase (Phase 2) of the 7th Blind Test, but the investigation here has been extended beyond the results we submitted for the test. In the sections that follow, we investigate the crystal energy landscapes for both species with and without conformational energy corrections, examine the nature of the conformational energy errors, and compare the performance of a variety of semi-empirical, DFT, and correlated wavefunction methods for describing those conformational energies. Notably, while modeling the conformational energies of molecule XXXI proves relatively straightforward, our results highlight the substantial difficulties that remain when trying to study highly flexible species such as molecule XXXII.

2. Computational Methods

During Phase 2 of the 7th Blind Test, the Cambridge Crystallographic Data Centre (CCDC) provided 100 candidate structures of molecule XXXI and 500 candidate structures of molecule XXXII. The methods we used to select, refine, and rank the

crystal structures from these lists are described below. To refer to structures from these sets, we adopt the nomenclature that, for example, #055 refers to structure 55 from original structure list provided for that species by the CCDC.

2.1. Crystal structure selection and geometry optimizations

Molecule XXXI: 99 of the 100 crystal structures provided during Phase 2 of the Blind Test were fully relaxed (both atomic positions and lattice parameters) with periodic DFT. The remaining structure, #89 (Z=18), was omitted for computational expedience, though this structure was later revealed to be experimental form C. The DFT calculations employed the B86bPBE density functional (Becke, 1986; Perdew *et al.*, 1996) and exchange-hole dipole moment (XDM) dispersion correction (Otero-de-la Roza & Johnson, 2012). The projector augmented wave (PAW) approach, a 50 Ry planewave cutoff and Monkhorst-Pack k-point grid spacing of at least 0.06 Å^{-1} were used. The variable-cell QuantumEspresso (Giannozzi *et al.*, 2017) crystal structure optimizations employed energy and geometry convergence criteria of `etot_conv_thr` $= 2 \times 10^{-6}$ a.u. and `forc_conv_thr` $= 6 \times 10^{-4}$ a.u.

Molecule XXXII: Given the large number of crystal structures and the large unit cell sizes for molecule XXXII, a hierarchical refinement was used to select structures for full DFT optimization and ranking. The 500 provided structures were initially optimized with the semi-empirical HF-3c model (Sure & Grimme, 2013) under periodic boundary conditions using Crystal17 (Dovesi *et al.*, 2018). The original solid-state implementation of HF-3c was used (Brandenburg & Grimme, 2014), rather than the rescaled s-HF-3c variant (Cutini *et al.*, 2016). When it became apparent during the course of the Blind Test that those structures differed too strongly from the final DFT ones to provide a reliable preliminary ranking, 481 of the structures were then loosely optimized with periodic DFT, again using the B86bPBE-XDM functional.

These preliminary DFT optimizations employed slightly less dense k-point grids and looser convergence criteria (`etot_conv_thr` = 10^{-4} a.u. and `forc_conv_thr` = 10^{-3} a.u). The remaining 19 structures were omitted due to their larger unit cells and comparatively high HF-3c energies. During the Blind Test, all 33 low-energy structures lying within ∼5 kJ/mol of the global minimum energy (based either on the DFT energies or the SCS-MP2D conformational energy-corrected single-point energies described below) were optimized more tightly with the same settings applied to Molecule XXXI. This 5 kJ/mol cutoff threshold was chosen due to practical time constraints during the test, but it unfortunately excluded the experimental structures.

After the Blind Test results were released, the set of tightly-refined crystal structures was enlarged for the present study to include all 50 most stable structures on the list we had submitted to the Blind Test, which corresponded to all structures within 7.9 kJ/mol of the global minimum (GM) at the level of B86bPBE-XDM with SCS-MP2D conformational energy corrections. This expanded list includes a total of 53 structures from the initial 500, and it includes the experimental forms A_{maj} and B.

2.2. Single-point energies

To address delocalization error or other limitations in the B86bPBE-XDM conformational energies, final single-point energy calculations were performed that correct the intramolecular conformational energies with a higher-level of theory (Greenwell & Beran, 2020),

$$E_{\text{crystal}}^{\text{corrected}} = E_{\text{crystal}}^{\text{DFT}} + \sum_{i}^{Z} \left(E_{\text{mon},i}^{\text{Higher}} - E_{\text{mon},i}^{\text{DFT}} \right) \tag{1}$$

In this expression, the periodic DFT energy of the crystal ($E_{\text{crystal}}^{\text{DFT}}$) is corrected via gas-phase monomer calculations performed with the same DFT model ($E_{\text{mon},i}^{\text{DFT}}$) and at a higher level of theory ($E_{\text{mon},i}^{\text{Higher}}$). The monomer geometries are extracted directly from each DFT-optimized crystal. The sum in Eq 1 runs over all monomers in the

unit cell, though exploitation of space-group symmetry reduces the number of unique monomers that need to be computed to one (if $Z' = 1$) or two (if $Z' = 2$). The "Higher" levels of theory used here include DLPNO-CCSD(T1) (Guo *et al.*, 2018), spin-component-scaled dispersion-corrected second-order Møller-Plesset perturbation theory (SCS-MP2D), and different density functionals. Further details of the gas-phase conformational energy calculations are discussed below.

2.3. Gas-phase conformational energies

Because a few of the crystals in the molecule XXXI and XXXII sets have $Z' = 2$, there are a total of 100 symmetrically unique conformations of molecule XXXI and 55 conformations of molecule XXXII. The B86bPBE-XDM gas-phase conformational energies in Eq 1 were computed in QuantumEspresso, using the same 50 Ry planewave cutoff and PAW potentials as for the crystal calculations. The molecules were placed in a large orthorhombic unit cell that ensures separation between any atoms in the molecule and its periodic images of at least 20 Å in all directions, and only the Γ-point was sampled.

SCS-MP2D conformational energies were obtained by first obtaining MP2 energies that were extrapolated to the complete-basis-set (CBS) limit from aug-cc-pVTZ and aug-cc-pVQZ (Dunning, 1989) results by combining Hartree-Fock (HF)/aug-cc-pVQZ with CBS-limit correlation energies according to (Helgaker *et al.*, 1997),

$$E_{\mathrm{CBS}}^{\mathrm{MP2}} = E_{\mathrm{aQZ}}^{\mathrm{HF}} + E_{\mathrm{CBS}}^{\mathrm{corr}} = E_{\mathrm{aQZ}}^{\mathrm{HF}} + \frac{4^3 E_{\mathrm{aQZ}}^{\mathrm{corr}} - 3^3 E_{\mathrm{aTZ}}^{\mathrm{corr}}}{4^3 - 3^3} \tag{2}$$

These MP2 calculations were performed with PSI4 (Smith *et al.*, 2020). The final SCS-MP2D energies were obtained from the MP2 results using the MP2D library (Greenwell & Beran, 2018).

DLPNO-CCSD(T1) energies were computed in Orca 5 (Neese, 2012) and extrapo-

lated to the CBS-limit via the focal point approach,

$$E_{\mathrm{CBS}}^{\mathrm{DLPNO-CCSD(T1)}} = E_{\mathrm{CBS}}^{\mathrm{MP2}} + E_{\mathrm{aXZ}}^{\mathrm{DLPNO-CCSD(T1)}} - E_{\mathrm{aXZ}}^{\mathrm{MP2}} \tag{3}$$

where "aXZ" refers to a basis set in the aug-cc-pVXZ family (Dunning, 1989). For molecule XXXI, the DLPNO-CCSD(T1) calculations use the aug-cc-pVTZ basis, `TightPNO` settings and `TCutMKN` = 10^{-4}. The molecule XXXI results here differ slightly from those submitted to the Blind Test, where the non-iterative triples variant DLPNO-CCSD(T0) (Riplinger *et al.*, 2013) was used instead of the iterative triples correction in DLPNO-CCSD(T1). The mean absolute deviation in the relative conformational energies between T1 and T0 triples across all molecule XXXI structures is only 0.04 kJ/mol, with a maximum deviation of 0.19 kJ/mol. For further convergence testing, we examined the difference between using aug-cc-pVDZ and aug-cc-pVTZ energies in the CBS-limit extrapolation of the DLPNO-CCSD(T) energies (Eq 3). At the DLPNO-CCSD(T0) level, the resulting conformational energies differed by only a mean absolute average of 0.10 kJ/mol (maximum 0.43 kJ/mol). Similar deviations are found if one omits the diffuse basis functions by using cc-pVTZ instead of aug-cc-pVTZ. Finally, the impact of tightening the `TCutPNO` parameter by an additional factor of 3 (to 3.33×10^{-8}) was also tested at the aug-cc-pVDZ level, and it altered the relative conformational energies only by a mean absolute average of 0.06 kJ/mol (maximum 0.15 kJ/mol). Therefore, the default `TightPNO` setting of `TCutPNO` = 1.0×10^{-7} was used instead for the final aug-cc-pVTZ calculations. Details of the convergence tests can be found in Section S1.3 of the Supporting Information (SI). Taken together, these results suggest that the DLPNO-CCSD(T1) molecule XXXI conformational energies are likely converged to within a few tenths of a kJ/mol or better.

Converging the DLPNO-CCSD(T) calculations for molecule XXXII proves more difficult. First, due to the large molecular size, the aug-cc-pVDZ basis set (1,585 basis functions) was used instead of aug-cc-pVTZ (3,047 basis functions). For molecule

XXXI, the convergence testing above found only modest differences between the two basis sets once extrapolated to the CBS limit. However, the discrepancies might be larger for molecule XXXII, due to its greater complexity and the more variable magnitudes of the intra-molecular non-covalent interactions across the different conformations. For example, the cc-pVDZ and aug-cc-pVDZ conformational energies of molecule XXXII vary by a mean absolute deviation of 0.39 kJ/mol, compared to 0.16 kJ/mol for molecule XXXI with the same basis sets. If one uses molecule XXXI as a guide, this would imply that similarly large differences would be found if one used aug-cc-pVTZ instead of aug-cc-pVDZ. Second, the molecule XXXII relative conformational energies are more sensitive to the numerical thresholds and triples treatment. Convergence testing found that tightening `TCutPNO` from 10^{-7} to 3.33×10^{-8} and employing the iterative T1 triples instead of non-iterative T0 altered the relative conformational energies by mean absolute 0.28 kJ/mol on average. While many of the conformational energies changed only ∼0.1–0.2 kJ/mol, the conformational energies for several highly-folded conformers changed by a much larger 1.0–1.3 kJ/mol (with sizable contributions arising from both tightening `TCutPNO` and using the iterative T1 triples correction). See SI Section 2.4 for details. Accordingly, the DLPNO-CCSD(T1) with the tighter `TCutPNO` setting was used for the final benchmark conformational energies here. Overall, given that the sensitivity of the benchmark XXXII conformational energies to the basis set and DLPNO parameters are a few times larger than for XXXI, we coarsely estimate that the molecule XXXII conformational energy uncertainties are probably at least ∼0.5 kJ/mol, and they may be larger for the highly folded conformations.

The gas-phase conformational energies have also been evaluated with several additional DFT and semi-empirical models: GGA functional PBE-D4 (Perdew *et al.*, 1996; Caldeweyher *et al.*, 2017), hybrid functional PBE0-D4 (Adamo & Barone, 1999),

range-separated hybrid meta-GGA ωB97M-V, and double-hybrid functional revDSD-PBEP86-D4. PBE and PBE0 are included because of their widespread use in crystal structure prediction, while ωB97M-V and revDSD-PBEP86-D4 are representative of the best-performing density functionals from large-scale benchmark studies (Martin & Santra, 2020; Řezáč, 2022). The PBE-D4, PBE0-D4, and ωB97M-V calculations were performed with PSI4, while the remaining methods were computed using Orca. Most of these DFT calculations were performed in the aug-cc-pVQZ basis set. However, the revDSD-PBEP86-D4 calculations for molecule XXXII employed def2-QZVP instead due to frequent issues in converging the self-consistent-field equations with the larger aug-cc-pVQZ basis. Test calculations on molecule XXXI found that the revDSD-PBEP86-D4 root-mean-square errors relative to DLPNO-CCSD(T1) differed by only 0.02 kJ/mol between def2-QZVP and aug-cc-pVQZ, suggesting that the impact of using the smaller basis set for molecule XXXII is also probably small.

Three inexpensive semi-empirical models are also considered: HF-3c, PBEh-3c (Grimme *et al.*, 2015), and r^{2}SCAN-3c (Grimme *et al.*, 2021). These were chosen because of their potential use for intermediate-level refinement/ranking of candidate crystal structures. These calculations were performed with Orca.

Finally, the sections below analyze the errors in the conformational energies. It is important to recognize that the error statistics obtained when comparing two methods can differ (sometimes significantly) depending on the choice of the reference conformer. The lowest-energy conformer is commonly chosen for the reference conformation. For molecule XXXII, however, it was found that a number of models disagree significantly with DLPNO-CCSD(T1) on the stability of this most stable conformer, and using this structure as the reference energy effectively imparts this disagreement into the relative conformational energies of all other conformations. To reduce the biases introduced by selecting any one particular conformation, the gas-phase conformational energies

discussed below are computed relative to the average conformational energy computed at each level of theory. This choice of the reference energy effectively means that the conformational energy errors offer insight into the distribution of conformational energies about the average conformational energy.

2.4. Gas-phase conformational energy scans

To understand the energetics of the molecular conformations found in the crystals better, a series of one-dimensional (1-D) gas-phase conformational scans were performed about selected dihedral angles in each molecule. For each chosen value of the selected dihedral angle, all other degrees of freedom were relaxed using the same B86bPBE-XDM model as described above. Single-point energies with various models were then performed as described in Section 2.3.

For molecule XXXI, a scan was performed about dihedral angle d_2 as defined in Figure 1. For molecule XXXII, 1-D scans were performed about each of d_1–d_4. For computational expedience, the molecule XXXII scans were performed on fragments of XXXII instead of the whole molecule. The fragment used for each scan is shown in the corresponding potential energy curve figure, and any truncated bonds were terminated with hydrogen atoms.

3. Results and Discussion

3.1. Molecule XXXI

Three polymorphs of XXXI have been found experimentally. Forms A and B are related enantiotropically, with form B (#025, $P2_1/c$, $Z = 4$, $Z' = 1$,) being the most stable form at lower temperatures, and form A ($P2_1/c$, $Z = 4$, $Z' = 1$) becoming the thermodynamically preferred polymorph at higher temperatures (a transition occurs at $\sim$55°C). Form A has major and minor disorder components, which are represented

via structures #098 (A_{maj}) and #001 (A_{min}), respectively. Form C (#089, $R\bar{3}$, $Z = 18$, $Z' = 1$) is solvent-templated and contains large void channels. Form C is the least stable form, at least when the pores are unoccupied, and it is omitted here due to its large unit-cell size (Section 2.1).

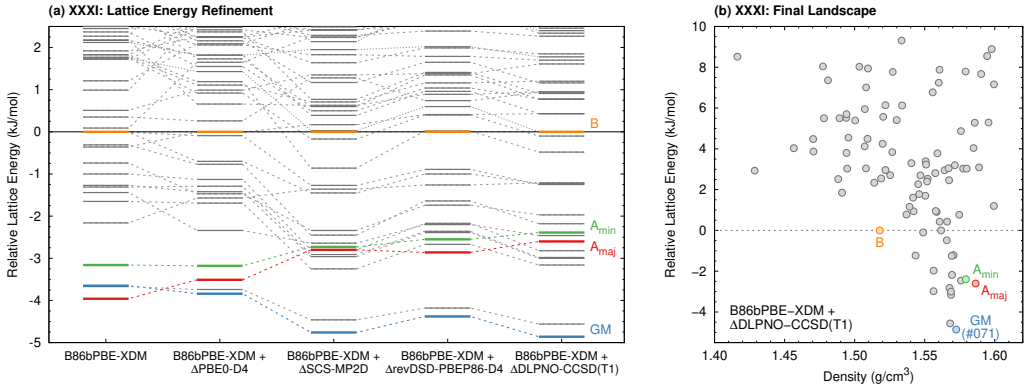

Fig. 2. (a) Impact of applying intramolecular conformational energy corrections to the relative B86bPBE-XDM lattice energies of the candidate crystal structures for molecule XXXI. (b) Final B86bPBE-XDM + ΔDLPNO-CCSD(T1) crystal energy landscape for molecule XXXI. Form B is the most stable form experimentally. Tabulated energetics and additional CSP landscapes can be found in SI Section S1.1

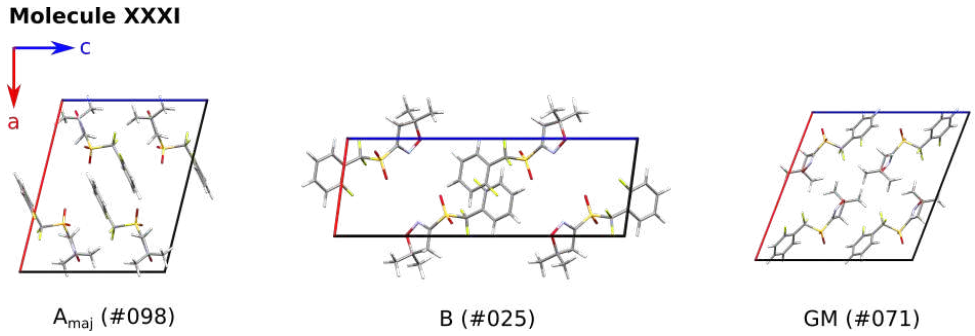

Fig. 3. Crystal structures of the molecule XXXI experimental polymorphs, A_{maj} and B, and the predicted global minimum structure GM.

Figure 2a plots the crystal energy landscape obtained after optimizing the 99 structures with periodic B86bPBE-XDM. This DFT functional predicts the major disorder component of form A_{maj} to be the global minimum (GM). Form B is the most

stable experimentally, but B86bPBE-XDM predicts that it lies at rank 16, about 4 kJ/mol higher than A_{maj}. The minor component A_{min} lies at rank 4, only 0.8 kJ/mol above A_{maj}. Correcting the B86bPBE-XDM conformational energies with DLPNO-CCSD(T1) shifts the relative lattice energies modestly (Figure 2a), and Figure 2b plots the final crystal energy landscape. Candidate structure #071 shifts from rank 3 at the B86bPBE-XDM level to become the new GM after the DLPNO-CCSD(T1) correction. The A_{maj} and A_{min} structures are destabilized and now sit at ranks 7 and 9, or 2.2–2.5 kJ/mol above the GM. Form B now lies 4.8 kJ/mol above the GM at rank 17. Though these conformational energy corrections shift the relative energies, they do not alter the qualitative stability ordering among the three experimental forms. The modest impact of the conformational energy corrections on the relative energies of the GM and experimental structures can be understood from the close similarities of the conformations found in all four crystal structures (Figure 3).

The conformational energy-corrected relative lattice energies still disagree with the experimental room temperature stability ordering by placing A_{maj} and A_{min} below form B. No vibrational free energy calculations are performed here, though results reported by other Blind Test participants suggest that the inclusion of vibrational free contributions can preferentially stabilize form B relative to form A at room temperature.(Hunnisett *et al.*, 2024) Overall, the DLPNO-CCSD(T1) conformational energy corrections are modest across the full set of crystal structures: the average magnitude of the shift is 1.1 kJ/mol relative to Form B, and the largest change is 2.9 kJ/mol.

Next, we examine the role of the conformational energies in the relative lattice energies for the full set of candidate crystal structures more carefully. Because one of the crystal structures has $Z' = 2$, the set of 99 crystal structures contains 100 symmetrically-unique monomer conformations. The gas-phase DLPNO-CCSD(T1)

conformational energies of these structures span a roughly 15 kJ/mol range, with the conformations found in the GM and experimental structures mostly lying near the middle of the range (6.9–8.0 kJ/mol above the lowest-energy conformation). A_{maj} lies moderately higher at 11.1 kJ/mol.

The key conformational flexibility in molecule XXXI involves the central dihedral angle d_2 between the two rings, as defined in Figure 1. Secondary conformational flexibility involving dihedral angles d_1 and d_3 impacts the relative orientations of the two rings, with the fluorine on the benzene ring typically being either *syn*- or *anti*- relative to the isoxazoline heteroatoms. Figure 4 superimposes the distribution of dihedral angles d_2 for the set of monomer conformations onto gas-phase potential energy scans of d_2 computed with B86bPBE-XDM and DLPNO-CCSD(T1). In 78% of the structures, molecule XXXI adopts an extended conformation with d_2 lying in the range ±150–180°. This group includes forms A_{maj}, A_{min}, B, and the GM. Another 16% of the conformations adopt a more "folded" structure, with d_2 dihedral angles of ±40–80°. These two clusters of dihedral angles generally lie near local energy minima of the gas-phase conformational energy profile. The remaining 6% of monomers have intermediate d_2 angles in the ±120–140° range, which correspond to conformations that are energetically-unfavorable in the gas-phase.

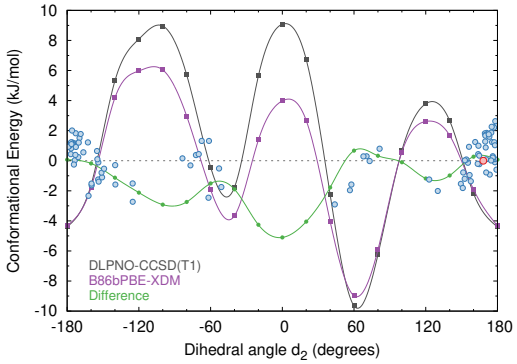

Fig. 4. Gas-phase conformational energy profile for rotation about the central dihedral angle of Molecule XXXI as computed with B86bPBE-XDM (purple) and DLPNO-CCSD(T1) (gray), and the difference between the two (green). For comparison, the figure also plots the energy differences between the DFT and coupled cluster models for each of the conformations found in the set of crystals as a function of the d_2 dihedral angle (blue circles). For consistency with the crystal results in Figure 2, all energies here are plotted relative to the molecular conformation found in the form B crystal (red circle).

None of these three molecule XXXI dihedral angles impacts the extent of π-conjugation in the molecule. Nevertheless, B86bPBE-XDM exhibits errors up to $\sim$5 kJ/mol relative to DLPNO-CCSD(T1) (Figure 4). The most notable discrepancies stem from B86bPBE-XDM underestimating the stability of the conformations with greater overlap of the ring systems. Similar behavior has been observed previously for B86bPBE-XDM and π-π interactions (Beran *et al.*, 2022*b*). Although the crystal conformation energy differences between B86bPBE-XDM and DLPNO-CCSD(T1) are not completely captured by the dihedral angle d_2 descriptor, Figure 4 shows that B86bPBE-XDM errors in the crystal conformations (blue points) largely track with the difference between B86bPBE-XDM and DLPNO-CCSD(T1) computed on the gas-phase scan (green curve). The results for other electronic structure methods on this scan can be found in SI Section S1.3.

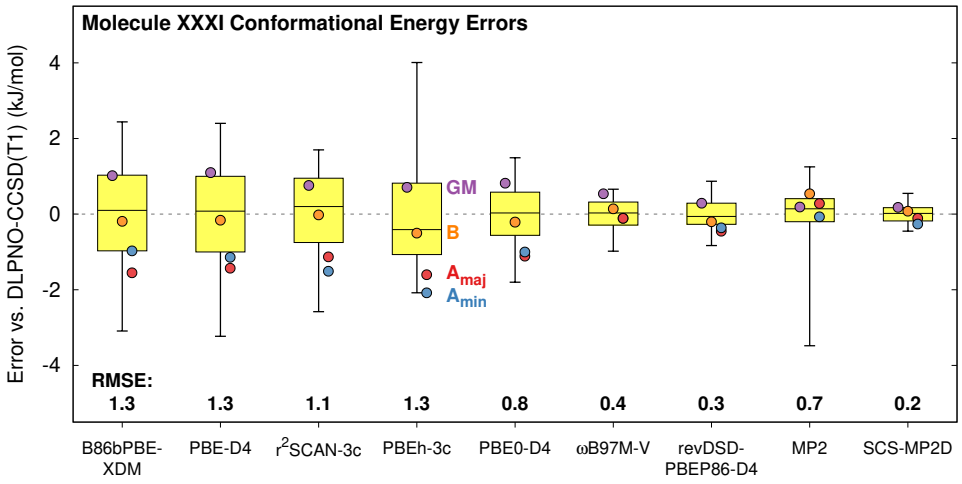

Fig. 5. Gas-phase molecule XXXI conformational energy errors relative to DLPNO-CCSD(T1) for several electronic structure methods, in kJ/mol. The root-mean-square error is also indicated. The conformational energies at each level of theory are defined relative to the average conformational energy in the set.

Although B86bPBE-XDM predicts the XXXI conformational energies reasonably well, it is interesting to investigate what level of theory is required to achieve a more faithful description of the conformational energies for the moderately complex molecule XXXI that does not appear to exhibit strong π-conjugation delocalization error effects. Figure 5 plots the error distributions for several different electronic structure models. Raw conformational energies are tabulated in SI Section S1.2. In these box plots, the center line indicates the median error, the yellow box contains 50% of the errors, and the whiskers show the most extreme errors. Figure 5 shows that both the B86bPBE-XDM and PBE-D4 GGA functionals perform similarly, with root-mean-square (rms) errors of 1.3 kJ/mol. The hybrid functional PBE0-D4 reduces the rms error by 40% to 0.8 kJ/mol. PBE0-D4 makes the largest improvements relative to the GGAs for the most strained conformations ($d_2 \sim 120$–$140°$) and near the $d_2 \sim \pm 60°$ basins. Unsurprisingly, the inclusion of van der Waals dispersion corrections is important for these conformational energies, and omitting the corrections increases the rms

errors ∼2–3-fold (data not shown).

The range-separated hybrid meta-GGA ωB97M-V and the double-hybrid revDSD-PBEP86-D4 functionals perform even better than PBE0-D4, with rms errors of 0.4 kJ/mol and 0.3 kJ/mol, respectively. MP2 performs well for the extended conformations which comprise most of the conformations in this set, but it overestimates the favorable π-π interactions between rings in the more-folded conformations, leading to a larger 0.7 kJ/mol rms error. The SCS-MP2D dispersion correction addresses this problem and reproduces the DLPNO-CCSD(T1) conformational energies with an rms error of only 0.2 kJ/mol.

Examining the largest errors in the Figure 5 conformational energy error distributions, the GGA errors lie within roughly ±3 kJ/mol of DLPNO-CCSD(T1), the global hybrid PBE0-D4 functional clearly improves the accuracy to around ±2 kJ/mol, the range-separated hybrid meta-GGA ωB97M-V and double-hybrid revDSD-PBEP86-D4 obtain errors within ±1 kJ/mol, and SCS-MP2D achieves ±0.6 kJ/mol. The good fidelity among revDSD-PBEP86-D4, SCS-MP2D, and DLPNO-CCSD(T1) is further apparent in Figure 6, which shows the strong correlations (R^2 = 0.94–0.97) between the conformational energy corrections achieved when any of these three models is used to correct B86bPBE-XDM. As a result, a similar crystal energy landscape is obtained regardless of which of these methods (or even PBE0-D4) is used to correct the B86bPBE-XDM lattice energies (Figure 2a). The excellent consistency among the different methods also provides confidence in the accuracy of the benchmark conformational energies.

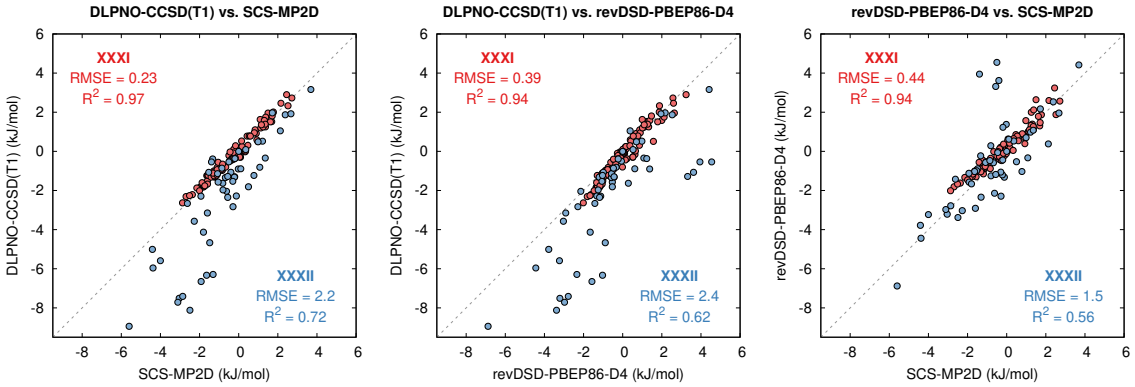

Fig. 6. Correlations between the B86bPBE-XDM conformational energy corrections as computed with SCS-MP2D, revDSD-PBEP86-D4, and DLPNO-CCSD(T1) for molecules XXXI (red) and XXXII (blue). The conformational energy corrections were evaluated relative to form B of both species.

Finally, given the interest in using inexpensive semi-empirical models for intermediate refinement/ranking of structures, we tested three of Grimme's semi-empirical "3c" methods: HF-3c, PBEh-3c, and r^{2}SCAN-3c. HF-3c calculations are very fast but also rather unreliable, with an rms error of 2.8 kJ/mol and maximum errors up to 8 kJ/mol (SI Table S3). PBEh-3c and r^{2}SCAN-3c offer better accuracy, with rms errors of 1.3 kJ/mol (max 4.0 kJ/mol) and 1.1 kJ/mol (max 2.6 kJ/mol), respectively (Figure 5). PBEh-3c is not recommended because it exhibits particularly large errors for certain conformations and is computationally expensive compared to the other "3c" methods tested, but the excellent balance of low-computational cost and good accuracy of r^{2}SCAN-3c is quite promising.

3.2. Molecule XXXII

Molecule XXXII is much larger than XXXI and has many more conformational degrees of freedom. It has a complex solid-form landscape, with at least 8 anhydrate forms, four hydrates, six solvates, and seven other transient or unidentified forms

having been observed. The Blind Test focused on the two anhydrates whose crystal structures have been determined, Forms A and B. The experimental screening report (Hunnisett *et al.*, 2024) provided to the Blind Test organizers indicates that Form B (#232, $P\bar{1}$, $Z = 4$, $Z' = 2$) is believed to be the thermodynamically stable polymorph, while Form A (#317, $P\bar{1}$, $Z = 2$, $Z' = 1$) is thermodynamically metastable, at least at room temperature and above. Form A exhibits disorder in rotation of the the difluoro-methyl group, and structure #317 corresponds to the major component of the disorder, A_{maj}. It should be noted that the experimental investigation of the molecule XXXII solid-form landscape was challenging, and this understanding may be incomplete.(Hunnisett *et al.*, 2024)

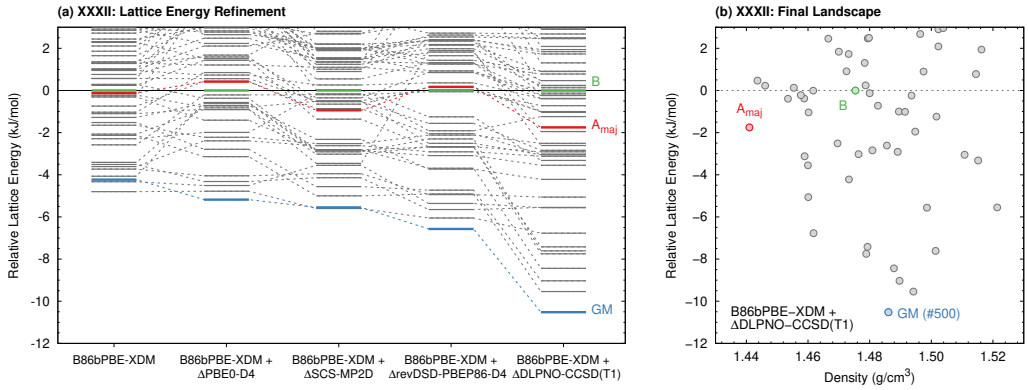

Fig. 7. (a) Impact of applying intramolecular conformational energy corrections to the relative B86bPBE-XDM lattice energies of the candidate crystal structures for molecule XXIII. (b) Final B86bPBE-XDM + ΔDLPNO-CCSD(T1) crystal energy landscape for molecule XXXII. Form B is the most stable form experimentally. Tabulated energetics and additional CSP landscapes can be found in SI Section S2.1

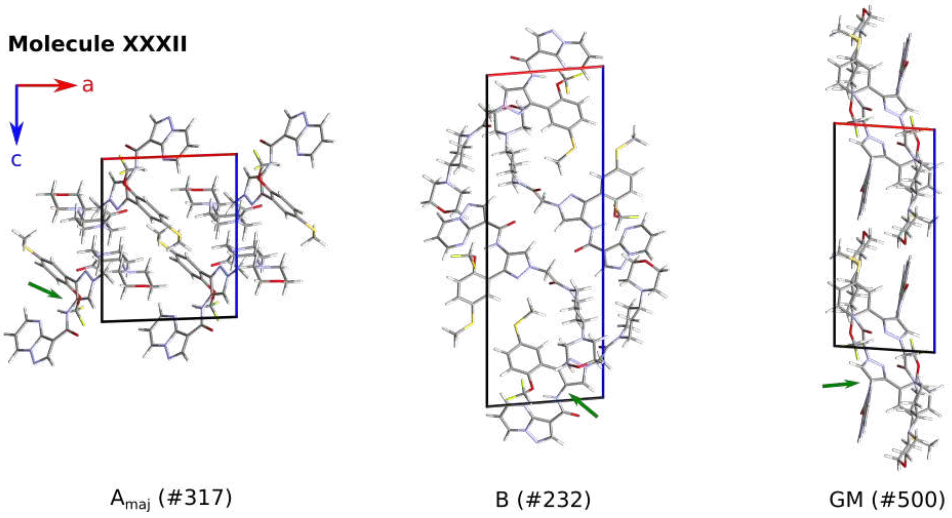

Fig. 8. Crystal structures of the molecule XXXII polymorphs A_{maj} and B, and the predicted GM structure. Green arrows point to key dihedral angle d_2, which is 108° in the GM structure versus nearly planar in the experimental polymorphs.

From the initial set of 500 molecule XXXII crystal structures, 53 low-energy structures were selected for full B86bPBE-XDM geometry refinement as described in Section 2.1. The resulting crystal energy landscapes computed with B86bPBE-XDM and three different conformational energy correction models are plotted in Figure 7, and selected crystal structures are shown in Figure 8. At the B86bPBE-XDM level, forms A_{maj} and B lie 0.1 kJ/mol apart and 4.7–4.8 kJ/mol above the lowest-energy structure (#423). Applying either SCS-MP2D or revDSD-PBEP86-D4 conformational energy corrections moderately alters the B86bPBE-XDM landscape, stabilizing some of the lowest-energy structures relative to A_{maj} and B by a few tenths of a kJ/mol, and making structure #500 (GM) similar to (SCS-MP2D) or lower in energy (revDSD-PBEP86-D4) than #423. However, applying the DLPNO-CCSD(T1) conformational correction alters the landscape considerably. Although forms A_{maj} and B lie within 2 kJ/mol of one another across all four landscapes, the DLPNO-CCSD(T1) conformational correction stabilizes a number of other low-energy structures substantially. As a result, A_{maj} now lies 8.8 kJ/mol above the GM, while the experimentally-preferred

form B lies 10.5 kJ/mol above it. In other words, the DLPNO-CCSD(T1) conformational energy correction has a substantial impact, approximately doubling the energy window that separates the experimental forms from the most stable form on the landscape.

Some of the lower-energy structures on the crystal landscape might conceivably correspond to some of the other 6 uncharacterized anhydrate forms that have been observed experimentally. Nevertheless, it is surprising that the reportedly experimentally-preferred form B would be computed to lie 10.5 kJ/mol above the GM. The vast majority of observed crystal polymorphs (including conformational polymorphs) lie within 10 kJ/mol of one another (Nyman & Day, 2015; Cruz-Cabeza *et al.*, 2015), and the errors in the relative lattice energies computed with quantum chemistry are typical considerably smaller than 10 kJ/mol (Whittleton *et al.*, 2017*a*; Whittleton *et al.*, 2017*b*; Hoja *et al.*, 2019; Greenwell *et al.*, 2020; Beran *et al.*, 2022*a*).

The large differences between the final DLPNO-CCSD(T1)-corrected molecule XXXII landscape and the experimental understanding suggest that there are likely problems with the computational and/or experimental results. Potential issues might include the neglect of finite-temperature contributions that significantly reorder the structures on the landscape, unusually large errors in the quantum chemistry lattice energy calculations, and/or the incomplete experimental understanding of the system. Addressing any gaps in the experimental understanding is beyond the scope of this work. We have not performed free energy calculations on the crystals due to the large unit cell sizes. However, a couple other participating groups in the Blind Test did perform free energy calculations, and forms A_{maj} and B remained at least several kJ/mol above those groups' respective GM structures (Hunnisett *et al.*, 2024).

Here, we investigate one aspect of the accuracy of the lattice energy calculations: the accuracy of the conformational energies. Molecule XXXII exhibits a wide variety of

conformations across the 55 symmetrically-unique molecular structures extracted from the 53 crystal structures. They range from highly-extended to somewhat folded, and span a gas-phase energy window of more than 40 kJ/mol. The most stable conformations adopt highly-folded structures with strong intramolecular interactions. Interestingly, the GM crystal structure contains the least stable conformation in this set, lying 43 kJ/mol above the folded conformation in structure #331 in the gas phase. In contrast, the conformational energies of the experimental polymorphs are approximately average within the set (see SI Section S2.2). In the condensed phase, extended conformations will be less penalized due to the ability to form intermolecular interactions with the surrounding environment. For example, placing the models in a methanol or water polarizable continuum solvent model reduces the energy window spanned by the conformations moderately, but the GM conformation still remains among the least stable conformations and lies ∼25 kJ/mol above the structure #331 conformation or ∼10–20 kJ/mol above the form A_{maj} or B conformations.

Among the many conformational degrees of freedom in molecule XXXII, dihedral angles d_1–d_4 (Figure 1) have the largest potential impact on the extent of π-conjugation and therefore to manifest delocalization error issues. For d_1, all 55 conformers examined adopt a narrow range of angles near 45° (or 135°). The absence of significant variations in the extent of π-conjugation across d_1 among the polymorphs means that any errors in describing the energetics associated with dihedral d_1 should largely cancel in the relative lattice energies. Similarly, the intramolecular N-H$\cdots$N hydrogen bond prevents dihedral angle d_3 from deviating significantly from planarity in this set of structures. However, dihedral angles d_2 and d_4 vary more widely across the different structures and prove to be the most important with regard to π-conjugation-related delocalization error in practice.

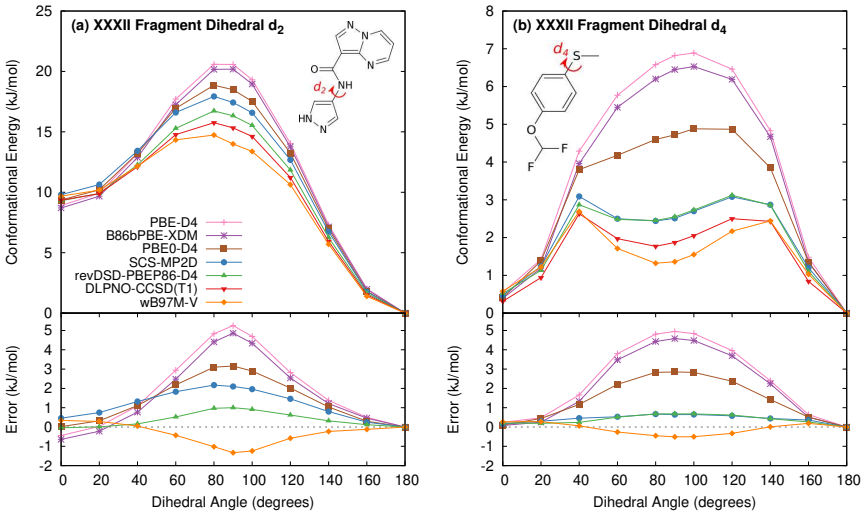

Fig. 9. Potential energy scans for key dihedrals d_2 and d_4 using the molecular fragments of molecule XXXII shown. The upper panel shows the potential energy curve, while the lower one plots the errors relative to the DLPNO-CCSD(T1).

Specifically, whereas most of the molecule XXXII conformations adopt roughly planar d_2 and d_4 angles, 8 structures rotate the amide torsion d_2 out of the plane to ∼100°–110°, and 3 rotate the thioether out of the plane (d_4) to ∼40–65°. Both changes decrease the amount of π-conjugation in the molecule, making these structures potentially problematic for density functionals that exhibit substantial delocalization error. To test this, Figure 9 plots conformational energy scans performed for d_2 and d_4 on the fragments of molecule XXXII shown. Similar scans for dihedral angles d_1 and d_3 are provided in the SI Section S2.3. For both d_2 and d_4, GGA functionals B86bPBE-XDM and PBE-D4 overestimate the torsional barriers by up to ∼4–5 kJ/mol compared to DLPNO-CCSD(T1)—i.e. they over-stabilize the more conjugated planar structures. Problematic delocalization error issues have been observed for both of these functional group types previously (Beran *et al.*, 2022*b*). The hybrid PBE0-D4 functional reduces the error somewhat, while further improvements are obtained with the higher-level methods. Interestingly, whereas most of the models tested here overestimate the

torsion barriers, ωB97M-V underestimates them. SCS-MP2D performs very well for dihedral d_4, but it overestimates the barrier for dihedral d_2 by up to 2 kJ/mol, versus only 1 kJ/mol for revDSD-PBEP86-D4. Note that in contrast to d_4, the orientation of the –$OCHF_2$ group *para*- to the thioether exhibits much smaller B86bPBE-XDM conformational energy errors.

The errors observed for the 11 aforementioned molecule XXXII crystal conformations that deviate from planarity about d_2 or d_4 are consistent with these fragment molecule scans. B86bPBE-XDM and PBE-D4 overestimate the DLPNO-CCSD(T1) conformational energies for the conformations with non-planar d_2 and d_4 by an average 4.6 and 5.2 kJ/mol, respectively, with maximum errors of nearly 7 kJ/mol. Switching to the hybrid PBE0-D4 functional reduces that to 3.3 kJ/mol (max 4.6 kJ/mol), SCS-MP2D to 2.5 kJ/mol (max 4.2 kJ/mol) and revDSD-PBEP86-D4 to 2.2 kJ/mol (max 3.7 kJ/mol). Returning to the crystal polymorph energies in Figure 7, almost all of the large disagreements between the SCS-MP2D and DLPNO-CCSD(T)-corrected lattice energies occur for structures that exhibit non-planar amide bonds d_2. DLPNO-CCSD(T1) stabilizes conformations with non-planar amides moreso than SCS-MP2D or revDSD-PBEP86-D4. As a result, B86bPBE-XDM with DLPNO-CCSD(T1) conformational energy corrections predicts those non-planar amide structures to be the most stable ones on the entire landscape (Figures 7 and 8).

In addition to the conformations that are under-stabilized by B86bPBE-XDM, seven other molecular XXXII conformations are over-stabilized by ∼3–5 kJ/mol relative to DLPNO-CCSD(T1) (3.9 kJ/mol on average, see SI Table S6 for details). No unifying structural trends were identified among these structures. The largest-error case adopts a fairly “folded” conformation, while others are more extended. Errors associated with the conformational energies of the saturated six-membered rings also appear to play a role in some cases. These cases serve as a reminder that the conformational energies

of highly flexible molecules can be difficult to model correctly even without changes in π-conjugation.

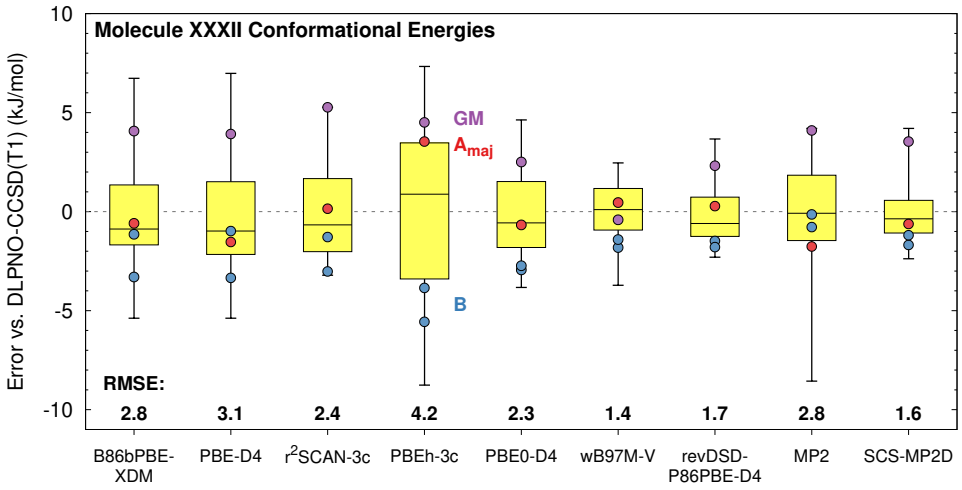

Fig. 10. Gas-phase molecule XXXII conformational energy errors (kJ/mol) vs DLPNO-CCSD(T1) for several electronic structure methods. The conformational energies at each level of theory are defined relative to the average conformational energy in the set

Figure 10 plots the conformational energy error distributions obtained using various electronic structure models. Similar to molecule XXXI, improving the electronic structure model generally reduces the errors relative to DLPNO-CCSD(T1). The 1.4–1.7 kJ/mol rms errors for the three best models, ωB97M-V, SCS-MP2D, and revDSD-P86PBE-D4 are about half as large as those for the GGA functionals (2.8–3.1 kJ/mol). Hybrid PBE0-D4 lies in between the two sets with an rms error of 2.3 kJ/mol. Among the "3c" methods, only r^{2}SCAN-3c gives reasonable accuracy, with an rms error of 2.4 kJ/mol, slightly larger than PBE0-D4. However, it is notable that the widths of the error distributions and the root-mean-square errors for molecule XXXII are substantially larger than for XXXI. For example, SCS-MP2D gives an rms error of only 0.2 kJ/mol for molecule XXXI, compared to 1.6 kJ/mol for XXXII. Moreover, the conformational energy corrections computed with SCS-MP2D, revDSD-PBEP86-D4,

and DLPNO-CCSD(T1) all disagree considerably with one another (Figure 6), with R^2 values of 0.56–0.72. This markedly contrasts the high consistency and excellent R^2 values of 0.94–0.97 found for molecule XXXI.

The large discrepancies between DLPNO-CCSD(T1) and most of the other methods might raise questions about the reliability of the DLPNO-CCSD(T1) benchmarks. As discussed in Section 2.3, however, test calculations suggest the DLPNO-CCSD(T1) relative conformational energies for most conformations appear to be converged to within ∼0.5 kJ/mol with regard to the DLPNO numerical thresholds, basis set, and use of T1 instead of T0 triples. The DLPNO-CCSD(T) calculations that proved most difficult to converge involved the highly folded conformations (especially #331 and #120), but the agreement among SCS-MP2D, revDSD-PBEP86-D4, and DLPNO-CCSD(T1) is reasonable for those conformations (SI Table S6). We further test the DLPNO-CCSD(T1) calculations by comparing them against canonical density-fitted CCSD(T)/aug-cc-pVDZ energy calculations for the small fragment molecule scans about d_2 and d_4. Those tests found that DLPNO-CCSD(T1) differs from full CCSD(T) by 0.4 kJ/mol or less for d_2, and 0.1 kJ/mol or less for d_4. This does not rule out the possibility that the local approximations in DLPNO-CCSD(T1) become more problematic in the full molecule, but the fact that the errors between SCS-MP2D and DLPNO-CCSD(T1) are similar in magnitude for both the small fragments and the full molecules suggests that the small fragments provide a reasonable model system. In other words, the DLPNO-CCSD(T1) energies appear robust across several different potential sources of error.

Due to the large size of molecule XXXII, we have not benchmarked the performance of the B86bPBE-XDM intermolecular interactions against higher-level electronic structure methods. However, we observe that the energy difference between forms A and B and the energies of those forms relative to the GM obtained with our

conformationally-corrected B86bPBE-XDM results appear to be generally consistent with the results from other participating groups in the Blind Test that applied hybrid density functionals to the full crystals (Hunnisett *et al.*, 2024). This provides cause for some optimism that the intermolecular energies associated with molecule XXXII may be easier to model than the conformational energies, and that the crystal lattice energy landscapes here are reasonably well-converged with respect to the quantum chemistry treatment.

Overall, the impact of conformational corrections beyond GGA DFT are much more important for molecule XXXII than for molecule XXXI. Even with the conformationally-corrected lattice energies, however, the predicted polymorph stabilities for molecule XXXII do not agree with the reported experimental stabilities of forms A and B, and the experimental forms lie surprisingly high above the predicted GM on the crystal energy landscape. Computing accurate conformational energies for molecule XXXII proves challenging even with state-of-the-art electronic structure methods. Nevertheless, between our own results and those from other groups participating in the Blind Test, we did not identify any evidence that the errors in the relative lattice energies would be large enough to account for the apparent disagreement between the computed landscape and the reported experimental interpretation. Further experimental work to solve additional crystal forms and to determine the crystal structures of the unknown forms would be very useful to resolving the discrepancies between theory and experiment. In closing, we note the possibility that it may be difficult to crystallize the GM and the other most stable crystal structures identified here, since the need to adopt the highly-unstable non-planar amide conformations could hinder their crystallization kinetics (Bhardwaj *et al.*, 2019; Abramov *et al.*, 2020).

4. Conclusions

The Blind Tests provide an excellent opportunity to assess the quality of state-of-the-art crystal structure prediction techniques. The 7th Blind Test contained several systems whose complexity is comparable to that frequently faced in industrial applications. Focusing on molecules XXXI and XXXII, we investigated the role of conformational energies in the crystal energy landscapes and sought to understand how well widely-used electronic structure methods can capture them. Molecule XXXI proves relatively straightforward: While typical GGA functionals exhibit modest errors in the conformational energies, these errors can be ameliorated readily using hybrid DFT or more advanced electronic structure models. Among the semi-empirical methods, r^2SCAN-3c proves the most promising, with accuracy intermediate between the GGA and hybrid functionals at a low computational cost.

Molecule XXXII proves much more difficult, due to its extremely high conformational flexibility, wide range of conformational energies, and some examples of delocalization error derived from changes in the π-conjugation. Even the best range-separated and double-hybrid density functionals or dispersion-corrected MP2 methods tested here exhibit errors up to several kJ/mol relative to the DLPNO-CCSD(T1) conformational energies. Regardless, the results demonstrate the importance of refining the conformational energies beyond the GGA or even hybrid DFT levels for challenging CSP cases. The present study also demonstrates how consideration of multiple levels of theory in the energy calculations can help assess the errors and uncertainties associated with the predicted crystal energy landscapes.

Interestingly, for both molecules XXXI and XXXII, the relative lattice energies computed here disagree with what has been inferred from experiment. For XXXI, the discrepancies can probably be attributed to our neglect of finite temperature effects and/or the typical uncertainties associated with quantum chemistry calculations. For

molecule XXXII, however, the discrepancies between the computed crystal energy landscape and the experimental understanding appear too large to be explained by such factors. Additional work to quantify the modeling errors associated with the intermolecular contribution to the lattice energies would be valuable. At the same time, further experimental efforts to solve the unknown molecule XXXII crystal structures and reveal their relative stabilities are clearly needed. Finally, further investigations should also consider whether the crystallization of the thermodynamically most-stable forms on the computational crystal energy landscape might be hindered by their highly-unfavorable intramolecular conformations.

Supporting Information: Crystal energy landscapes in graphical and tabulated forms, tables of the conformational energies, additional conformational energy scans, and data from the DLPNO-CCSD(T) convergence testing are provided in PDF format. B86bPBE-XDM optimized crystal structures and the molecular geometries extracted from the crystals are provided in CIF and XYZ formats, respectively.

Acknowledgements G.J.O.B. gratefully acknowledges funding from the National Science Foundation (CHE-1955554) and supercomputer time from ACCESS (CHE110064). Computations were also performed using the computer clusters and data storage resources of the UC Riverside HPCC, which were funded by grants from NSF (MRI-2215705, MRI-1429826) and NIH (1S10OD016290-01A1). Additional support for this work to P.A.U. came from the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under contract number DE-SC0014664. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of DOE, ORAU,

or ORISE.

References

Abramov, Y. A., Zhang, P., Zeng, Q., Yang, M., Liu, Y. & Sekharan, S. (2020). *Cryst. Growth Des.* **20**(3), 1512–1525.
https://pubs.acs.org/doi/10.1021/acs.cgd.9b01153

Adamo, C. & Barone, V. (1999). *J. Chem. Phys.* **110**, 6158.

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 Crystallographica Section B Structural Science*, **67**(6), 535–551.
https://scripts.iucr.org/cgi-bin/paper?S0108768111042868

Becke, A. D. (1986). *J. Chem. Phys.* **85**, 7184–7187.

Beran, G. J. O. (2019). *CrystEngComm*, **21**, 758–764.

Beran, G. J. O. (2023). *Chem. Sci.* **14**, 13290–13312.

Beran, G. J. O., Greenwell, C., Cook, C. & Řezáč, J. (2023). *Acc. Chem. Res.* **56**, 3525–3534.

Beran, G. J. O., Sugden, I. J., Greenwell, C., Bowskill, D. H., Pantelides, C. C. & Adjiman, C. S. (2022*a*). *Chem. Sci.* **13**, 1288–1297.

Beran, G. J. O., Wright, S. E., Greenwell, C. & Cruz-Cabeza, A. J. (2022*b*). *J. Chem. Phys.* **156**, 104112.

Bhardwaj, R. M., McMahon, J. A., Nyman, J., Price, L. S., Konar, S., Oswald, I. D. H., Pulham, C. R., Price, S. L. & Reutzel-Edens, S. M. (2019). *J. Am. Chem. Soc.* **141**(35), 13887–13897.
https://pubs.acs.org/doi/10.1021/jacs.9b06634

Bowskill, D. H., Sugden, I. J., Konstantinopoulos, S., Adjiman, C. S. & Pantelides, C. C. (2021). *Ann. Rev. Chem. Biomol. Eng.* **12**(1), 593–623.
https://www.annualreviews.org/doi/10.1146/annurev-chembioeng-060718-030256

Brandenburg, J. G. & Grimme, S. (2014). In *Top. Curr. Chem.*, edited by S. Atahan-Evrenk & A. Aspuru-Guzik, vol. 345 of *Topics in Current Chemistry*, pp. 1–24. Cham: Springer International Publishing.
https://link.springer.com/10.1007/128

Braun, D. E., Lingireddy, S. R., Beidelschies, M. D., Guo, R., Müller, P., Price, S. L. & Reutzel-Edens, S. M. (2017). *Cryst. Growth Des.* **17**(10), 5349–5365.

Braun, D. E., McMahon, J. A., Bhardwaj, R. M., Nyman, J., Neumann, M. A., Van De Streek, J. & Reutzel-Edens, S. M. (2019). *Cryst. Growth Des.* **19**(5), 2947–2962.

Caldeweyher, E., Bannwarth, C. & Grimme, S. (2017). *J. Chem. Phys.* **147**(3), 034112.
http://aip.scitation.org/doi/10.1063/1.4993215

Cook, C. J., Li, W., Lui, B. F., Gately, T. J., Al-Kaysi, R. O., Mueller, L. J., Bardeen, C. J. & Beran, G. J. O. (2022). *Chem. Sci.* **14**, 937–949.

Cook, C. J., Perry, C. J. & Beran, G. J. O. (2023). *J. Phys. Chem. Lett.* **14**, 6823–6831.

Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. (2015). *Chem. Soc. Rev.* **44**, 8619–8635.
http://xlink.rsc.org/?DOI=C5CS00227C

Cutini, M., Civalleri, B., Corno, M., Orlando, R., Brandenburg, J. G., Maschio, L. & Ugliengo, P. (2016). *J. Chem. Theory Comput.* **12**(7), 3340–3352.
https://pubs.acs.org/doi/10.1021/acs.jctc.6b00304

Day, G. M., Cooper, T. G., Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., Della Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., Leusen, F. J. J., Kendrick, J., Price, S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G. W. A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., van de Streek, J., Wolf, A. K. & Schweizer, B. (2009). *Acta Cryst. B*, **65**(Pt 2), 107–125.
http://www.ncbi.nlm.nih.gov/pubmed/19299868

Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Leusen, F. J. J., Liang, C., Pantelides, C. C., Karamertzanis, P. G., Price, S. L., Lewis, T. C., Nowell, H., Torrisi, A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U. & Verwer, P. (2005). *Acta Cryst. B*, **61**(Pt 5), 511–27.
http://www.ncbi.nlm.nih.gov/pubmed/16186652

Dovesi, R., Erba, A., Orlando, R., Zicovich-Wilson, C. M., Civalleri, B., Maschio, L., Rérat, M., Casassa, S., Baima, J., Salustro, S. & Kirtman, B. (2018). *WIREs Comput. Mol. Sci.* **8**(4), e1360.
http://doi.wiley.com/10.1002/wcms.1360

Dunning, T. H. (1989). *J. Chem. Phys.* **90**, 1007–1023.

Firaha, D., Liu, Y. M., van de Streek, J., Sasikumar, K., Dietrich, H., Helfferich, J., Aerts, L., Braun, D. E., Broo, A., DiPasquale, A. G., Lee, A. Y., Le Meur, S., Nilsson Lill, S. O., Lunsmann, W. J., Mattei, A., Muglia, P., Putra, O. D., Raoui, M., Reutzel-Edens, S. M., Rome, S., Sheikh, A. Y., Tkatchenko, A., Woollam, G. R. & Neumann, M. A. (2023). *Nature*, **623**(7986), 324–328.
https://www.nature.com/articles/s41586-023-06587-3

Gately, T. J., Cook, C. C., Almuzarie, R., Islam, I., Gardner, Z. T., Iuliucci, R. J., Al-Kaysi, R. O., Beran, G. J. O. & Bardeen, C. J. (2022). *Cryst. Growth Des.* **22**, 7298–7307.

Gately, T. J., Sontising, W., Easley, C. J., Islam, I., Al-Kaysi, R. O., Beran, G. J. O. & Bardeen, C. J. (2021). *CrystEngComm*, **23**, 5931–5943.

Giannozzi, P., Andreussi, O., Brumme, T., Bunau, O., Buongiorno Nardelli, M., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Cococcioni, M., Colonna, N., Carnimeo, I., Dal Corso, A., de Gironcoli, S., Delugas, P., DiStasio, R. A., Ferretti, A., Floris, A., Fratesi, G., Fugallo, G., Gebauer, R., Gerstmann, U., Giustino, F., Gorni, T., Jia, J., Kawamura, M., Ko, H.-Y., Kokalj, A., Küçükbenli, E., Lazzeri, M., Marsili, M., Marzari, N., Mauri, F., Nguyen, N. L., Nguyen, H.-V., Otero-de-la Roza, A., Paulatto, L., Poncé, S., Rocca, D., Sabatini, R., Santra, B., Schlipf, M., Seitsonen, A. P., Smogunov, A., Timrov, I., Thonhauser, T., Umari, P., Vast, N., Wu, X. & Baroni, S. (2017). *J. Phys. Condens. Mat.* **29**(46), 465901.
http://stacks.iop.org/0953-8984/29/i=46/a=465901

Greenwell, C. & Beran, G. J. O. (2018). The MP2D software library.
https://github.com/Chandemonium/MP2D

Greenwell, C. & Beran, G. J. O. (2020). *Cryst. Growth Des.* **20**, 4875–4881.

Greenwell, C. & Beran, G. J. O. (2021). *J. Mater. Chem. C*, **9**, 2848–2857.

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.

Greenwell, C., Řezáč, J. & Beran, G. J. O. (2022). *Phys. Chem. Chem. Phys.* **24**, 3695–3712.

Grimme, S., Brandenburg, J. G., Bannwarth, C. & Hansen, A. (2015). *J. Chem. Phys.* **143**(5).

Grimme, S., Hansen, A., Ehlert, S. & Mewes, J.-M. (2021). *J. Chem. Phys.* **154**(6), 064103. https://doi.org/10.1063/5.0040021 https://aip.scitation.org/doi/10.1063/5.0040021

Guo, Y., Riplinger, C., Becker, U., Liakos, D. G., Minenkov, Y., Cavallo, L. & Neese, F. (2018). *J. Chem. Phys.* **148**(1), 011101.
http://dx.doi.org/10.1063/1.5011798

Helgaker, T., Klopper, W., Koch, H. & Noga, J. (1997). *J. Chem. Phys.* **106**, 9639–9646.

Hoja, J., Ko, H.-Y., Neumann, M. A., Car, R., DiStasio, R. A. & Tkatchenko, A. (2019). *Science Adv.* **5**(1), eaau3338.
http://advances.sciencemag.org/lookup/doi/10.1126/sciadv.aau3338

Hunnisett, L. M., Abraham, N. S., Aitipamula, S. *et al.* (2024). *Acta Cryst. B*, **in preparation**.

Karamertzanis, P. G., Day, G. M., Welch, G. W. A., Kendrick, J., Leusen, F. J. J., Neumann, M. A. & Price, S. L. (2008). *J. Chem. Phys.* **128**(24), 244708. http://www.ncbi.nlm.nih.gov/pubmed/18601366

LeBlanc, L. M., Dale, S. G., Taylor, C. R., Becke, A. D., Day, G. M. & Johnson, E. R. (2018). *Angew. Chem. Int. Ed.* **57**(45), 14906–14910. http://doi.wiley.com/10.1002/anie.201809381

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst. B*, **56**(4), 697–714. http://scripts.iucr.org/cgi-bin/paper?S0108768100004584

Mardirossian, N. & Head-Gordon, M. (2016). *J. Chem. Phys.* **144**(21), 214110. http://dx.doi.org/10.1063/1.4952647 http://aip.scitation.org/doi/10.1063/1.4952647

Martin, J. M. L. & Santra, G. (2020). *Israel J. Chem.* **60**(8-9), 787–804. https://onlinelibrary.wiley.com/doi/10.1002/ijch.201900114

Mortazavi, M., Hoja, J., Aerts, L., Quéré, L., van de Streek, J., Neumann, M. A. & Tkatchenko, A. (2019). *Commun. Chem.* **2**(1), 70. http://dx.doi.org/10.1038/s42004-019-0171-y http://www.nature.com/articles/s42004-019-0181-9

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. A., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2002). *Acta Cryst. B*, **58**(4), 647–661. http://scripts.iucr.org/cgi-bin/paper?S0108768102005669

Neese, F. (2012). *WIREs Comput. Molec. Sci.* **2**, 73–78.

Neumann, M. A., Leusen, F. J. J. & Kendrick, J. (2008). *Angew. Chem. Int. Ed.* **47**(13), 2427–2430.

Neumann, M. A., van de Streek, J., Fabbiani, F. P. A., Hidber, P. & Grassmann, O. (2015). *Nat. Comm.* **6**(May), 7793. http://www.nature.com/doifinder/10.1038/ncomms8793

Nikhar, R. & Szalewicz, K. (2022). *Nat. Comm.* **13**(1), 3095. https://www.nature.com/articles/s41467-022-30692-y

Nyman, J. & Day, G. M. (2015). *CrystEngComm*, **17**(28), 5154–5165. http://dx.doi.org/10.1039/C5CE00045A

Nyman, J., Yu, L. & Reutzel-Edens, S. M. (2019). *CrystEngComm*, **21**(13), 2080–2088. http://xlink.rsc.org/?DOI=C8CE01902A

Oganov, A. R. (2018). *Faraday Disc.* **211**, 643–660. https://pubs.rsc.org/en/content/articlepdf/2018/fd/c8fd90033g http://xlink.rsc.org/?DOI=C8FD90033G

Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865.

Perry, C. J. & Beran, G. J. O. (2023). *Cryst. Growth Des.* **23**, 8352–8360.

Price, A. J., Mayo, R. A., Otero-De-la Roza, A. & Johnson, E. R. (2023). *CrystEngComm*, **25**(6), 953–960.

Price, A. J., Otero-De-La-Roza, A. & Johnson, E. R. (2022). *Chem. Sci.* **14**(5), 1252–1262.

Price, S. L. (2008). *Int. Rev. Phys. Chem.* **27**(3), 541–568.

Reilly, A. M., Cooper, R. I., Adjiman, C. S., Bhattacharya, S., Boese, A. D., Brandenburg, J. G., Bygrave, P. J., Bylsma, R., Campbell, J. E., Car, R., Case, D. H., Chadha, R., Cole, J. C., Cosburn, K., Cuppen, H. M., Curtis, F., Day, G. M., DiStasio Jr, R. A., Dzyabchenko, A., van Eijck, B. P., Elking, D. M., van den Ende, J. A., Facelli, J. C., Ferraro, M. B., Fusti-Molnar, L., Gatsiou, C.-A., Gee, T. S., de Gelder, R., Ghiringhelli, L. M., Goto, H., Grimme, S., Guo, R., Hofmann, D. W. M., Hoja, J., Hylton, R. K., Iuzzolino, L., Jankiewicz, W., de Jong, D. T., Kendrick, J., de Klerk, N. J. J., Ko, H.-Y., Kuleshova, L. N., Li, X., Lohani, S., Leusen, F. J. J., Lund, A. M., Lv, J., Ma,

Y., Marom, N., Masunov, A. E., McCabe, P., McMahon, D. P., Meekes, H., Metz, M. P., Misquitta, A. J., Mohamed, S., Monserrat, B., Needs, R. J., Neumann, M. A., Nyman, J., Obata, S., Oberhofer, H., Oganov, A. R., Orendt, A. M., Pagola, G. I., Pantelides, C. C., Pickard, C. J., Podeszwa, R., 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. B*, **72**(4), 439–459.
http://scripts.iucr.org/cgi-bin/paper?S2052520616007447

Riplinger, C., Pinski, P., Becker, U., Valeev, E. F. & Neese, F. (2016). *J. Chem. Phys.* **144**(2), 024109.
http://dx.doi.org/10.1063/1.4939030

Riplinger, C., Sandhoefer, B., Hansen, A. & Neese, F. (2013). *J. Chem. Phys.* **139**(13), 134101.
http://aip.scitation.org/doi/10.1063/1.4821834

Otero-de-la Roza, A. & Johnson, E. R. (2012). *J. Chem. Phys.* **136**(17), 174109.
http://www.ncbi.nlm.nih.gov/pubmed/22583212

Otero-de-la Roza, A., LeBlanc, L. M. & Johnson, E. R. (2019). *J. Chem. Theory Comput.* **15**, 4933–4944.

Santra, G., Sylvetsky, N. & Martin, J. M. L. (2019). *J. Phys. Chem. A*, **123**(24), 5129–5143.
https://pubs.acs.org/doi/10.1021/acs.jpca.9b03157

Smith, D. G. A., Burns, L. A., Simmonett, A. C., Parrish, R. M., Schieber, M. C., Galvelis, R., Kraus, P., Kruse, H., Di Remigio, R., Alenaizan, A., James, A. M., Lehtola, S., Misiewicz, J. P., Scheurer, M., Shaw, R. A., Schriber, J. B., Xie, Y., Glick, Z. L., Sirianni, D. A., O'Brien, J. S., Waldrop, J. M., Kumar, A., Hohenstein, E. G., Pritchard, B. P., Brooks, B. R., Schaefer, H. F., Sokolov, A. Y., Patkowski, K., DePrince, A. E., Bozkaya, U., King, R. A., Evangelista, F. A., Turney, J. M., Crawford, T. D. & Sherrill, C. D. (2020). *J. Chem. Phys.* **152**(18), 184108.
https://doi.org/10.1063/5.0006002 http://aip.scitation.org/doi/10.1063/5.0006002

Sure, R. & Grimme, S. (2013). *J. Comp. Chem.* **34**(19), 1672–1685.
http://doi.wiley.com/10.1002/jcc.23317

Tan, M., Shtukenberg, A. G., Zhu, S., Xu, W., Dooryhee, E., Nichols, S. M., Ward, M. D., Kahr, B. & Zhu, Q. (2018). *Faraday Disc.* **211**, 477–491.
http://xlink.rsc.org/?DOI=C8FD00039E

Thomas, S. P. & Spackman, M. A. (2018). *Austral. J. Chem.* **71**(4), 279.
http://www.publish.csiro.au/?paper=CH17620

Thompson, H. P. G. & Day, G. M. (2014). *Chem. Sci.* **5**(8), 3173–3182.
http://xlink.rsc.org/?DOI=c4sc01132e

Řezáč, J. (2020). *J. Chem. Theory Comput.* **16**(4), 2355–2368.
https://pubs.acs.org/doi/10.1021/acs.jctc.9b01265

Whittleton, S. R., Otero-de-la Roza, A. & Johnson, E. R. (2017*a*). *J. Chem. Theory Comput.* **13**(2), 441–450.
http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b00679

Whittleton, S. R., Otero-de-la Roza, A. & Johnson, E. R. (2017*b*). *J. Chem. Theory Comput.* **13**(11), 5332–5342.
http://pubs.acs.org/doi/10.1021/acs.jctc.7b00715

Řezáč, J. (2020). *J. Chem. Theory Comput.* **16**(10), 6305–6316.

Řezáč, J. (2022). *Phys. Chem. Chem. Phys.* **24**(24), 14780–14793.
http://xlink.rsc.org/?DOI=D2CP01602H

Synopsis

The conformational energies and crystal energy landscapes for molecules XXXI and XXXII from the 7th Blind Test of Crystal Structure Prediction are investigated with a variety of electronic structure methods. Whereas molecule XXXI can be modeled relatively straightforwardly, molecule XXXII proves challenging even for state-of-the-art quantum chemistry techniques.

Graphical Abstract