EDGE ARTICLE

Check for updates

From polymorphs to cocrystals and salts: successfully predicting axitinib's challenging crystal forms

Gregory J. O. Beran [ID] *

The experimental development of cancer drug axitinib was disrupted by two surprise discoveries of new, more stable crystal polymorphs. Organic crystal structure prediction can help de-risk against such occurrences, but previous attempts failed to predict the theoretically challenging axitinib conformational polymorph lattice energies reliably. Here, we demonstrate how modern crystal structure prediction methodologies can not only successfully predict the problematic axitinib crystal structures, but how they can also accurately distinguish between salt and cocrystal forms in three axitinib multi-component crystals, which has been a long-standing challenge for organic crystal modeling. These successes derive from addressing the density-driven delocalization error found in commonly-used generalized gradient approximation density functional theory models through the application of intramolecular energy corrections and/or hybrid density functionals. Notably, the simultaneous combination of both approaches provides more robust lattice energy predictions than either individual approach.

1 Introduction

Axitinib, Pfizer's tyrosine kinase inhibitor for advanced renal cell carcinoma, has a notoriously complex solid-form landscape. Its development provides a case study in formulation challenges, with five known conformational polymorphs and dozens of solvates, cocrystals, and salts.[1–4] Initial formulation efforts focused on form IV, but these plans were upended by the surprise discovery of a more stable form XXV.[1] Soon after, an even more stable polymorph, form XLI, was found serendipitously and became the final commercial product.[2]

Cases such as axitinib, the recalls of ritonavir[5,6] and rotigotine,[7,8] and other drugs with complex solid-form landscapes[9–11] have motivated the use of crystal structure prediction (CSP) to help de-risk against undiscovered polymorphs.[12–14] When reliable, CSP can predict the potential for discovering more stable polymorphs,[8,15] help solve structures of polymorphs for which single-crystal X-ray diffraction data is unavailable,[11,16–20] assess possible disorder or crystallization difficulties,[9,21–25] anticipate hydrate formation,[10,26–29] and identify useful coformer species when developing salts and cocrystals.[30–36] On the other hand, erroneous CSP predictions can send experimental researchers in pursuit of phantom polymorphs,[11] potentially delaying a drug or even causing it to be canceled.

Attempts to predict the axitinib crystal energy landscape a decade ago met with only partial success due to its challenging conformational polymorphism. After a couple preliminary

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

studies from Pfizer,[37,38] Vasileiadis and co-workers carried out a more thorough CSP in 2015.[39] Despite the challenging intramolecular flexibility of axitinib, their careful searches over candidate structures containing a single molecule in the asymmetric unit ($Z' = 1$) successfully generated a crystal landscape containing all four known $Z' = 1$ polymorphs (Fig. 1). Unfortunately, the lattice energy models employed ranked the polymorph stabilities poorly. Notably, the experimentally most-stable form XLI occurs at rank 108, +11.5 kJ mol^{-1} above the
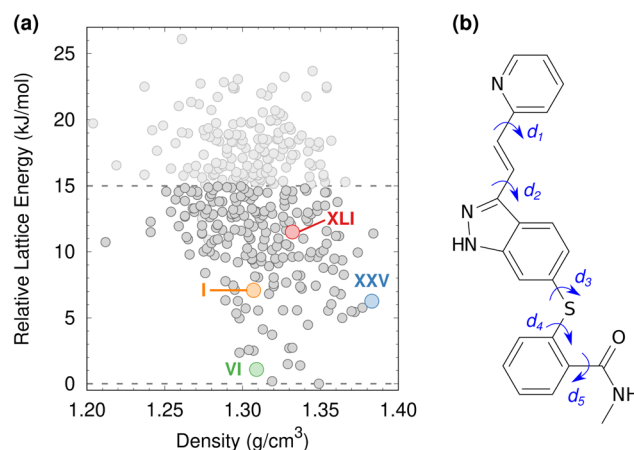

Fig. 1 (a) Previously-published $Z' = 1$ crystal energy landscape showing the relative lattice energies and densities of the predicted axitinib crystal structures.[39] The experimental (colored) and low-energy structures (<15 kJ mol^{-1}) indicated in dark gray were refined here. (b) Structure of axitinib highlighting five key dihedral angles.

global minimum, and higher than forms VI (rank 3, +1.1 kJ mol^{-1}), XXV (rank 26, +6.3 kJ mol^{-1}), and I (rank 32, +7.1 kJ mol^{-1}). Experiment also suggests forms XXV and VI are close in energy, contrary to the 5 kJ mol^{-1} separation predicted in the CSP.

Since then, the widespread shift toward using periodic density functional theory (DFT) models for the crystal structure refinements and lattice energies has substantially improved crystal energy rankings for pharmaceutical-sized molecules, as evidenced by the progress in the recent blind tests of crystal structure prediction.[40–44] Less computationally demanding generalized gradient approximation (GGA) density functionals are often used for initial or routine CSP studies,[45,46] though switching to more reliable and computationally expensive hybrid density functionals for the final energy ranking stage can be important.[47–52] Augmenting the DFT treatment with corrections for van der Waals dispersion is essential, typically *via* Grimme's D3/D4 corrections,[53–55] the many-body dispersion (MBD) approach,[56–58] or the exchange-hole dipole moment (XDM) model.[59] Machine-learned potentials trained on DFT data are also rapidly improving and reducing the computational costs associated with generating crystal energy landscapes.[60–65] Advances in treating finite-temperature crystal thermochemistry are further improving the utility of CSP for predicting structures and properties at ambient conditions.[64,66–72] With these techniques, successful DFT-driven CSPs continue to accumulate for species ranging from small, rigid molecules[51,73–75] to larger, flexible pharmaceuticals.[8–11,15,21,34,46,49,66,68,76,77]

Despite numerous successes, challenging CSP cases remain. Many of these difficult systems exhibit density-driven delocalization error—the tendency of approximate density functionals to overly delocalize electron densities.[78] Delocalization error creates a myriad of problems in CSP. For example, it over-stabilizes molecular conformations with extended π-conjugation, which can lead to erroneous energy rankings in conformational polymorphs.[46,79–84] Even larger energy errors can occur for tautomeric polymorphs (a.k.a desmotropes) when tautomerization changes the π-conjugation character.[75] Intermolecular delocalization error can cause spurious proton transfer in acid–base cocrystals, incorrectly predicting salt formation.[85] Delocalization error also tends to overestimate the strength of hydrogen and halogen bonds,[86] which distorts crystal landscapes that contain varied hydrogen- or halogen-bonded packing motifs.

Two primary strategies have emerged to address these delocalization error challenges in recent years. First, delocalization error is pronounced in GGA functionals, but hybrid density functionals are less affected thanks to the inclusion of some amount of exact exchange (EXX) in the functional. The extent of reduction is determined by the specific admixture of exact exchange: whereas ~20–25% EXX is typical for general-purpose global hybrid functionals such as B3LYP[87] or PBE0,[88] larger fractions around ~50% EXX are often recommended for systems with substantial delocalization error issues.[78] Such hybrid functionals can substantially improve the quality of crystal structure predictions in difficult delocalization error

cases,[50,51,75,85,89,90] but there remains the challenge that the optimal fraction of exact exchange for a given system is not always obvious *a priori*.[89] Moreover, the use of hybrid functionals with ~25–50% EXX alone appears to be insufficient to predict reliable energy differences for tautomeric polymorphs, for example.[75]

The second strategy applies intramolecular energy corrections computed at a higher-level of theory to the periodic DFT treatment.[91,92] Because the corrections are computed on isolated molecules, they can be predicted very reliably using advanced electronic structure techniques such as double-hybrid density functionals or correlated wavefunction methods. Of course, an entirely intramolecular correction does not address intermolecular delocalization error. Nevertheless, this approach has proved effective in small molecules,[80–82,91] pharmaceuticals,[79,83] and photomechanical crystals.[73,93,94] For the ROY molecule, for example, combining a periodic GGA functional with higher-level intramolecular energy corrections enabled the first CSP[81] that ranked the polymorphs correctly, differentiated energetically between the routine and harder-to-crystallize/metastable polymorphs, demonstrated that the lowest-energy crystal forms had already been discovered experimentally, and even included the subsequently-discovered O22 polymorph.[95]

The solid forms of axitinib similarly present significant delocalization error challenges: its wide variety of conformational polymorphs with varying degrees of π conjugation make the lattice energies difficult to predict reliably.[79] Further challenges arise when considering multi-component crystals of axitinib, where the experimental salt *versus* cocrystal character changes across different carboxylic acid coformers.[3] A decade after the last attempt,[39] the present study revisits the difficult crystal structure prediction of axitinib to assess current techniques for 0 K organic crystal lattice energy ranking. The results here will demonstrate that while routine GGA density functionals are inadequate for this system, state-of-the-art intramolecular correction and/or hybrid DFT techniques could have anticipated that form XLI is the global minimum structure and would have revealed that most of the lowest-energy predicted polymorphs of axitinib have already been discovered experimentally. These same techniques also overcome a long-standing challenge by successfully differentiating between salt and cocrystal forms of three axitinib + carboxylic acid formulations. Importantly, while both strategies for addressing delocalization error are effective, this work highlights how the simultaneous combination of hybrid DFT with intramolecular corrections mitigates the shortfalls of each approach and enables more robust crystal lattice energy predictions. Overall, these axitinib results showcase the tremendous progress made in the CSP of difficult pharmaceutical systems over the past decade.

2 Computational methods

2.1 Crystal structure geometry optimizations

This study begins with the initial axitinib crystal energy landscape from ref. 96 and shown in Fig. 1. 204 of the 207 crystal structures lying within 15 kJ mol^{-1} of the global minimum

energy structure were optimized here. The three remaining structures within this energy window were omitted for computational expedience due to their large, 16-molecule (736 atoms) unit cells. Those omitted structures range from $+10.9$–12.2 kJ mol^{-1} on the original landscape, and all three adopt packing motifs that consistently prove energetically unfavorable with the electronic structure models used here, as will be discussed further below. Because the previous CSP only searched over $Z' = 1$ structures, the initial experimental $Z' = 2$ structure of form IV was obtained from the Cambridge Structure Database (reference code VUSDIX01 (ref. 1)), relaxed, and added to the landscape. Initial experimental multi-component crystal structures[3] for axitinib with fumaric (UZAKEO), suberic (UZAKAK), and *trans*-cinnamic acid (UZAKIS) were also obtained from the database.

All single- and multi-component crystal structures were fully relaxed (both atomic positions and lattice parameters) with periodic planewave DFT using the B86bPBE density functional[97,98] and exchange-hole dipole moment (XDM) dispersion correction[59] in Quantum Espresso v6.5.[99] Initial optimizations employed the projector augmented wave treatment of the core electrons, a 40 Ry planewave cutoff, and Monkhorst k-point grids with density of at least 0.05 Å^{-1} were used. Structures lying within 15 kJ mol^{-1} of the global minimum either before or after the intramolecular energy correction (as described below) was applied were then further refined with a larger 50 Ry planewave cutoff.

To map out the potential energy curves for the proton transfer coordinates in the multi-component crystals, a series of constrained optimizations were performed, starting from the fully-relaxed structure and freezing the key axitinib N–H distances at various values for each axitinib molecule in the unit cell. The lattice parameters were held fixed at their fully-relaxed values during the constrained proton transfer optimizations due to software limitations.

2.2 Single-point energy refinements

Two different approaches were used to refine the crystal energies. First, because the inclusion of some exact exchange in the density functional reduces density-driven delocalization error, periodic DFT was employed with both generalized-gradient approximation (GGA) and hybrid density functionals, using either the B86bPBE-XDM (0% EXX), B86bPBE-25-XDM (25% EXX), and B86bPBE-50-XDM (50% EXX)[50] or PBE-D3(BJ) (0% EXX) and PBE0-D3(BJ) (25% EXX)[54,88,98] families of functionals. These single-point energy calculations were performed in FHI-Aims version 240920.2.[100,101] GGA calculations employed the triple-ζ "Tight" numerical atomic orbital basis set distributed with the code. To make the hybrid functional calculations more affordable, the "Tight" basis results were approximated from a mixture of "Tight" and double-ζ "Light" basis set calculations according to:[51]

$$E_{\text{hybrid}}^{\text{tight}} \approx E_{\text{hybrid}}^{\text{light}} + (E_{\text{GGA}}^{\text{tight}} - E_{\text{GGA}}^{\text{light}}) \quad (1)$$

Second, the intramolecular correction approach[79] refines single-point periodic GGA or hybrid DFT energies according to,

$$E_{\text{crystal}}^{\text{corrected}} = E_{\text{crystal}}^{\text{DFT}} + \sum_i^Z \left(E_{i,\text{gas}}^{\text{higher}} - E_{i,\text{gas}}^{\text{DFT}} \right) \quad (2)$$

The terms in parentheses correspond to the energy difference between the same DFT functional used for the periodic crystal and some higher-level electronic structure method, as calculated on the isolated molecules using molecular geometries extracted directly from the crystal. The sum runs over all Z molecules in the unit cell, though space group symmetry often greatly reduces the number of unique electronic structure calculations that actually need to be performed. The choice of the higher-level method is flexible. For the pure axitinib CSP, domain-local pair natural orbital coupled cluster singles, doubles, and perturbative triples (DLPNO-CCSD(T1))[102,103] was chosen. The DLPNO-CCSD(T1) calculations were run in Orca v6.0 (ref. 104) using the aug-cc-pVDZ basis set, `TightPNO` settings, and `TCutMKN` to 10^{-4}. They were then approximately extrapolated to the complete-basis-set (CBS) limit *via* the focal point method,[105]

$$E_{\text{DLPNO-CCSD(T1)}}^{\text{CBS}} \approx E_{\text{DLPNO-CCSD(T1)}}^{\text{aDZ}} + (E_{\text{MP2}}^{\text{CBS}} - E_{\text{MP2}}^{\text{aDZ}}). \quad (3)$$

The CBS-limit MP2 results in eqn (3) were obtained by extrapolating the correlation energy E_{corr} computed in the aug-cc-pVTZ and aug-cc-pVQZ basis sets[106] and pairing it with the aug-cc-pVQZ Hartree–Fock energy E_{HF},

$$E_{\text{MP2}}^{\text{CBS}} \approx E_{\text{HF}}^{\text{aQZ}} + \left(\frac{4^3 E_{\text{corr}}^{\text{aQZ}} - 3^3 E_{\text{corr}}^{\text{aTZ}}}{4^3 - 3^3} \right) \quad (4)$$

For the acid–base salt/cocrystal cases, for which the molecular identity varies from neutral to charged across the proton transfer coordinate, the "intramolecular" corrections were performed on a dimer unit containing both the proton-donor carboxylic acid species and the axitinib acceptor molecule. In addition, because the electronic structure of charged functional groups can vary considerably between the gas- and condensed phases, the intramolecular calculations were embedded in a polarizable continuum model. The optimal choice of the dielectric constant used in the PCM is potentially ambiguous: dielectric constants of ~ 3 are often recommended for organic crystals comprised of typical neutral species, while larger dielectric constants are recommended for zwitterionic species or organic salts. For example, in the 6th blind test, Day and co-workers found that a dielectric constant of ~ 7 improved their predictions in such systems.[42]

Fortunately, although the energies of individual species are sensitive to the chosen dielectric constant, the difference between the low-level and high-level method energies in eqn (2) is much less sensitive. For example, varying the dielectric constant over ~ 3–7 alters the "intramolecular" correction energies in the multi-component axitinib crystals by just a few tenths of a kJ mol^{-1}. Therefore, an intermediate dielectric constant of $\varepsilon = 4.9$ (chloroform) has been adopted here. Note that PCM embedding for the intramolecular correction terms in the pure axitinib CSP case is unnecessary, since it impacts the

relative polymorph energies by a mere 0.1 kJ mol^{-1} on average. See SI Section S3 for additional details.

To ensure consistency in the PCM treatment between the low- and high-level calculations in the intramolecular correction, both were performed in Orca. Since the B86bPBE-XDM family of models are currently unavailable in Orca, the intramolecular corrections were applied to PBE-D3(BJ) and PBE0-D3(BJ) crystal energies. Furthermore, the double-hybrid revDSD-PBEP86-D4 functional[107] was chosen for the higher-level method instead of DLPNO-CCSD(T1), since the former is compatible with the PCM model. All of these Orca DFT calculations employed the def2-QZVP basis set and conductor-like PCM model.

3 Results and discussion

3.1 Crystal energy landscape of neat axitinib

Consider first the crystal structure prediction of axitinib itself. The previous CSP by Vasileiadis and co-workers[39] searched over unit cells from 59 common space groups that contain a single molecule in the asymmetric unit ($Z' = 1$). It resulted in a crystal energy landscape containing 361 crystal structures, including all four reported $Z' = 1$ crystal polymorphs. However, the landscape ranked those known polymorphs poorly, as discussed above.

Here, optimizing the low-energy crystal structures indicated in Fig. 1 plus experimental form IV ($Z' = 2$) with periodic B86bPBE-XDM dramatically transforms the energy landscape (Fig. 2a). Just 40 of the 204 structures within 15 kJ mol^{-1} of the global minimum in the earlier landscape remain within that energy window; all others now lie higher in energy. Given that most experimentally-observed polymorphs under ordinary conditions lie within a ~10 kJ mol^{-1} energy window,[108,109] this suggests that the vast majority of structures on the original landscape were not experimentally relevant. A large fraction of the structures eliminated by DFT refinement contain one of three common structural motifs whose stabilities were evidently exaggerated by the potential used in ref. 39: structures with doubly hydrogen-bonded indazole dimers, trimers with two

indazole groups hydrogen bonded to a third molecule's indazole group, and intramolecular hydrogen-bonding N–H⋯S interactions involving the amide nitrogen (SI Section S2.1).

Encouragingly, the B86bPBE-XDM refinement also places the four most important experimental structures near the bottom of the crystal energy landscape (Fig. 2a). Only form I, which is the least-stable polymorph experimentally, lies higher in energy (+11.9 kJ mol^{-1}). On the other hand, the crystal structure ordering of the remaining four experimental polymorphs remains incorrect, with form XXV erroneously predicted to be the most stable, and form XLI lying 3 kJ mol^{-1} higher at rank 3.

To understand this B86bPBE-XDM landscape better, consider the electronic structure of the axitinib molecule. It contains five key dihedral angles that can potentially impact the extent of π conjugation (Fig. 1). Surveying the distribution of dihedral angles among the structures in the Fig. 2 landscapes, we find that dihedrals d_1 and d_2 vary minimally in practice (up to 25° out of plane, but often by less than 10°). Similarly, angle d_3 mostly varies between 50–90° relative to the indole ring, preventing significant conjugation between the sulfur lone pairs and the indole π electron cloud. In contrast, dihedral angles d_4 and d_5 range from ~0–70° relative to the phenyl ring, leading to considerable variation in the extent of conjugation between the phenyl ring and the sulfur lone pairs and/or the amide group across the crystal structures.

Density-driven delocalization error in GGA functionals such as B86bPBE-XDM artificially stabilizes the intramolecular conformations that delocalize electron density through more extended π conjugation. In other words, structures with d_4 and/or d_5 angles near 0° will be over-stabilized relative to those closer to 90°. Indeed, comparison of the gas-phase B86bPBE-XDM and DLPNO-CCSD(T1) conformational energies from the 54 most stable structures shows good correlations ($R^2 > 0.7$) between the intramolecular B86bPBE-XDM conformational energy error and both dihedral angles d_4 and d_5 (SI Section S2.3). In contrast, the B86bPBE-XDM conformational errors are completely uncorrelated with d_1, d_2, or d_3 ($R^2 < 0.02$). Further evidence for the role of density-driven delocalization error
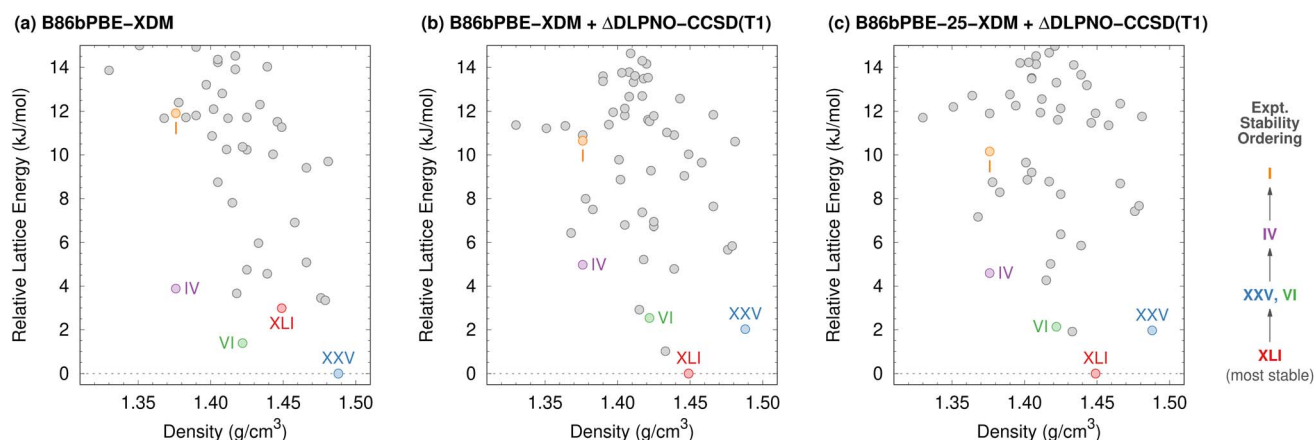

Fig. 2 Comparison of the axitinib crystal energy landscapes at 0 K computed with (a) B86bPBE-XDM, (b) B86bPBE-XDM + intramolecular DLPNO-CCSD(T1) corrections, and (c) B86bPBE-25-XDM + intramolecular DLPNO-CCSD(T1) corrections. The finite-temperature experimental stability ordering is indicated at right.

comes from comparing gas-phase PBE-D4 and density-corrected[110,111] PBE-D4 conformational energies against the coupled cluster benchmarks. Whereas PBE-D4 exhibits a root-mean-square error of 3.5 kJ mol^{-1} relative to the DLPNO-CCSD(T1), evaluating the PBE-D4 energies on the Hartree–Fock densities (that do not suffer the same delocalization error) reduces the rms error to just 1.6 kJ mol^{-1}.

On the axitinib CSP landscape, the experimentally most-stable polymorph form XLI adopts a rather unique intra-molecular conformation in which the molecule folds to create π–π interactions between the amide group and the indole ring (Fig. 3).[79] The associated large values of dihedral angles d_4 (68°) and d_5 (55°) disrupt the conjugation of both sulfur and amide with the phenyl ring. As a result, the form XLI intramolecular conformation is artificially destabilized at the GGA level relative to many other candidate structures that have more planar d_4 and d_5 angles.

To refine the 0 K crystal energy landscape, intramolecular DLPNO-CCSD(T1) single-point energy corrections are applied to the B86bPBE-XDM lattice energies according to eqn (2). The magnitude of these corrections on the 54 lowest-energy structures is 2.9 kJ mol^{-1} on average and reaches 6 kJ mol^{-1}—enough to reorder the low-energy structures considerably. Crucially, form XLI becomes the most stable form on the corrected landscape (Fig. 2b), in agreement with experiment. The refined energies also reproduce the correct ordering among the experimental forms: XLI < XXV ~ VI < IV ≪ I. The very similar 0 K stabilities of forms XXV and VI is consistent with experimental observation that they are very close in energy. In addition, the higher energy of form IV comports with the experimental evidence that form XXV is more stable than form IV below 75 °C.

Fairly similar crystal energy landscapes are also obtained if one refines the polymorph energies using the hybrid DFT analogs of B86bPBE-XDM with either 25% or 50% EXX (Fig. 4) instead of applying the intramolecular correction. Both hybrid functionals correctly predict form XLI to be the most stable and reproduce the qualitative experimental stability ordering. Increasing the fraction of exact exchange further shifts
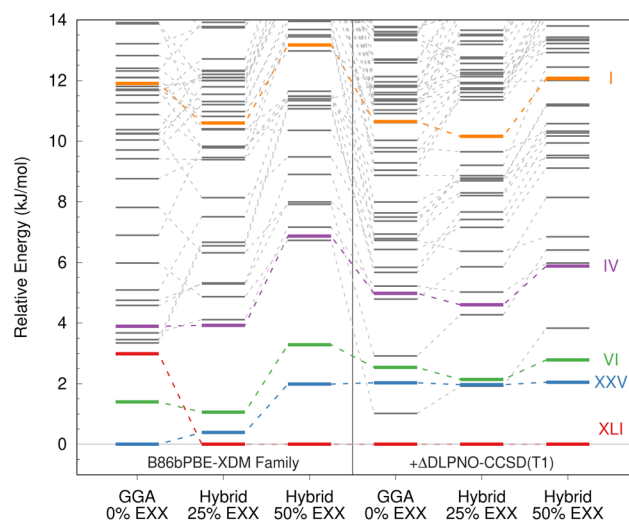

Fig. 4 Comparison of crystal energy landscapes for different base density functionals before (left side) and after (right side) the DLPNO-CCSD(T1) intramolecular correction is applied.

a number of the hypothetical structures higher in energy relative to most of the experimentally observed forms. Unfortunately, it is not always *a priori* obvious what fraction of exact exchange is optimal for a given system, and Fig. 4 shows that the relative energies of the low-energy crystal structures can vary by multiple kJ mol^{-1} between 25% and 50% EXX.

In this context, we observe a secondary benefit of the intramolecular correction approach for challenging systems such as axitinib: combining hybrid DFT and the intramolecular DLPNO-CCSD(T1) conformational energy correction somewhat reduces the variability in the relative polymorph energies as a function of the fraction of exact exchange. Similar behavior was also recently observed in cases of tautomeric polymorphism.[75] Pairing any of the three base density functionals with the DLPNO-CCSD(T1) correction gives results that generally lie intermediate between the uncorrected 25% and 50% EXX hybrid DFT energies.

That said, the effects of refining the intermolecular interactions *via* the use of hybrid DFT are non-negligible. These effects can be understood by considering, for example, that form XLI adopts extended networks of strong hydrogen bonds,[2] and increased exact exchange should disproportionately impact the description of polarization in those hydrogen bond networks compared to the weaker polarization effects present in other crystal structures.

Overall, given the fairly routine nature of the intermolecular interactions among these crystal structures, the author suspects that the 25% EXX hybrid functional is probably sufficient to describe the intermolecular interactions when paired with the DLPNO-CCSD(T1) correction that addresses the intramolecular delocalization error. Previous benchmark studies of intermolecular interactions and crystal lattice energies found that for systems without major delocalization error issues, B86bPBE-25-XDM performed as well as or slightly better than B86bPBE-50-XDM.[50,51] Therefore, the remaining discussion focuses on the B86bPBE-25-XDM + ΔDLPNO-CCSD(T1) landscape shown in
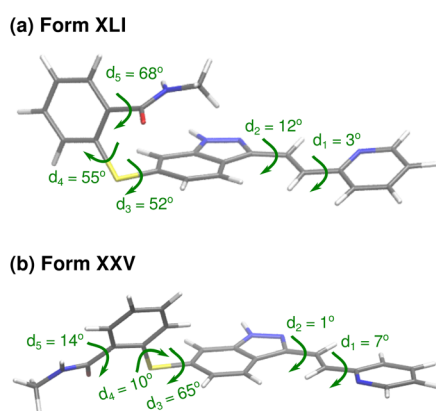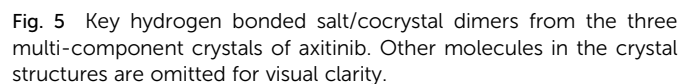

Fig. 3 The (a) form XLI molecular conformation is more folded than (b) the more typical form XXV one, leading to considerable deviations from planarity in dihedral angles d_4 and d_5 in particular.

Fig. 2c. This landscape reveals that many of the lowest-energy structures are already known experimentally (forms XLI, XXV, VI and IV). This is reminiscent of the earlier CSP study of the highly polymorphic molecule ROY, where careful electronic structure treatment produced a 0 K landscape in which 9 of the 12 lowest-energy predicted structures corresponded to experimentally known polymorphs.[81] It emphasizes that even challenging crystal structures are increasingly becoming predictable (even if difficult examples remain[43,44,84]).

Among the as-yet undiscovered low-energy predicted structures in Fig. 2c, several contain partial packing motifs that are analogous to those found the more stable forms XXV or VI. Accordingly, one might expect forms XXV and VI would crystallize preferentially over the higher-energy alternative structures. On the other hand, both the rank 2 and 5 structures in Fig. 2c exhibit distinct packing motifs not found in the other experimental structures (Section S2.4). The intramolecular conformations adopted in these two candidates appear to be readily accessible energetically as well, suggesting they might be crystallizable. It would be interesting to compute their room-temperature free energies, though that lies outside the scope of the current study. It is also worth noting that a sixth axitinib polymorph with undisclosed crystal structure is reported in the patent literature,[112] raising the question of whether one of these two candidate structures (or any others on the CSP landscape) might correspond to the newer polymorph.

In summary, the first successful 0 K crystal structure prediction of the challenging cancer drug axitinib has been achieved. The predicted crystal energy landscape suggests that the most stable polymorph of axitinib, form XLI, was indeed eventually found in the lab, thanks to extensive experimental efforts and a couple doses of serendipity. Several other lowest-energy polymorphs have also been found experimentally, and further investigation into the structure of the purported sixth polymorph would be worthwhile.

Methodologically, the central difficulty in the DFT-driven crystal structure prediction of axitinib lay in addressing the intramolecular delocalization error *via* higher-level intramolecular conformational corrections, hybrid density functionals, or a combination of both. Fig. 4 provides some insight into the likely uncertainty associated with the 0 K lattice energy predictions. Importantly, combining any of the GGA or hybrid DFT functionals tested here with intramolecular corrections increases the consistency among many of the predicted polymorph lattice energy differences, enhancing overall confidence in the energy landscape. This observation reflects how the most challenging aspect of the axitinib conformational polymorph energy ranking stems from difficulties in modeling the intramolecular conformational energies. In contrast, the next section focuses on multi-component crystals of axitinib for which the intermolecular errors become key.

3.2 Salts and cocrystals of axitinib

Salt and cocrystal forms of pharmaceuticals are often employed to improve their solubility and bioavailability. Numerous axitinib salts with strong acids were reported early on.[113] More recently, Ren and co-workers[3] sought to make neutral cocrystals of axitinib with good dissolution profiles by cocrystallizing it with weaker carboxylic acids. They found that axitinib forms 1 : 1 neutral cocrystals with both suberic and *trans*-cinnamic acids. With fumaric acid, however, axitinib forms a 2 : 3 salt cocrystal in which one fumaric acid donates one proton to each axitinib molecule, forming a salt, while the third fumaric acid remains neutral.

Whether an acid–base multi-component crystal remains as a neutral cocrystal, converts to a salt, or adopts any intermediate state in the salt-cocrystal continuum depends on the pK_a differences between the acidic and basic sites. Characterizing the behavior is important due to the regulatory distinctions between the salt and neutral forms. Unfortunately, determining proton positions from X-ray diffraction can be challenging, and the salt *vs.* cocrystal character is often inferred indirectly from heavy-atom positions (by examining C–O bond lengths in a carboxylic acid, for example). Alternatively, techniques such as solid-state nuclear magnetic resonance or neutron diffraction can resolve protonation states more clearly.

In principle, computational modeling ought to be able to help characterize the salt/cocrystal behaviors. However, in addition to the potential importance of nuclear quantum effects for the acidic proton,[114] density-driven delocalization error in GGA functionals tends to over-stabilize salt forms relative to neutral ones.[115] While careful computational protocols may offer a pragmatic strategy for converging GGA structure optimizations to cocrystal forms,[116] the most robust solution to date has been to employ hybrid density functionals, especially those containing ~50% exact exchange, to reduce the delocalization error.[85,90] Here, we show how the intramolecular correction and hybrid functional strategies can both be used to predict the correct proton transfer behavior in these acid–base crystals of axitinib.

Fig. 5 shows the key hydrogen bond in each of the three axitinib + carboxylic acid crystals studied by Ren *et al.*, and Fig. 6 plots potential energy curves computed for the key proton
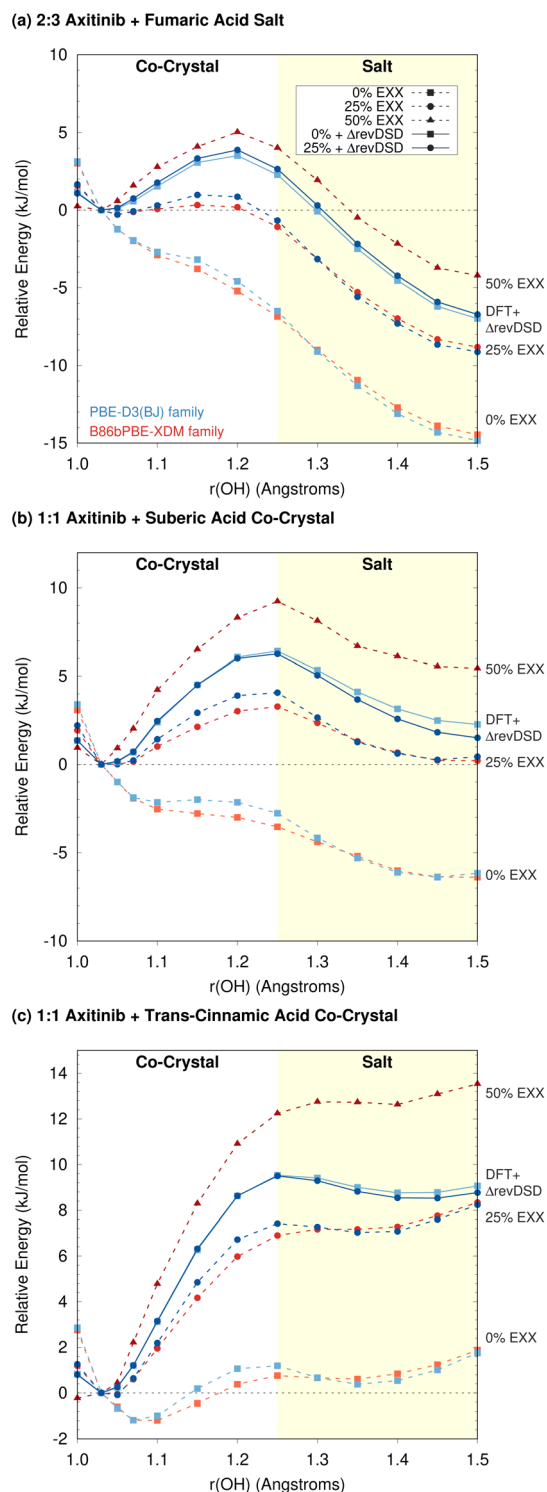

Fig. 5 Key hydrogen bonded salt/cocrystal dimers from the three multi-component crystals of axitinib. Other molecules in the crystal structures are omitted for visual clarity.

Fig. 6 Solid-state potential energy scans for cocrystal to salt proton transfer in multi-component crystals of axitinib with (a) fumaric acid, (b) suberic acid, and (c) *trans*-cinnamic acid. Dotted red curves correspond to B86bPBE-XDM type functionals and blue to PBE-D3(BJ) type functionals with 0%, 25%, or 50% exact exchange (EXX). The solid blue curves represent the PBE-D3(BJ) functionals after applying a revDSD-PBEP86-D4 correction.

transfer coordinate in each crystal. These curves were obtained by performing a series of constrained optimizations in the periodic crystals at the B86bPBE-XDM level of theory in which the important carboxylic acid O–H distances in the O–H···N hydrogen bonds were frozen at selected values while all other atomic positions were relaxed. Single-point energies were then computed on each optimized geometry using a variety of models.

Consider first functionals in the B86bPBE-XDM family. As expected, delocalization error causes the B86bPBE-XDM GGA (0% EXX) to substantially over-stabilize the salt forms, leading to barrierless conversion from neutral to salt for both the fumaric and suberic acid cases (Fig. 6a and b). While this prediction is qualitatively correct for fumaric acid, it contradicts experiment for suberic acid. For *trans*-cinnamic acid, B86bPBE-XDM correctly predicts a cocrystal, though the salt form lies only 2 kJ mol^{-1} higher in energy. Switching to the hybrid analogs with 25% (B86bPBE-25-XDM) or 50% (B86bPBE-50-XDM) exact exchange increasingly destabilizes the salt form, and now the models correctly predict cocrystals for suberic and *trans*-cinnamic acids and a fumarate salt at 50% EXX. The PBE-D3(BJ) GGA and PBE0-D3(BJ) hybrid (25% EXX) functionals produce potential energy curves that are very similar to the B86bPBE-XDM family ones as well.

One can also apply the intramolecular correction approach of eqn (2) to these cases. To provide a consistent treatment across the proton transfer potential energy curve, the dimers shown in Fig. 5 are treated a single "intramolecular" correction unit. Moreover, the intramolecular correction calculations are embedded in a chloroform polarizable continuum model to better describe the charged salt crystals (see Computational methods section for details). Applying a revDSD-PBEP86-D4 double-hybrid density functional correction to the PBE-D3(BJ) family of functionals leads to very interesting results: first, despite the large differences between the uncorrected PBE-D3(BJ) and PBE0-D3(BJ) potential energy curves, the two curves become nearly identical after applying the revDSD-PBEP86-D4 correction. This highlights how the key modeling challenge lies in describing the specific acid–base behavior, rather then the rest of the crystal.

Second, the experimentally-reported salt/cocrystal behavior is reproduced for all three axitinib carboxylic acid crystals after the correction has been applied. Given the small proton-transfer barriers (<10 kJ mol^{-1}), interconversion between neutral and salt forms should be possible at room temperature in all three cases. Therefore, the room-temperature populations of salt and cocrystal forms were estimated for each crystal based on the relative revDSD-PBE0-D3(BJ)-corrected PBE0-D3(BJ) energies. Whereas the fumaric acid crystal contains >90% salt and the *trans*-cinnamic acid one has >95% cocrystal, the suberic acid crystal is predicted to occur as a mixture of roughly two-thirds cocrystal and one-third salt.

Third, the revDSD-PBEP86-D4-corrected energies in Fig. 6 lie intermediate between those obtained from the uncorrected 25% and 50% EXX B86bPBE-XDM functionals, similar to what was found for the pure axitinib crystal energy landscapes in

Fig. 4. While these energies cannot currently be confirmed experimentally, they are suggestive that the optimal fraction of exact exchange in these systems might lie somewhere in between 25% and 50%.

Overall, these results demonstrate once again the complementarity of hybrid DFT and intramolecular corrections for developing more robust insights into challenging crystal behaviors. With these models, one can successfully discriminate the subtle energetic differences among a set of related coformers that lead to distinct salt *versus* cocrystal behaviors. It should also be noted that the combination of constrained GGA-level optimizations plus refined single-point energies provide a useful and relatively inexpensive computational protocol for understanding these systems without the need for expensive hybrid DFT geometry optimizations.[85,90]

4   Conclusions

Axitinib highlights the difficult experimental and computational challenges researchers can face when developing solid-state drug formulations. On the computational side, density-driven delocalization error causes GGA-type density functionals to incorrectly predict the most stable crystal structures of axitinib and to overstabilize the salt forms for some of its multi-component crystals. This work shows how these problems in axitinib can be overcome through the use of hybrid DFT functionals and/or intramolecular energy corrections. As a result, the present study produced the first 0 K crystal energy landscape of axitinib that appears to be consistent with experiment. The CSP reveals that most of the low-energy predicted structures are known experimentally and that form XLI is indeed the global minimum structure at 0 K. The landscape also contains two low-energy structures with unique packing motifs that have not yet been found in the lab. Performing free energy calculations on these structures to determine their relevance near ambient conditions and investigating possible connections to the reported sixth polymorph of axitinib would be worthwhile. This study further demonstrates that the same techniques can correctly characterize the structural character of acid–base multi-component crystals of axitinib, which can have important regulatory implications. The methods clearly distinguished the axitinib salt formed with fumaric acid from the cocrystals involving suberic or *trans*-cinnamic acids.

From a methodological standpoint, this study addresses a potential source of confusion about the relationships between different methods found in the literature by demonstrating the complementarity of the intramolecular correction and hybrid DFT approaches. DFT conformational energy errors, whether arising from delocalization error[75,79,82,83] or other density functional limitations[84] are common in conformational polymorphs, and intramolecular corrections provide a pragmatic solution. This paper also demonstrates, for the first time, how local energy corrections computed on the key acid–base dimer can resolve the long-standing dilemma that GGA density functionals artificially favor charged salt forms over their neutral cocrystal counterparts. Such corrections can potentially be evaluated at considerably lower cost than periodic hybrid DFT functionals.

On the other hand, the inequivalent treatment of intra- and intermolecular interactions provided by the intramolecular correction approach can potentially introduce biases into the computed polymorph energy differences. For example, crystal landscapes where there is a competition between intra- and intermolecular hydrogen bonding can potentially be challenging with this approach.[82]

Periodic hybrid density functional methods avoid this issue by applying a single, unified model chemistry for all interactions in the crystal. However, such functionals are not always sufficiently accurate,[75] and the optimal admixture of exact exchange can be ambiguous.[89] In this context, the increased robustness of the crystal energies obtained from the simultaneous combination of hybrid DFT and intramolecular corrections is encouraging. While earlier studies in the literature have combined both approaches,[66,75] the increased robustness of the predictions that results has not been widely recognized. The combination of these and many other recent advances are pushing us closer to the day when crystal structures will truly be predictable.

Author contributions

G. J. O. B. conceived the study, collected and analyzed the data, and wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI). Supplementary information: DFT-optimized crystal structures, additional crystal energy landscape data, crystal structure images, and analysis of polarizable continuum model embedding. See DOI: **https://doi.org/10.1039/d5sc06271c**.

Acknowledgements

Notes and references

1 B. P. Chekal, A. M. Campeta, Y. A. Abramov, N. Feeder, P. P. Glynn, R. W. McLaughlin, P. A. Meenan and R. A. Singer, *Org. Process Res. Dev.*, 2009, **13**, 1327–1337.

2 A. M. Campeta, B. P. Chekal, Y. A. Abramov, P. A. Meenan, M. J. Henson, B. Shi, R. A. Singer and K. R. Horspool, *J. Pharm. Sci.*, 2010, **99**, 3874–3886.

3 B.-Y. Ren, X.-L. Dai, J. Wang, C. Wu, T.-B. Lu and J.-M. Chen, *CrystEngComm*, 2021, **23**, 5504–5515.

4 Y. Pan, T. Xiao, Y. Wang, Z. Pan, S. Du and F. Xue, *Molecules*, 2024, **29**, 4696.

5 J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter and J. Morris, *Pharm. Res.*, 2001, **18**, 859–866.

6 S. R. Chemburkar, J. Bauer, K. Deming, H. Spiwek, K. Patel, J. Morris, R. Henry, S. Spanton, W. Dziki, W. Porter, J. Quick, P. Bauer, J. Donaubauer, B. A. Narayanan, M. Soldani, D. Riley and K. Mcfarland, *Org. Process Res. Dev.*, 2000, **4**, 413–417.

7 I. B. Rietveld and R. Céolin, *J. Pharm. Sci.*, 2015, **104**, 4117–4122.

8 M. Mortazavi, J. Hoja, L. Aerts, L. Quéré, J. van de Streek, M. A. Neumann and A. Tkatchenko, *Commun. Chem.*, 2019, **2**, 70.

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

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

11 R. M. Bhardwaj, J. A. McMahon, J. Nyman, L. S. Price, S. Konar, I. D. H. Oswald, C. R. Pulham, S. L. Price and S. M. Reutzel-Edens, *J. Am. Chem. Soc.*, 2019, **141**, 13887–13897.

12 S. L. Price and S. M. Reutzel-Edens, *Drug Discovery Today*, 2016, **21**, 912–923.

13 S. L. Price, D. E. Braun and S. M. Reutzel-Edens, *Chem. Commun.*, 2016, **52**, 7065–7077.

14 M. A. Neumann and J. van de Streek, *Faraday Discuss.*, 2018, **211**, 441–458.

15 M. A. Neumann, J. van de Streek, F. P. A. Fabbiani, P. Hidber and O. Grassmann, *Nat. Commun.*, 2015, **6**, 7793.

16 E. Salager, G. M. Day, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, *J. Am. Chem. Soc.*, 2010, **132**, 2564–2566.

17 S. Askin, J. K. Cockcroft, L. S. Price, A. D. Gonçalves, M. Zhao, D. A. Tocher, G. R. Williams, S. Gaisford and D. Q. M. Craig, *Cryst. Growth Des.*, 2019, **19**, 2751–2757.

18 H. Polyzois, R. Guo, V. K. Srirambhatla, M. Warzecha, E. Prasad, A. Turner, G. W. Halbert, P. Keating, S. L. Price and A. J. Florence, *Cryst. Growth Des.*, 2022, **22**, 4146–4156.

19 M. Baias, C. M. Widdifield, J.-N. Dumez, H. P. G. Thompson, T. G. Cooper, E. Salager, S. Bassil, R. S. Stein, A. Lesage, G. M. Day and L. Emsley, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8069.

20 M. Balodis, M. Cordova, A. Hofstetter, G. M. Day and L. Emsley, *J. Am. Chem. Soc.*, 2022, **144**, 7215–7223.

21 G. R. Woollam, M. A. Neumann, T. Wagner and R. J. Davey, *Faraday Discuss.*, 2018, **211**, 209–234.

22 M. Habgood, R. Grau-Crespo and S. L. Price, *Phys. Chem. Chem. Phys.*, 2011, **13**, 9590–9600.

23 A. Torrisi, C. K. Leech, K. Shankland, W. I. F. David, R. M. Ibberson, J. Benet-Buchholz, R. Boese, M. Leslie, C. R. A. Catlow and S. L. Price, *J. Phys. Chem. B*, 2008, **112**, 3746–3758.

24 R. C. B. Copley, S. A. Barnett, P. G. Karamertzanis, K. D. M. Harris, B. M. Kariuki, M. Xu, E. A. Nickels, R. W. Lancaster and S. L. Price, *Cryst. Growth Des.*, 2008, **8**, 3474–3481.

25 D. E. Braun, J. A. McMahon, L. H. Koztecki, S. L. Price and S. M. Reutzel-Edens, *Cryst. Growth Des.*, 2014, **14**, 2056–2072.

26 D. E. Braun, P. G. Karamertzanis and S. L. Price, *Chem. Commun.*, 2011, **47**, 5443–5445.

27 D. E. Braun, R. M. Bhardwaj, A. J. Florence, D. A. Tocher and S. L. Price, *Cryst. Growth Des.*, 2013, **13**, 19–23.

28 D. E. Braun, H. Oberacher, K. Arnhard, M. Orlova and U. J. Griesser, *CrystEngComm*, 2016, **18**, 4053–4067.

29 E. C. Dybeck, A. Thiel, M. J. Schnieders, F. C. Pickard, G. P. Wood, J. F. Krzyaniak and B. C. Hancock, *Cryst. Growth Des.*, 2023, **23**, 142–167.

30 A. J. Cruz-Cabeza, G. M. Day and W. Jones, *Chem.–Eur. J.*, 2008, **14**, 8830–8836.

31 A. J. Cruz-Cabeza, S. Karki, L. Fábián, T. Friić, G. M. Day, W. Jones, T. Friščić, G. M. Day, W. Jones, T. Friscić, G. M. Day, W. Jones, T. Friić, G. M. Day, W. Jones, T. Friscić, G. M. Day, W. Jones, T. Friić, G. M. Day, W. Jones, T. Friščić, G. M. Day and W. Jones, *Chem. Commun.*, 2010, **46**, 2224–2226.

32 C. R. Taylor and G. M. Day, *Cryst. Growth Des.*, 2018, **18**, 892–904.

33 N. Issa, P. G. Karamertzanis, G. W. A. Welch and S. L. Price, *Cryst. Growth Des.*, 2009, **9**, 442–453.

34 G. Sun, Y. Jin, S. Li, Z. Yang, B. Shi, C. Chang and Y. A. Abramov, *J. Phys. Chem. Lett.*, 2020, **11**, 8832–8838.

35 I. J. Sugden, D. E. Braun, D. H. Bowskill, C. S. Adjiman and C. C. Pantelides, *Cryst. Growth Des.*, 2022, **22**, 4513–4527.

36 Y. A. Abramov, L. Iuzzolino, Y. Jin, G. York, C.-h. Chen, C. S. Shultz, Z. Yang, C. Chang, B. Shi, T. Zhou, C. Greenwell, S. Sekharan and A. Y. Lee, *Mol. Pharmaceutics*, 2023, **20**, 3380–3392.

37 Y. A. Abramov, *J. Phys. Chem. A*, 2011, **115**, 12809–12817.

38 D. Lupyan, Y. A. Abramov and W. Sherman, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 1195–1205.

39 M. Vasileiadis, C. C. Pantelides and C. S. Adjiman, *Chem. Eng. Sci.*, 2015, **121**, 60–76.

40 G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallogr., Sect. B:Struct. Sci.*, 2009, **65**, 107–125.

41 M. A. Neumann, F. J. J. Leusen and J. Kendrick, *Angew. Chem., Int. Ed.*, 2008, **47**, 2427–2430.

42 A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma,

J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu and C. R. Groom, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**, 439–459.

43 L. M. Hunnisett, J. Nyman, N. Francia, N. S. Abraham, C. S. Adjiman, S. Aitipamula, T. Alkhidir, M. Almehairbi, A. Anelli, D. M. Anstine, J. E. Anthony, J. E. Arnold, F. Bahrami, M. A. Bellucci, R. M. Bhardwaj, I. Bier, J. A. Bis, A. D. Boese, D. H. Bowskill, J. Bramley, J. G. Brandenburg, D. E. Braun, P. W. V. Butler, J. Cadden, S. Carino, E. J. Chan, C. Chang, B. Cheng, S. M. Clarke, S. J. Coles, R. I. Cooper, R. Couch, R. Cuadrado, T. Darden, G. M. Day, H. Dietrich, Y. Ding, A. DiPasquale, B. Dhokale, B. P. van Eijck, M. R. J. Elsegood, D. Firaha, W. Fu, K. Fukuzawa, J. Glover, H. Goto, C. Greenwell, R. Guo, J. Harter, J. Helfferich, D. W. M. Hofmann, J. Hoja, J. Hone, R. Hong, G. Hutchison, Y. Ikabata, O. Isayev, O. Ishaque, V. Jain, Y. Jin, A. Jing, E. R. Johnson, I. Jones, K. V. J. Jose, E. A. Kabova, A. Keates, P. F. Kelly, D. Khakimov, S. Konstantinopoulos, L. N. Kuleshova, H. Li, X. Lin, A. List, C. Liu, Y. M. Liu, Z. Liu, Z.-p. Liu, J. W. Lubach, N. Marom, A. A. Maryewski, H. Matsui, A. Mattei, R. A. Mayo, J. W. Melkumov, S. Mohamed, Z. Momenzadeh Abardeh, H. S. Muddana, N. Nakayama, K. S. Nayal, M. A. Neumann, R. Nikhar, S. Obata, D. O'Connor, A. R. Oganov, K. Okuwaki, A. Otero-de-la Roza, C. C. Pantelides, S. Parkin, C. J. Pickard, L. Pilia, T. Pivina, R. Podeszwa, A. J. A. Price, L. S. Price, S. L. Price, M. R. Probert, A. Pulido, G. R. Ramteke, A. U. Rehman, S. M. Reutzel-Edens, J. Rogal, M. J. Ross, A. F. Rumson, G. Sadiq, Z. M. Saeed, A. Salimi, M. Salvalaglio, L. Sanders de Almada, K. Sasikumar, S. Sekharan, C. Shang, K. Shankland, K. Shinohara, B. Shi, X. Shi, A. G. Skillman, H. Song, N. Strasser, J. van de Streek, I. J. Sugden, G. Sun, K. Szalewicz, B. I. Tan, L. Tan, F. Tarczynski, C. R. Taylor, A. Tkatchenko, R. Tom, M. E. Tuckerman, Y. Utsumi, L. Vogt-Maranto, J. Weatherston, L. J. Wilkinson, R. D. Willacy, L. Wojtas, G. R. Woollam, Z. Yang, E. Yonemochi, X. Yue, Q. Zeng, Y. Zhang, T. Zhou, Y. Zhou, R. Zubatyuk and J. C. Cole, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2024, **80**, 517–547.

44 L. M. Hunnisett, N. S. Abraham, S. Aitipamula, T. Alkhidir, M. Almehairbi, A. Anelli, D. M. Anstine, J. E. Anthony, J. E. Arnold, F. Bahrami, M. A. Bellucci, G. J. O. Beran, R. M. Bhardwaj, R. Bianco, I. Bier, J. A. Bis, A. D. Boese, J. Bramley, D. E. Braun, P. W. V. Butler, J. Cadden, S. Carino, C. Červinka, E. J. Chan, C. Chang, S. M. Clarke, S. J. Coles, C. J. Cook, R. I. Cooper, G. M. Day, W. Deng, H. Dietrich, Y. Ding, A. DiPasquale, B. Dhokale, B. P. van Eijck, M. R. J. Elsegood, D. Firaha, N. Francia, W. Fu, K. Fukuzawa, N. Galanakis, H. Goto, C. Greenwell, R. Guo, J. Harter, J. Helfferich, J. Hoja, J. Hone, R. Hong, M. Hušák, Y. Ikabata, O. Isayev, O. Ishaque, V. Jain, Y. Jin, A. Jing, E. Johnson, I. Jones, K. V. J. Jose, E. A. Kabova, A. Keates, P. F. Kelly, J. Klimeš, V. Kostková, H. Li, X. Lin, A. List, C. Liu, Y. M. Liu, Z. Liu, I. Lončarić, J. W. Lubach, J. Ludík, A. A. M. andk Noa Marom, H. Matsui, A. Mattei, R. A. Mayo, J. Melkumov, B. Mladineo, S. Mohamed, Z. S. M. Abardeh, H. Muddana, N. Nakayama, K. S. Nayal, M. A. Neumann, R. Nikhar, J. Nyman, S. Obata, D. O'Connor, A. R. Oganov, K. Okuwaki, A. O. de-la Roza, S. Parkin, A. Parunov, R. Podeszwa, A. J. A. Price, L. S. Price, S. L. Price, M. R. Probert, A. Pulido, G. R. Ramteke, A. U. Rehman, S. M. Reutzel-Edens, J. Rogal, M. J. Ross, A. F. Rumson, G. Sadiq, Z. M. Saeed, A. Salimi, M. Salvalaglio, K. Sasikumar, S. Sekharan, K. Shankland, B. Shi, X. Shi, K. Shinohara, H. Song, N. Strasser, J. van de Streek, I. J. Sugden, G. Sun, K. Szalewicz, L. Tan, K. Tang, F. Tarczynski, C. R. Taylor, A. Tkatchenko, R. Tom, P. Touš, M. E. Tuckerman, P. A. Unzeta, Y. Utsumi, L. Vogt-Maranto, J. Weatherston, L. A. Wilkinson, R. D. Willacy, L. Wojtas, W. Wood, G. R. Woollam, Y. Yang, Z. Yang, E. Yonemochi, X. Yue, Q. Zeng, T. Zhou, Y. Zhou, R. Zubatyuk and J. C. Cole, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2024, **80**, 548–574.

45 S. R. Whittleton, A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 441–450.

46 S. R. Whittleton, A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 5332–5342.

47 L. Kronik and A. Tkatchenko, *Acc. Chem. Res.*, 2014, **47**, 3208–3216.

48 N. Marom, R. A. DiStasio, V. Atalla, S. Levchenko, A. M. Reilly, J. R. Chelikowsky, L. Leiserowitz and A. Tkatchenko, *Angew. Chem., Int. Ed.*, 2013, **52**, 6629–6632.

49 J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio and A. Tkatchenko, *Sci. Adv.*, 2019, **5**, eaau3338.

50 A. J. Price, A. Otero-De-La-Roza and E. R. Johnson, *Chem. Sci.*, 2022, **14**, 1252–1262.

51 A. J. Price, R. A. Mayo, A. Otero-De-la Roza and E. R. Johnson, *CrystEngComm*, 2023, **25**, 953–960.

52 O. A. Loboda, G. A. Dolgonos and A. D. Boese, *J. Chem. Phys.*, 2018, **149**, 124104.

53 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

54 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

55 E. Caldeweyher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2017, **147**, 034112.

56 R. A. DiStasio, O. A. von Lilienfeld and A. Tkatchenko, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14791–14795.

57 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.

58 A. Ambrosetti, A. M. Reilly, R. A. DiStasio and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.

59 A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **136**, 174109.

60 R. J. Clements, J. Dickman, J. Johal, J. Martin, J. Glover and G. M. Day, *MRS Bull.*, 2022, **47**, 1054–1062.

61 V. Gharakhanyan, Y. Yang, L. Barroso-Luque, M. Shuaibi, D. S. Levine, K. Michel, V. Bernat, M. Dzamba, X. Fu, M. Gao, X. Liu, K. Noori, L. J. Purvis, T. Rao, B. M. Wood, A. Rizvi, M. Uyttendaele, A. J. Ouderkirk, C. Daraio, C. L. Zitnick, A. Boromand, N. Marom, Z. W. Ulissi and A. Sriram, *arXiv*, 2025, DOI: 10.48550/arXiv.2508.02641.

62 K. S. Nayal, D. O'Connor, R. Zubatyuk, D. M. Anstine, Y. Yang, R. Tom, W. Deng, K. Tang, N. Marom and O. Isayev, *Cryst. Growth Des.*, 2025, **25**, 9092–9106.

63 P. W. V. Butler, R. Hafizi and G. M. Day, *J. Phys. Chem. A*, 2024, **128**, 945–957.

64 V. Kapil and E. A. Engel, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, 1–8.

65 S. Wengert, G. Csányi, K. Reuter and J. T. Margraf, *J. Chem. Theory Comput.*, 2022, **18**, 4586–4593.

66 D. Firaha, Y. M. Liu, J. van de Streek, K. Sasikumar, H. Dietrich, J. Helfferich, L. Aerts, D. E. Braun, A. Broo, A. G. DiPasquale, A. Y. Lee, S. Le Meur, S. O. Nilsson Lill, W. J. Lunsmann, A. Mattei, P. Muglia, O. D. Putra, M. Raoui, S. M. Reutzel-Edens, S. Rome, A. Y. Sheikh, A. Tkatchenko, G. R. Woollam and M. A. Neumann, *Nature*, 2023, **623**, 324–328.

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

68 G. Sun, X. Liu, Y. A. Abramov, S. O. Nilsson Lill, C. Chang, V. Burger and A. Broo, *Cryst. Growth Des.*, 2021, **21**, 1972–1983.

69 M. Yang, E. Dybeck, G. Sun, C. Peng, B. Samas, V. M. Burger, Q. Zeng, Y. Jin, M. A. Bellucci, Y. Liu, P. Zhang, J. Ma, Y. A. Jiang, B. C. Hancock, S. Wen and G. P. F. Wood, *Cryst. Growth Des.*, 2020, **20**, 5211–5224.

70 N. P. Schieber, E. C. Dybeck and M. R. Shirts, *J. Chem. Phys.*, 2018, **148**, 144104.

71 C. Červinka and G. J. O. Beran, *Chem. Sci.*, 2018, **9**, 4622–4629.

72 C. Cook, J. L. McKinley and G. J. O. Beran, *J. Chem. Phys.*, 2021, **154**, 134109.

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

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

75 C. J. Perry, S. A. Ramos, M. C. Phelps, L. J. Mueller and G. J. O. Beran, *J. Am. Chem. Soc.*, 2025, **147**, 26865–26876.

76 J. Kendrick, G. A. Stephenson, M. A. Neumann and F. J. J. Leusen, *Cryst. Growth Des.*, 2013, **13**, 581–589.

77 P. Zhang, G. P. F. Wood, J. Ma, M. Yang, Y. Liu, G. Sun, Y. A. Jiang, B. C. Hancock and S. Wen, *Cryst. Growth Des.*, 2018, **18**, 6891–6900.

78 K. R. Bryenton, A. A. Adeleke, S. G. Dale and E. R. Johnson, *WIREs Computational Molecular Science*, 2023, **13**, e1631.

79 C. Greenwell and G. J. O. Beran, *Cryst. Growth Des.*, 2020, **20**, 4875–4881.

80 C. Greenwell and G. J. O. Beran, *J. Mater. Chem. C*, 2021, **9**, 2848–2857.

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

82 G. J. O. Beran, S. E. Wright, C. Greenwell and A. J. Cruz-Cabeza, *J. Chem. Phys.*, 2022, **156**, 104112.

83 S. K. Rai, A. Gunnam, G. J. O. Beran, J. A. Kaduk and A. K. Nangia, *Cryst. Growth Des.*, 2023, **23**, 1179–1188.

84 G. J. O. Beran, C. C. Cook and P. A. Unzueta, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2024, **80**, 606–619.

85 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.

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

87 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

88 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158.

89 R. A. Mayo, A. J. A. Price, A. Otero-de-la Roza and E. R. Johnson, *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.*, 2024, **80**, 595–605.

90 K. M. Bal and A. Collas, *CrystEngComm*, 2024, **26**, 6765–6773.

91 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.

92 G. J. O. Beran, C. Greenwell, C. Cook and J. Řezáč, *Acc. Chem. Res.*, 2023, **56**, 3525–3534.

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

94 C. J. Cook, C. J. Perry and G. J. O. Beran, *J. Phys. Chem. Lett.*, 2023, **14**, 6823–6831.

95 J. Weatherston, M. R. Probert and M. J. Hall, *J. Am. Chem. Soc.*, 2025, **147**, 11949–11954.

96 M. Vasileiadis, A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman and C. C. Pantelides, *Acta Crystallogr., Sect. B:Struct. Sci.*, 2012, **68**, 677–685.

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

98 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

99 P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car,

C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni, *J. Phys.:Condens. Matter*, 2017, **29**, 465901.

100 V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, *Comput. Phys. Commun.*, 2009, **180**, 2175–2196.

101 X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter and M. Scheffler, *New J. Phys.*, 2012, **14**, 053020.

102 C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.

103 Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo and F. Neese, *J. Chem. Phys.*, 2018, **148**, 011101.

104 F. Neese, *WIREs Computational Molecular Science*, 2012, **2**, 73–78.

105 A. G. Császár, W. D. Allen and H. F. Schaefer, *J. Chem. Phys.*, 1998, **108**, 9751–9764.

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

107 G. Santra, N. Sylvetsky and J. M. L. Martin, *J. Phys. Chem. A*, 2019, **123**, 5129–5143.

108 J. Nyman and G. M. Day, *CrystEngComm*, 2015, **17**, 5154–5165.

109 A. J. Cruz-Cabeza, S. M. Reutzel-Edens and J. Bernstein, *Chem. Soc. Rev.*, 2015, **44**, 8619–8635.

110 E. Sim, S. Song, S. Vuckovic and K. Burke, *J. Am. Chem. Soc.*, 2022, **144**, 6625–6639.

111 S. Nam, E. Cho, E. Sim and K. Burke, *J. Phys. Chem. Lett.*, 2021, **12**, 2796–2804.

112 B. Pipal, V. Singh, and A. Chaturvedi, WO Pat., 2016/178150A1, 2016.

113 Q. Ye, R. M. Hart, R. Kania, M. Ouellette, Z. P. Wu and S. E. Zook, *US Pat.*, 20060094763A1, 2006.

114 J. R. Štoček, O. Socha, I. Císařová, T. Slanina and M. Dračínský, *J. Am. Chem. Soc.*, 2022, **144**, 7111–7116.

115 A. Otero-de-la Roza, G. A. DiLabio and E. R. Johnson, *J. Chem. Theory Comput.*, 2016, **12**, 3160–3175.

116 Y. A. Abramov and J. Wang, *Cryst. Growth Des.*, 2024, **24**, 4017–4024.