Chemical Science

ROYAL SOCIETY OF CHEMISTRY

PERSPECTIVE

Cite this: *Chem. Sci.*, 2023, **14**, 13290

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 27th July 2023
Accepted 2nd November 2023

DOI: 10.1039/d3sc03903j

rsc.li/chemical-science

Frontiers of molecular crystal structure prediction for pharmaceuticals and functional organic materials

Gregory J. O. Beran

The reliability of organic molecular crystal structure prediction has improved tremendously in recent years. Crystal structure predictions for small, mostly rigid molecules are quickly becoming routine. Structure predictions for larger, highly flexible molecules are more challenging, but their crystal structures can also now be predicted with increasing rates of success. These advances are ushering in a new era where crystal structure prediction drives the experimental discovery of new solid forms. After briefly discussing the computational methods that enable successful crystal structure prediction, this perspective presents case studies from the literature that demonstrate how state-of-the-art crystal structure prediction can transform how scientists approach problems involving the organic solid state. Applications to pharmaceuticals, porous organic materials, photomechanical crystals, organic semi-conductors, and nuclear magnetic resonance crystallography are included. Finally, efforts to improve our understanding of which predicted crystal structures can actually be produced experimentally and other outstanding challenges are discussed.

1 Introduction

Molecular organic crystals occur across many areas of chemistry. The majority of small-molecule pharmaceuticals are administered in crystalline form.[1] Molecular crystals are key components of fertilizers,[2,3] pesticides,[4,5] and pigments.[6] They can function as field effect transistors, light-emitting diodes, and photovoltaic cells.[7] Porous organic crystals can perform gas storage and separations.[8] Crystalline order can enable highly-selective solid-state syntheses,[9] while crystalline phase transitions and solid-state chemical reactions create the basis for new mechanically-responsive "dynamic" materials.[10–12]

The properties and functions of these crystals, including color, stability, solubility, carrier mobility, *etc.*, often depend strongly on the crystal packing. Notably, about half of all organic molecules are thought to exhibit polymorphism,[13] or the ability to adopt multiple distinct crystal packing motifs, and this creates both challenges and opportunities when working with organic materials. While the crystallization of the "wrong" polymorph can hinder the bioavailability of a pharmaceutical and force its recall, for example, the possibility to tailor crystal packing to achieve desired physical properties is alluring. Unfortunately, experimental polymorph control can be difficult, and even seemingly minor changes in the crystallization conditions or to the molecular structure can alter the crystal structure significantly. The choice of solvent system, heat, pressure, or time can similarly transform a system from one polymorph to another.

For these reasons, developing new organic materials often requires an understanding of the landscape of crystal structures that can occur for the species of interest. Given the difficulties in ensuring that all important crystal forms have been discovered experimentally, researchers have long sought the complementary ability to predict crystal polymorphs theoretically. Seventy years ago, science fiction author Robert Heinlein dreamed of a future when "mathematical chemists will design new materials, predict their properties, and tell engineers how

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

Gregory J. O. Beran

Gregory Beran earned his PhD at the University of California Berkeley in 2005, working with Prof. Martin Head-Gordon, and performed postdoctoral research with Prof. William H. Green at the Massachusetts Institute of Technology 2005–2007. He started his independent career at the University of California Riverside in 2007, where he is currently a full professor. His research focuses on the development and application of electronic structure methods for non-covalent interactions and organic crystal structure prediction.

to make them—without entering the laboratory."[14] Progress toward this goal remained slow for decades, and in 1998 Maddox famously referred to the difficulty in predicting crystal structures as "one of the continuing scandals in the physical sciences."[15] Since then, however, crystal structure prediction (CSP) has transformed from scandal to reality, and Heinlein's vision is finally now being realized for organic crystals.

Successful predictions continue to mount in recent Blind Tests of Crystal Structure Prediction which have been held every few years since 1999.[16–21] The results of the most recent 7th Blind Test will be published in the near future. The scope of successful predictions has progressed from small, rigid molecules to larger pharmaceutical-sized molecules with conformational flexibility and/or disorder, and from single-component crystals to multi-component hydrates, solvates, co-crystals, and salts. Even the definition of what constitutes a "successful" crystal structure prediction has evolved to become more stringent over time. In the first Blind Test, for example, simply finding the experimental crystal structure during the search procedure was considered a partial success, even if the energy model ranked it poorly. Today, a successful CSP is expected to predict both the structures and the relative stabilities accurately, and sometimes also how those stabilities vary with temperature and pressure.

Thanks to this progress, the pharmaceutical industry is rapidly adopting CSP to help de-risk against the unexpected appearance of new crystal forms or to narrow the search space of crystal co-formers to be screened experimentally.[22–31] CSP has expanded from a purely academic endeavor to one with multiple private companies developing software, creating new algorithms, and providing contract CSP services. Some larger pharmaceutical companies have their own internal CSP teams as well. Beyond pharmaceuticals, CSP is being used to understand or discover new functional organic materials. In all of these application areas, CSP is helping to solve difficult crystal structures, anticipating new crystal forms, guiding experimental researchers toward the discovery of those forms, and enabling rational materials design.[23,32–34]

This perspective article seeks to highlight what organic CSP can accomplish today, how it can transform the discovery and understanding for a broad range of problems in organic materials, and where major outstanding challenges remain. Section 2 discusses reasons why CSP is such a difficult problem, while Section 3 provides a high-level overview of the methods currently used to overcome those challenges. Section 4 presents a variety of recent case studies that highlight the diverse frontiers of CSP, including examples from pharmaceutical formulation, its incorporation into nuclear magnetic resonance (NMR) crystallography, the discovery of new, highly porous organic crystals, the study of photochemical transformations in the solid state, and efforts towards the rational design of new materials. Finally, Section 5 discusses several directions in the field that will likely prove important in the next few years. For further reading, readers are also referred to several excellent earlier reviews that focus on CSP methods and applications in greater technical detail.[22,23,34–44]

2 The crystal structure prediction challenge

The difficulty of crystal structure prediction stems from several factors (Fig. 1): first, the search space of potential structures is massive, including 230 possible space groups, one or more molecules in the asymmetric unit, and, for many species, a competition between intramolecular conformational and intermolecular packing forces. While some of these complexities can reasonably be managed by, for example, constraining the search to the most common space groups and/or to crystals with just one molecule in the asymmetric unit, plenty of experimental crystals lie outside these constraints. Moreover, the conformational degrees of freedom in many modern active pharmaceutical ingredients and other highly flexible molecules are harder to circumvent, and they can dramatically increase the search space and the resulting computational costs of the structure prediction.

Second, the energy differences separating crystal polymorphs are small. Nearly all experimentally-known crystal polymorphs lie within 10 kJ mol^{-1} of one another,[13,45] and the energy differences are often just ~1–2 kJ mol^{-1}. Those small energy differences manifest from competitions among the hydrogen bonding, electrostatics, induction/polarization, and van der Waals dispersion interactions within and between the molecules. Achieving kJ mol^{-1} resolution in modeling these diverse interactions can be difficult for both force fields and electronic structure methods, especially for conformational polymorphs[46] whose crystal structures result from the interplay between changes in intramolecular conformation and the intermolecular crystal packing.

Third, while CSP often focuses on predicting 0 K crystal lattice energies,[40] real-world crystal structures are determined by free energies at finite temperatures and pressures. In smaller molecules with limited flexibility, the differences between relative lattice energies and relative room-temperature free energies are usually small (<2 kJ mol^{-1}).[45,47] However, the magnitude of the relative entropic/free energy contributions can

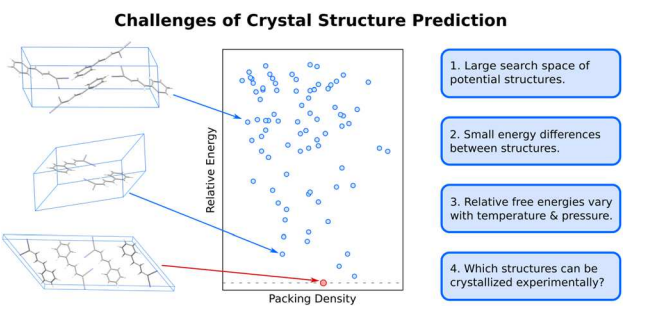

Fig. 1 Organic molecular crystal structure is difficult due to the large search space of potential structures (blue dots) on the 0 K crystal energy landscape which are separated by small lattice energy differences. Moreover, the relative free energies between polymorphs vary as a function of temperature and pressure, and not all thermodynamically feasible crystal structures can be readily crystallized experimentally.

increase significantly in large, flexible drug-like molecules[30,48] and disordered crystals.[49,50] Moreover, factors such as thermal expansion and dynamics can alter the finite-temperature crystal structures themselves. The magnitude of these effects is frequently modest, but not always.

Finally, the vast majority of CSP research has focused on the thermodynamic stability of the crystal, but polymorph crystallization is highly influenced by kinetics. CSP routinely predicts far more thermodynamically viable candidate structures than are ever observed experimentally. There are multiple reasons for this over-prediction of structures,[51] but crystallization kinetics are one major reason that more candidate polymorphs are not found experimentally. While there have been important advances in modeling organic crystal polymorph nucleation and growth in recent years,[52–55] the statistical mechanical sampling challenges and the need for accurate but computationally inexpensive potentials represent on-going hurdles to reliable prediction of crystallization kinetics. Moreover, CSP routinely focuses on infinite crystals, ignoring the surface energy contributions that depend on the size and shape of the finite crystallite. Surface energies can be relevant when considering the stability of nanocrystalline formulations or polar crystals, for example.[56–61] More detailed discussion of these issues is beyond the scope of this article.

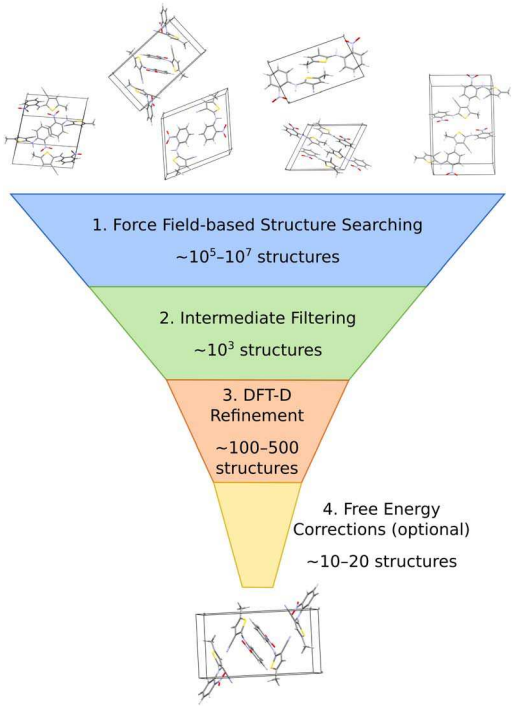

Fig. 2 A typical hierarchical crystal structure prediction approach might (1) generate and rank large numbers of candidate structures with an inexpensive force field, (2) refine many of the most promising structures with some method of intermediate accuracy and computational cost, (3) perform dispersion-corrected DFT refinement on a few hundred structures, and (4) perhaps end with free energy calculations on a small number of structures.

3 Current methods of crystal structure prediction

3.1 Overview of hierarchical crystal structure prediction

The most common organic CSP approaches employ hierarchical stages of structure refinement and ranking (Fig. 2). For example, the first stage in the hierarchy might employ an inexpensive force field potential to screen $\sim 10^5$–10^7 (pseudo-)randomly generated crystal structures, depending on the complexity of the species and the search space. The second-stage refines the $\sim 10^3$ lowest-energy structures with an intermediate-quality model. In the third stage, the few hundred most stable structures might then be refined further and ranked with dispersion-corrected density functional theory (referred to here as "DFT-D" for brevity, though many different dispersion-inclusive DFT models are used in practice). Optionally, one might perform final free-energy corrections for a handful of the most stable crystal structures to predict their stabilities at finite temperatures and pressures. More technical details have been reviewed elsewhere.[22,23,34–43]

A number of features factor into a successful crystal structure prediction. Ensuring a suitably-thorough search of crystal packing space is crucial. A routine search might focus only on crystals with a single molecule in the asymmetric unit ($Z' = 1$) and from the ~15–20 most common space groups that account for over 90% of observed organic crystals.[62] More exhaustive searches might consider all 230 space groups and/or crystals with $Z' > 1$. Significant additional complexity is introduced to the CSP for flexible molecules, due to the need to consider various equilibrium and non-equilibrium intramolecular conformations, or for multi-component crystals (co-crystals, solvates, hydrates, *etc.*), due to the much larger search space. Addressing these various complications can substantially increase the overall computational cost. Within the chosen search space, random[63–65] or low-discrepancy pseudo-random search approaches[66–68] are common in molecular CSP, though other global search algorithms such as simulated annealing,[69] particle-swarm optimization,[70–72] basin hopping[73] or evolutionary optimization[42,74–77] are also used.

The low-cost computational models used in the early stages of a hierarchical CSP enable broad searching, and the subsequent filtering out of poor candidates allows the more expensive methods to be applied only to the more promising candidates. Care must be taken to ensure the models used in the early stages are accurate enough to identify and select the relevant structures for later-stage refinement. For example, conventional off-the-shelf force fields are often not reliable enough for CSP. Section 3.2 will discuss how more customized potentials are often used instead.

Much of the current success in CSP stems from the widespread adoption of density functional theory for late-stage refinement and ranking. In the 4th Blind Test of CSP, a DFT-D-driven approach was the first to correctly predict the crystal structures of all the target molecules.[78] The development of accurate, non-empirical, and computationally efficient van der Waals dispersion corrections for DFT,[79] such as D3 and D4,[80–82] many-body dispersion (MBD),[83–85] and the exchange-hole dipole

moment (XDM) model,[86] has been particularly important. Generalized gradient approximation (GGA) density functionals used most frequently for computational expedience,[87–92] though refining the single-point energies with hybrid density functionals improves the results meaningfully.[93–98]

While many CSP studies finalize their predictions with DFT-D structures and lattice energies, others proceed further to consider finite-temperature free energies. Surveys of small molecule crystals have found that vibrational free energy contributions change polymorph stability orderings for ~10–20% of molecules at room temperature,[45,47] though the differences between lattice energies and free energies can increase for larger, more complex systems due to conformational flexibility or disorder. For simpler molecules, harmonic, quasi-harmonic, and/or other simplified anharmonic treatments capture the vibrational free energy contributions reasonably well. On the other hand, molecular dynamics-based approaches are potentially superior for describing more complex crystals, assuming a suitably accurate potential energy model. Such techniques will be discussed further in Section 3.2. Overall, the combination of accurate DFT-D models and (sometimes) vibrational free energy contributions frequently leads to successful crystal structure predictions, as demonstrated for many Blind Test targets[91,92,94,95,97] and for examples that will be discussed in Section 4.

In the end, performing a CSP produces a crystal energy landscape (Fig. 1), which is the set of predicted crystal structures and their relative lattice energies or free energies. Crystal energy landscapes at 0 K are often plotted as lattice energy *versus* crystal density, both because van der Waals forces generally favor more dense crystal packing motifs and because a scatter plot facilitates visualization of the large number of predicted structures. In some cases, one may simply wish to identify the most stable crystal structure(s). However, consideration of the full crystal energy landscape can provide valuable insights into the crystallization behaviors of a species[22,23] or help elucidate crystal structure–property relationships for materials design. Before discussing such applications in Section 4, we discuss several areas where methodological developments are actively underway.

3.2 Areas of active methodological developments

3.2.1 Improved models for early- and intermediate-stage structure refinement and ranking. In a hierarchical crystal structure prediction such as Fig. 2, the late-state DFT-D structure refinement and ranking typically consumes a large fraction of the total computational cost. Because the lower-cost intermediate stage models are generally less reliable than DFT-D, it is common to carry a relatively large number of structures forward to the DFT-D refinement to reduce the risk that an important structure is discarded early on (as happened in some cases during the sixth Blind Test[21]). Unfortunately, performing DFT-D refinement on many structures is computationally expensive. Therefore, the total computational cost of the CSP can potentially be reduced by improving the quality of the early/intermediate filtering model(s) so that fewer structures need to be carried forward to the DFT-D stage.

A number of strategies are currently being used to achieve this. One very successful approach involves parameterizing tailor-made force fields for each system based on DFT-D calculations.[99] The force fields often employ fairly standard functional forms, with terms describing the intramolecular geometry, short-range intermolecular repulsion, long-range London dispersion, point-charge or multipolar electrostatics, and sometimes induction/polarization,

$$U = U_{\text{intra}} + U_{\text{rep}} + U_{\text{disp}} + U_{\text{es}} + U_{\text{pol}} \quad (1)$$

but system-specific parameter tuning achieves higher accuracy than could typically be obtained with off-the-shelf force fields. Multiple force fields can be fitted to different subsets of data to predict and score structures independently, thereby potentially increasing the extent of the crystal packing space searched and providing insight into the uncertainties in the models.[72] Moreover, as the CSP proceeds, the force fields can be reparameterized iteratively based on the results of DFT-D structure refinement as well as monomer/dimer quantum mechanical benchmarks (Fig. 3).[72] Iterating the force field parameterization toward self-consistency with DFT-D helps ensure the search is performed with near-DFT-D quality. This iterative process also produces a more robust force field that can be used to evaluate finite-temperature free energy corrections.[48] Machine learning potentials represent a natural extension of this idea.[44,100,101]

Low-cost semi-empirical methods are similarly promising for intermediate refinement of crystal structures and lattice energies.[102–109] These can be further combined with Δ-ML, in which an ML model is trained to correct a simpler model up toward the quality of a more expensive one. Species-specific Δ-ML models have been used in CSP to correct semi-empirical density functional tight binding (DFTB) toward the accuracy of hybrid functional DFT-D,[110,111] or to correct GGA functionals up to hybrid functional DFT-D or correlated wave function methods.[112,113]

Finally, *ab initio* force fields fitted to symmetry-adapted perturbation theory (SAPT)[114,115] have also improved considerably. SAPT calculations naturally decompose the different types of intermolecular interactions (electrostatics, exchange-repulsion, *etc.*), which can be used to help ensure physically-sensible parameter fits in the potentials. Successful SAPT potentials could already be found in the literature 15 years ago,[116–120] but the algorithms and protocols have now matured to enable highly-automated fitting for organic molecules with modest conformational flexibility.[121,122] A recent study[123] of fifteen organic molecules found that this approach placed the experimental structure within the top 10–20 structures (and often in top 5). Subsequent DFT-D refinement of the top 20 structures generated by these potentials for each species ranked the experimental structure as the most stable one in every case. Thus, these potentials are very accurate on their own and can provide an excellent short-list of candidate structures for subsequent refinement with fully quantum mechanical approaches. The biggest outstanding question is how efficiently these fitting algorithms can be generalized to highly flexible molecules

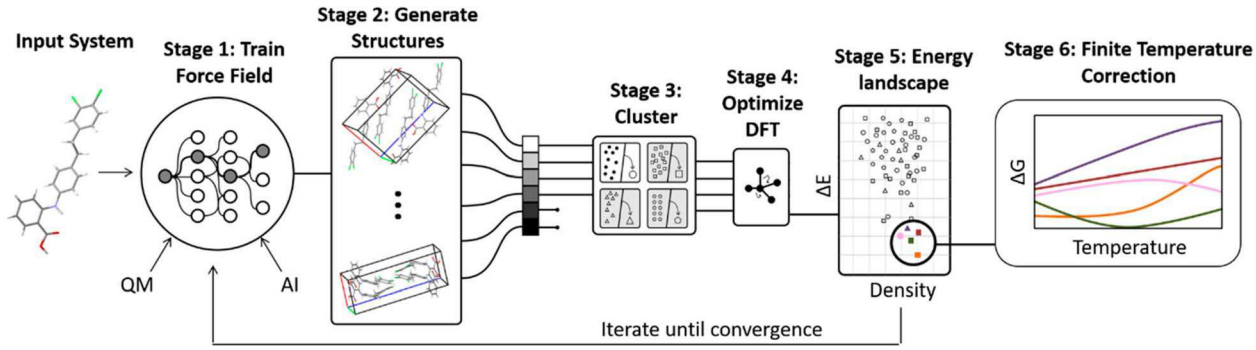

Fig. 3 Some CSP procedures involve iterative cycles of force field fitting, structure prediction, and DFT-D structure refinement until the force field and DFT-D crystal energy landscapes are suitably consistent. Adapted with permission from ref. 48. Copyright 2020 American Chemical Society.

3.2.2 Addressing DFT delocalization error in crystal structure prediction. Many CSP successes rely on dispersion-corrected DFT functionals. Commonly-used GGA and hybrid density functionals balance accuracy and computational cost and usually enable reliable refinement and ranking of hundreds of crystal structure candidates. However, approximate density functionals generally suffer from delocalization error (a.k.a. many-body self-interaction error),[124,125] which manifests as a spurious tendency to prefer overly delocalized electron densities. Delocalization error leads to systematic errors such as the underestimation of band gaps, underestimation of chemical reaction barriers, erroneous spin state energy differences, over-estimation of hydrogen bond strengths, and problematic conformational energies.

The impacts of delocalization error in CSP were first highlighted by Johnson and co-workers in the context of reanalyzing the conformational energies in candidate structures for Blind Test molecule X,[92] the lattice energies of halogen bonded crystals,[127] and, most dramatically, by showing how it could spontaneously convert neutral acid-base co-crystals to their charged salt forms.[128] The present author's group has since found many more examples where DFT delocalization error significantly impacts the relative stabilities for polymorphs of small molecules,[126,129,130] pharmaceuticals,[126] rubrene organic semiconductor materials,[131] and photochromic materials.[132–137] All of these systems have crystal structures which differ in the extent of π conjugation, either due to changes in the intramolecular conformation (conformational polymorphism) or chemical reactions that convert sp^2-hybridized atoms to sp^3-hybridized ones (*e.g.* cycloaddition reactions). Fig. 4 shows how DFT delocalization error over-stabilizes the more planar conformations of the ROY molecule,[126,138,139] the impacts of which will be discussed further in Section 4.3.

Delocalization error is particularly pronounced in GGA functionals such as PBE. Hybrid functionals such as PBE0 help mitigate the impacts of delocalization error,[96,97] though the necessary amount of exact exchange needed can vary.[97,126] Because the impacts of delocalization error on conformational energies are intramolecular in nature,[126,140] an alternative strategy can be to perform a simple conformational energy correction,

$$\tilde{E}_{\text{crystal}} = E^{\text{DFT}}_{\text{crystal}} - E^{\text{DFT}}_{\text{intra}} + E^{\text{High}}_{\text{intra}} \quad (2)$$

that computes the DFT crystal energy $E^{\text{DFT}}_{\text{crystal}}$ and replaces the DFT-D intramolecular energy $E^{\text{DFT}}_{\text{intra}}$ with one computed using a more advanced model that is free of delocalization error, $E^{\text{High}}_{\text{intra}}$, such as correlated wave function methods[126] advanced density functionals, or even density-corrected DFT.[141]

3.2.3 Improved treatment of finite-temperature free energies. Switching the focus from 0 K lattice energies (E) to finite-temperature Gibbs (G) or Helmholtz (F) free-energies,

$$G(T,P) = E + F_{\text{vib}}(T) + PV \quad (3)$$

can be important for making real-world predictions about the most stable polymorphs, polymorph phase transitions, the

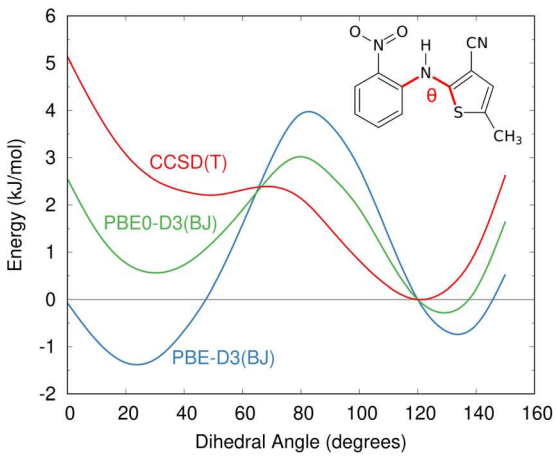

Fig. 4 Delocalization error in GGA and hybrid functionals such as PBE and PBE0 leads to over-stabilization of more planar conformations of the ROY molecular relative to those with a dihedral angle closer to 90°, as compared to high-level coupled cluster benchmarks.[126] This impacts the predicted crystal energy landscape, as will be discussed in Section 4.3.

formation of hydrates as a function of humidity, *etc.* The simplest approximation for these effects involves computing the static harmonic Helmholtz vibrational free energy contributions F_{vib} *via* lattice dynamics.

The quasi-harmonic approximation[142,143] refines the treatment further by approximating how the phonons and F_{vib} contributions change as a function of unit cell volume, which is especially important for the low-frequency modes.[144] The quasi-harmonic approximation enables predicting the temperature-dependent thermal expansion of the crystal lattice up to moderate temperatures, leading to improved-quality structures,[145–147] thermochemical properties,[148–151] spectroscopy,[152–154] and even polymorph phase diagrams.[152,155,156]

Lattice dynamics calculations are considerably more expensive than computing the energy, particularly due to the need to capture phonon dispersion. For this reason, they are typically computed with relatively inexpensive DFT-D functionals. A multi-level approach that combines a higher-level treatment of the phonons in the crystallographic unit cell with a lower-cost treatment in the supercell can reduce the costs further.[157,158] Although the quasi-harmonic approximation improves the description of lower-frequency modes, it does not address anharmonicities in the higher-frequency modes that are insensitive to the lattice parameters.[144] One simple approach for those phonon modes employs a 1-D anharmonic model to improve the description of each individual mode.[94,95] Vibrational self-consistent field calculations can capture anharmonicity more fully,[159,160] albeit at significantly higher computational cost.

Alternatively, molecular dynamics (MD) techniques can improve upon these static lattice dynamics approaches. MD simulations naturally capture anharmonicities.[144] Moreover, the finite-temperature dynamics will sometimes sample multiple minima on the potential energy surface, capturing contributions which would be missed entirely by (quasi-)harmonic models.[161,162] MD approaches are also inherently better-suited to describing dynamically disordered crystals.

One successful MD approach employs a pseudo-supercritical path approach to relate the free energies of the crystal polymorphs to that of an Einstein crystal reference state.[49,163,164] For example, a CSP study of the polymorphs of drug candidate oxabispidine found that the form A was several kJ mol^{-1} less stable than form B, contrary to experimental observations. However, applying this free energy correction approach on top of the 0 K lattice energy predictions demonstrated enantiotropic relationship, with form A becoming the thermodynamically preferred form near ambient temperature (Fig. 5).[30] Beyond classical molecular dynamics, path integral studies have shown that nuclear quantum effects can also be important for determining the relative polymorph stabilities in aspirin[165] and ices.[166]

The biggest challenge with MD approaches is the need for extensive sampling, which means that *ab initio* MD simulations are extraordinarily expensive computationally—*e.g.* ~2 million central processing unit (CPU) hours for paracetamol.[165] Therefore, inexpensive energy potentials must be used in practice. As noted before, standard force fields will frequently lack the requisite accuracy needed for CSP applications. However, good-quality tailored force fields and machine learning potentials being fitted as part of the search process (as described above) can also be used for the free energy simulations.[48,101] Re-weighting strategies that map from a low-cost free energy simulation to a higher-level one with only moderate sampling at the high level are also possible.[163,164]

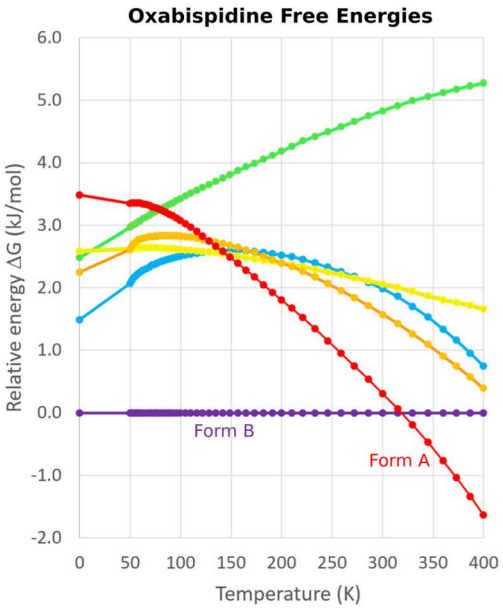

Fig. 5 Free energy calculations on experimentally-known forms A and B of oxabispidine and several other predicted polymorphs find that form A only becomes the most stable form near room temperature. Adapted with permission from ref. 30. Copyright 2021 American Chemical Society.

MD free energy approaches have benefits beyond simply predicting polymorph stabilities. Molecular crystal free energy landscapes tend to be smoother than lattice energy ones, with multiple lattice energy minima separated by small barriers coalescing into a single free energy well at finite temperatures. This feature enables reducing the number of predicted structures on a crystal energy landscape or even searching for crystal structures directly on the free energy landscape (see Section 4.9).

4 Selected applications at the frontiers of crystal structure prediction

Having discussed some of the model features that lead to successful crystal structure prediction, we now focus on case studies that demonstrate the range and capabilities of present-day CSP. These examples were chosen to highlight the diverse ways in which CSP can complement experiment across a broad range of organic materials, rather than aiming for a comprehensive review of the literature.

4.1 Pharmaceutical solid-form screening: rotigotine and galunisertib

Choosing a suitable solid form for manufacturing is an important step in pharmaceutical formulation. Researchers

desire crystals with suitable solubility profiles, mechanical properties, and stability. They want to avoid the surprise, late-stage appearance of new polymorphs with undesirable properties, such as those which necessitated the recall and reformulation of ritonavir[167,168] and rotigotine.[169,170] The risks are significant: it has been estimated that the most stable crystal form has not yet been discovered experimentally for some ~15–45% of small-molecule pharmaceuticals.[28] By providing a detailed understanding of the crystal energy landscape, CSP can complement experimental solid-form screening and help manage the risks of late-appearing polymorphs in pharmaceutical development.[22,23,171]

Consider two examples: rotigotine and galunisertib. Transdermal rotigotine patches are used to treat Parkinson's disease and restless leg syndrome. In 2008, the unexpected appearance of snowflake-like and highly insoluble crystals of a new crystal polymorph (form II) on the patches led to a major recall and restrictions on the drug in Europe, and its complete withdrawal from the U.S. market.[170] It took four years to reformulate the patches and return them to the U.S. market.[169]

Although CSP techniques were less mature in 2008, recent work by Mortazavi *et al.*[172] demonstrates how modern-day CSP techniques could have anticipated form II rotigotine. Starting from only the 2-D molecular structure of rotigotine, they employed a mixture of tailor-made force fields (fitted against DFT-D calculations), dispersion-corrected DFT, and harmonic vibrational enthalpy/free energy contributions to predict the most stable crystal structures of rotigotine, including both forms I and II (Fig. 6a). Their models indicate that form II is 7.6 kJ mol^{-1} more stable than form I, in exceptional agreement with the 7.5 kJ mol^{-1} measured experimentally (such excellent agreement probably reflects some fortuitous error cancellation). Today, a CSP prediction of a new polymorph that was so much more stable than the known form would warrant significant concern and would motivate further experimental screening efforts. Moreover, the higher packing density predicted for form II also would suggest that high-pressure experiments might facilitate its crystallization. In fact, similar CSP insights motivated the high-pressure crystallization experiments that discovered new polymorphs of dalcetrapib[27] and iproniazid,[171] as will be discussed in Section 4.9.

Overall, rotigotine has a sparse crystal energy landscape, with only two predicted crystal structures other than forms I and II in the low-energy (10 kJ mol^{-1}) region. One of those has stability intermediate between forms I and II. This putative "form III" has never been observed experimentally, and perhaps further investigations are warranted.

Whereas the CSP of rotigotine was performed long after its behavior was understood experimentally, CSP was directly integrated into the solid-form screening process for galunisertib.[26] This drug candidate for metastatic malignant cancer[173] has a complicated solid-form landscape: ten neat polymorphs and over 50 crystalline solvates have been discovered to-date. Its propensity for solvate formation complicated the experimental search for neat polymorphs, and a CSP was performed to identify any potentially important missing forms. The CSP revealed hundreds of potential crystal structures in the

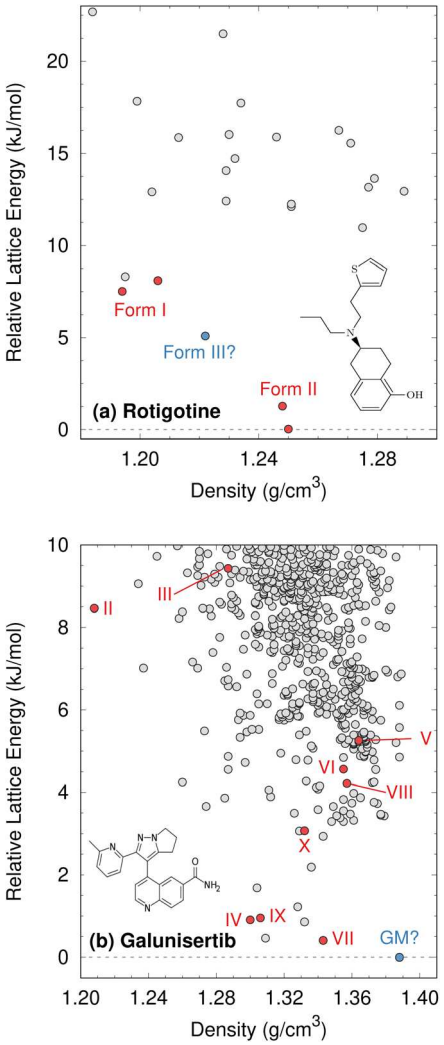

Fig. 6 Predicted crystal energy landscapes for (a) rotigotine[172] and (b) galunisertib.[26] Red points indicate experimentally-observed polymorphs. For rotigotine, a pair of static structures was identified for each of forms I and II which correspond to the two possible conformations of the disordered thiophene ring. The structures labeled "form III" for rotigotine and "GM" for galunisertib have not yet been found experimentally.

10 kJ mol^{-1} energy window (Fig. 6b). Such densely populated crystal energy landscapes are unfortunately more typical for pharmaceuticals than the sparser rotigotine one.

The initial CSP for galunisertib predicted seven of the ten of the polymorphs eventually found experimentally, but it missed the remaining three due to search constraints that had been imposed to expedite the CSP. As crystal forms lying outside the initial CSP search space were discovered experimentally, a second, broader CSP was performed using techniques very similar to those for rotigotine. This second landscape successfully predicted all experimentally-discovered polymorphs.

The galunisertib CSPs helped solve the crystal structures of forms VII and VIII. Experimental difficulties obtaining pure crystals of these forms complicated the powder X-ray diffraction patterns, but the structures were eventually solved using

comparisons against simulated powder diffraction patterns computed on candidate CSP structures. On the other hand, both CSPs predict that the most stable "global minimum" crystal structure has not yet been found experimentally, despite extensive efforts. This unrealized global minimum structure highlights two potential issues in CSP which will be discussed later: when are the accuracy limitations of widely-used DFT-D models problematic (Section 4.3)? When are predicted crystal structures actually crystallizable (Section 4.9)?

4.2 Addressing the further complexities in pharmaceutical crystal structure prediction

With mounting numbers of successful polymorph predictions for neat pharmaceuticals, CSP techniques are increasingly being applied to more complicated aspects of pharmaceutical formulation,[23] including disorder and multi-component hydrates, solvates, and co-crystals. Cases such as the experimental cancer drug gandotinib,[25] with its multiple hydrates, disorder, and difficulties crystallizing various forms exemplify the real-world complexities of pharmaceutical solid-form landscapes.

Consider first disorder, which is present in ~20–25% of crystal structures.[50] Static disorder results from molecules adopting a statistical distribution of different configurations or orientations in the lattice, while dynamic disorder is associated with the finite-temperature motions of molecules in the crystal. The distinction between the two types of disorder is not always sharp, however, and it can even vary with temperature.[174] Both types of disorder can stabilize a crystal structure entropically. Typical CSP protocols neglect disorder, though the prediction of multiple closely-related crystal structures with similar lattice energies can be suggestive of a greater likelihood for disorder to occur in the experimental crystal structures.[175–177]

To obtain more quantitative results, disorder needs to incorporated into the models. Dynamic disorder can potentially be described *via* molecular dynamics simulations,[49,161,175] for example, while a symmetry-adapted ensemble model which includes weighted energy contributions from all the configurationally unique structures is often used to treat crystals with static disorder.[178] Such descriptions are considerably more computationally demanding than conventional static structure models, unfortunately. A symmetry-adapted ensemble for a system with N disordered sites having two possible states each requires evaluating the energy for 2^N possible configurations, though symmetry reduces the number of unique configurations in practice.

Accounting for the effects of disorder can be important. A CSP study on the antihistamine medication loratadine,[50] for example, found multiple crystal structures corresponding to different components of the disorder. The initial landscape suggested that form I was relatively high in energy compared to other predicted forms. Form I became the most stable form only after it was modeled with a symmetry-adapted ensemble. Similarly, the initial predicted crystal energy landscape of gandotinib suggested that the most stable crystal polymorph had not yet been found experimentally. However, accounting for the disorder in form I made it isoenergetic to the predicted global minimum structure.[25]

Multi-component crystals are also extremely common in pharmaceuticals. Incorporation of water or other solvent molecules into a molecular crystal structure occurs frequently. In other cases, the active pharmaceutical ingredient (API) is deliberately crystallized as a salt (*e.g.* with hydrochloride) or with inactive co-formers to improve their solid form properties. CSP of multi-component systems can be considerably more difficult than single-component systems. Predicting when the co-crystal is thermodynamically preferred can be done pretty reliably.[179–181] On the other hand, the presence of multiple species increases the crystal packing search space considerably,[182] especially when multiple potential stoichiometries need to be considered.[183–185]

Despite these challenges, clear progress is being made. A number of successful hydrate predictions have been performed,[187,188] including ones that predicted the correct stoichiometries.[185] The use of free energy calculations to compare the stabilities of different co-crystal stoichiometries has also been demonstrated.[31] Data-driven algorithms can identify plausible locations for water molecules within an anhydrous crystal structure, enabling a high-throughput screen of potential hydrates from an existing CSP landscape.[189] Separately, Dybeck *et al.* impressively demonstrated that with the help of one experimentally determined co-existence point, the phase boundary between anhydrate and hydrate forms could be predicted as a function of temperature and relative humidity to within 10% relative humidity of experiment (Fig. 7).[186] An example of co-crystal stoichiometry prediction was also included in the recent 7th Blind Test, with results to be published soon.

For an example of a successful CSP applied to a multi-component salt crystal, consider the recent studies of the sleep-related drug candidate B5.[190] Whereas the neutral form of B5 has just one important crystal form,[190] understanding the solid-form landscape of its hydrochloride salt B5HCl proved much more difficult.[177] Extraordinary experimental effort was required to uncover two neat polymorphs of B5HCl, a dihydrate, and 11 alcohol solvates of B5HCl. The concurrent CSP study made

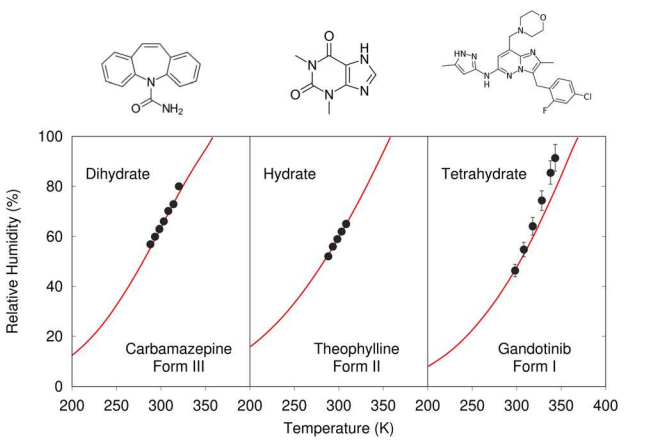

Fig. 7 Predicted phase-boundary between hydrate and anhydrate forms of three drugs as a function of temperature and relative humidity. They show nearly quantitative agreement between theory (red lines) and experimentally-derived coexistence points.[186]

several contributions to the eventual understanding: It highlighted the stability of form I, which helped explain its insolubility in various solvents and the difficulties in producing other crystal forms experimentally. It showed that the experimental forms discovered included examples of all major packing motifs found on the computational landscape, suggesting that the experimental screen was suitably complete. Moreover, the large number of closely related crystal structures on the computed landscape also pointed to the likelihood of disorder, especially for one particular conformation of the $B5H^+$ molecule. This helped rationalize the experimentally-observed disorder and difficulty in growing crystals that were suitable for diffraction.

Finally, a typical solid-form screen might consider multiple different possible co-formers, potentially multiplying the number of CSPs that may need to be performed. Sugden and co-workers recently demonstrated one clever approach for simplifying this task.[191] Their standard CSP approach employs pre-fitted local approximate potentials to describe important intramolecular conformational flexibility in their molecules,[192–194] and generating those models from quantum mechanical calculations requires non-trivial effort. However, by creating a library of these conformational energy models for commonly-used co-former species in advance, they can quickly run a CSP for a given API with a whole suite of potential co-formers. After fitting the local approximate models for the new API, they can run a CSP to screen each API + co-former combination in just ~2–3 days on a moderately sized cluster. Testing on three different drug molecules found that these relatively fast CSPs proved sufficiently accurate to rule out co-former candidates that were unlikely to form co-crystals experimentally, even if additional effort might be to refine the predictions for the most promising co-formers.

4.3 ROY and the impacts of DFT delocalization error

While CSP is increasingly successful, factors such as DFT delocalization error can still lead to incorrect predictions. The ROY molecule, so named for its vibrant red, orange, and yellow crystals,[195] is a classic example of polymorphism and holds the current world record with 12 fully-characterized polymorphs,[196–204] plus a thirteenth incompletely characterized form.[139,205] Several ROY polymorphs were discovered/solved in the past few years.[200–204] Despite the importance of this system to the field of polymorphism, predicted crystal energy landscapes of ROY were highly inconsistent with experimental polymorph stabilities rankings[139,200,206] until very recently. The conformational flexibility of the ROY molecule (Fig. 8) is the primary factor behind ROY's colors[207–210] and its propensity for polymorphism, but it also caused problems for CSP. As shown in Fig. 4, DFT delocalization error over-stabilizes the orange and red polymorphs, which have more planar intramolecular conformations exhibiting extended π conjugation, relative to the yellow polymorphs with their nearly perpendicular conformations that localize π electron density onto each ring.[126,138–140] These systematic biases found for GGA and hybrid density functionals[140] can be larger than the total experimental energy range spanned by the polymorphs.[195]

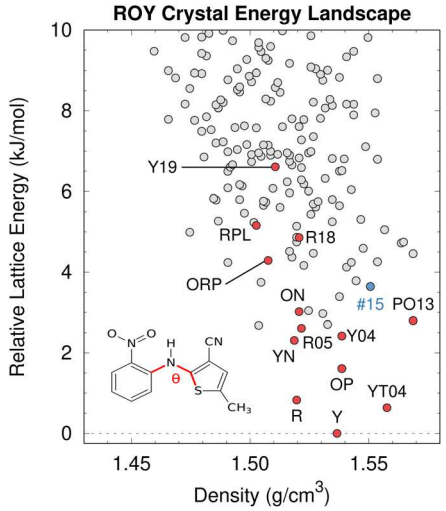

Fig. 8 After addressing the DFT delocalization error issues, the predicted crystal energy landscape of ROY shows that the lowest-energy polymorphs have already been discovered experimentally (red). Interestingly, the hypothetical rank #15 structure in blue is predicted to become the most stable structure near 10 GPa. Figure adapted from ref. 129.

Fortunately, correcting the ROY intramolecular conformational energy contribution to the lattice energy using correlated wave function methods[126,211] or density-corrected DFT[141] dramatically improves the crystal energy rankings relative to experiment.[129] As shown in Fig. 8, the resulting landscape reveals that the nine of the twelve lowest-energy candidate structures on the ROY landscape have already been crystallized experimentally. The four higher-energy forms (including the proposed-but-unconfirmed structure of the RPL polymorph[139]) are known to be metastable and/or were difficult to crystallize, suggesting that their less stable lattice energies are plausible. Interestingly, the calculations also suggest that the rank #15 structure on the landscape becomes the most stable form at high-pressure. While previous experimental high-pressure studies have not discovered any new polymorphs,[210,212] this prediction suggests further efforts to produce high-pressure forms may be worthwhile.

Overall, the ROY system highlights how, despite many successful DFT-driven structure predictions, there are cases where the most frequently-used DFT-D functionals are inadequate. Similar problematic DFT delocalization error issues occur with conventional DFT functionals for the anti-cancer drugs galunisertib (Section 4.1) and axitinib,[126] the photomechanical materials discussed in Section 4.5, and a number of other examples mentioned in Section 3. Fortunately, these errors can be overcome through intramolecular energy corrections of the sort used for ROY or the selection of a suitable density functional.[97]

4.4 Discovery of new porous organic materials

Porous materials are useful for gas storage and separations, but rationally designing porous organic crystals is difficult. Beyond the usual difficulty in intuiting the relationship between

molecular structure and crystal packing, porous organic crystals are exceptional because they counter the general thermodynamic preference toward dense crystal packing motifs. However, Pulido *et al.*[213] demonstrated how CSP and energy-structure–function maps could be used to drive experimental discovery of new porous organic crystals. They began by performing CSP on a series of candidate molecular building blocks. As expected, the lowest-energy structures were densely packed and non-porous. However, they identified several interesting "spikes" higher on the crystal energy landscape which corresponded to unusually stable porous structures (Fig. 9). While these porous structures lay tens of kJ mol^{-1} above the global minimum—much too energetically unstable to crystallize on their own—the authors recognized that these putative porous structures would be dramatically stabilized by guest solvent molecules adsorbed within the pores.

Next, they computationally characterized the "function" of every structure on their predicted crystal energy landscapes—their methane storage capacity and their potential for hydrocarbon separations. From this combined understanding of a molecule's propensity to form porous structures and the resulting functional properties, they identified the molecule **T2** (Fig. 9) as a promising candidate for new experimental crystallization screening studies. They discovered three new porous polymorphs of **T2** in addition to a previously known one. The new γ form of this molecule set a record for ultra-high-surface-area organic materials (3425 m^2 g^{-1}). The experiments confirmed several predicted properties of these crystals, including surface area, gas storage capacities, and some gas separation properties. In select cases, however, adsorption of larger molecules during the gas separation testing induced experimental phase transitions that would have been difficult to anticipate computationally. Nevertheless, this study highlights how structure prediction and energy–structure–function maps can drive experimental discovery.[214] Subsequent applications of these or similar concepts to porous materials[8,215–217] and organic semi-conductors[218–221] have further confirmed the role of CSP in the design and discovery of new organic materials.

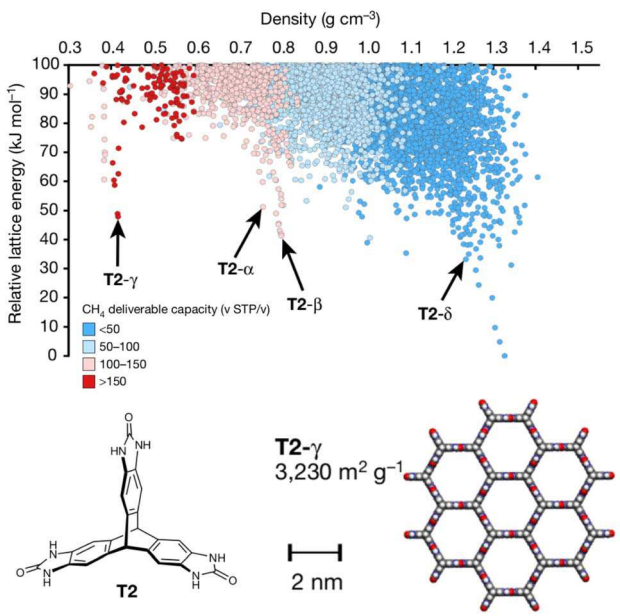

Fig. 9 Crystal energy landscape for the molecule T2, color coded by the methane storage capacity. The α–δ polymorphs represent particularly stable porous crystals that have been crystallized experimentally. T2-γ set a record for the highest surface area among porous organic molecular crystals. Adapted with permission from ref. 213. Copyright 2017 Springer Nature.

While CSP often strives to accurately predict the most stable crystal structure using high-quality energy models, such detail is not always required to establish useful design principles for organic materials. Following on work that showed how the combined efforts of computation and experiments could lead to the rapid discovery of new porous organic cage materials for gas separation and storage applications,[222,223] Wolpert and Jelfs demonstrated that a simple coarse-grained model could give meaningful insights into how porous organic cage molecules pack in the solid state.[224] Rather than perform traditional, expensive CSP on each new large organic cage molecule, their simplified model represents the cages as octahedra with "patches" on each face that distinguish between whether they contain either an arene group or an open pore. After expressing the intermolecular interactions between patches *via* a simple model Hamiltonian, they explored the types of packing motifs that were preferred across a range of interaction parameter strengths. They developed a general understanding for how the chemical features of the cage molecules translate into the resulting crystal packing. Moreover, they can determine the key patch parameter values using just gas-phase dimer DFT-D calculations, and then predict the preferred crystal packing for a given species. The simplicity of this approach makes it highly amenable to high-throughput screening. Coarse-grained approaches like this can significantly narrow the molecular search space for new functional materials before performing more detailed CSP studies or experiments.

4.5 Establishing design principles for organic photomechanical engines

In organic photomechanical crystals, solid-state photochemical reactions can induce elongation, compression, twisting, bending, jumping, cracking, splitting, and other deformations of the crystal.[10–12] Particular interest lies in harnessing these structural transformations to do mechanical work, such as for light-driven actuators[225–230] or locomotion.[231,232] Designing such materials requires understanding how molecular structure translates to crystal structure, how the crystal deforms due to the solid-state chemical reaction, and what anisotropic work is performed by that deformation. Characterizing these transformations experimentally is frequently challenging, since the solid-state chemical reactions may be incomplete, may be carried out in nanocrystals instead of bulk single crystals, and often produce short-lived and highly metastable polymorphs.

These systems therefore represent an excellent opportunity for first-principles structure prediction techniques. Until recently, the use of CSP in understanding solid-state reactions has been rare.[233] We have now shown how CSP can be used as

part of a strategy to predict or even design large photomechanical responses.[135] The approach starts by predicting the crystal structures of the reactants and products (Fig. 10a). However, the bigger challenge lies in determining which of the many predicted product crystal structures is relevant. Because the solid-state reaction generates the product molecules within the crystal packing of the reactant, the product polymorph is typically thermodynamically metastable and often lies outside the typical ~10 kJ mol^{-1} energy window associated with viable polymorph crystallization. Since CSP normally focuses on identifying the most stable crystal form(s), we instead apply topochemical principles to predict the solid-state transformation (Fig. 10b) and to identify the correct structure on the crystal energy landscape (Fig. 10a).

Once the structural transformations are known, the amount of work performed can be predicted. In the spirit of gas heat engines, we defined an idealized photomechanical engine cycle[135] that enables computing the maximum work potentially performed by a given solid-state reaction (Fig. 10c). The engine model assumes the reaction occurs quickly and completely, thereby generating the product within the unit cell parameters of the reactant. Relaxation the crystal relieves the internal stress, deforms the crystal, and performs work.

Studies of solid-state [4 + 4] anthracene photodimerization[135,136] and diarylethene ring opening and closing[137] have revealed a number of important insights and design principles for organic photomechanical materials. First, the unique combination of high elastic modulus and large strains means that photomechanical organic crystals exhibit exceptional theoretical work densities up to at least 200 MJ m^{-3}. If these could be realized experimentally, they would be several orders of magnitude larger than the work densities of elastomers or inorganic piezoelectrics. Second, the crystal packing proves crucial: for one diarylethene derivative studied, the maximum anisotropic work density differs 40-fold between two crystal forms. Based on the modest number of cases studied thus far, the range of work capacities across different polymorphs is broader than the range found for different photochromic species/reactions. This suggests that researchers should increase their emphasis on crystal engineering in selecting their target photochrome, especially since packing has a much larger impact on the mechanical response than do minor photochrome modifications (*e.g.* halogenation).[136] Parallel alignments of the molecules in the crystal generally produce the especially large anisotropic deformations and work densities. Finally, the research has identified how the "memory" of the reactant crystal packing throughout the photochemical engine transformations biases them to produce greater work in the forward stroke direction than in the reverse one, enabling net work to be accomplished.

The need to understand solid-state chemical reactions extends beyond photomechanical systems. For example, solid-state photochemical degradation is a significant issue for pharmaceuticals, and crystal packing impacts photostability.[234] Alternatively, crystal packing can be used to perform solid-state photochemical synthesis with precise stereochemistry and quantitative yields.[235–237] Solid-state oxidations, reductions, isomerizations, bonds formations/cleavages, and many other reactions are also possible.[238] Inducing solid-state reactions *via* mechanical grinding or milling (mechanochemistry) creates further synthetic possibilities.[239,240] CSP can help engineer crystal packing motifs that either facilitate or inhibit solid-state reactions.

4.6 Rational design of organic semi-conductors

The ultimate promise of structure prediction lies in the complete rational design of new materials, starting from the molecular building blocks. Because the relationships between molecular structure and materials properties frequently cannot be intuited *a priori*, this rational design will likely require high-throughput crystal structure prediction to map these relationships across many candidate species.

True CSP-driven rational design of organic materials is still in its infancy, but Cheng *et al.* provided an intriguing peek at the future possibilities with their evolutionary exploration of

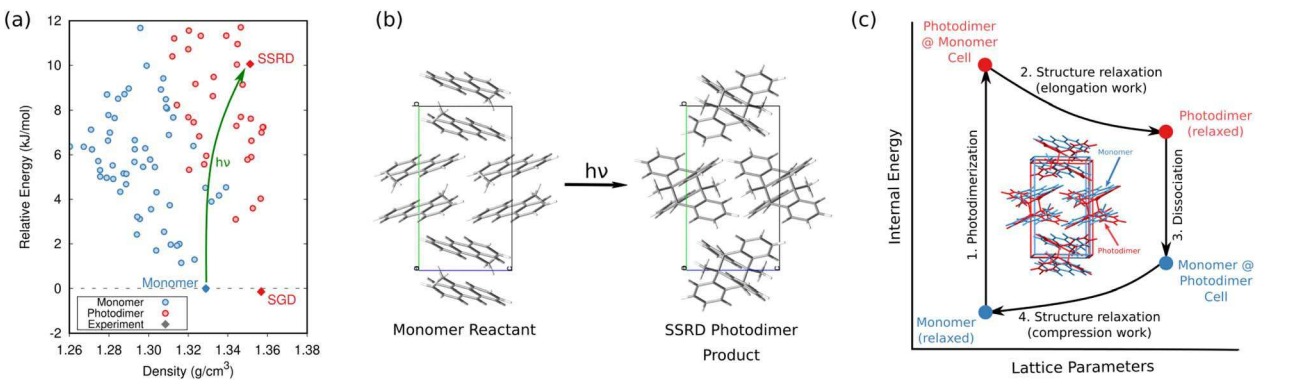

Fig. 10 (a) CSP for the 9-methyl anthracene monomer reactant (blue) and photodimer product (red). The experimental monomer (solid blue) photochemically transforms to the solid-state reacted dimer (SSRD) polymorph, which differs from the solution-grown dimer (SGD) polymorph obtained from solution-phase crystallization. (b) Scheme showing the solid-state topochemical [4 + 4] photodimerization. (c) The photomechanical engine cycle involves fast photodimerization/dissociation reactions (Steps 1 & 3) followed by structural relaxations (Steps 2 & 4) which perform work. Figure adapted from ref. 135.

organic semiconductor materials.[220] Building on earlier work that had manually examined small numbers of species,[218,219] they searched a chemical space of ~68 000 aza-substituted pentacene isomers with (nearly) arbitrary connections between the five aromatic rings. Predicting crystal structures for every one of these molecules would be utterly impractical. Instead, they first screened candidate species using simple molecular-based fitness functions optimized *via* an evolutionary algorithm.

Next, crystal structure prediction was performed on the most promising molecular structures. In any medium- or high-throughput scenario, the models used to predict crystal structures will likely sacrifice some accuracy to achieve faster throughput. As a result, the predicted global minimum crystal structure for each species is less likely to correspond to the actual experimental crystal structures. To compensate, they evaluated the performance of each species by computing both the electron mobility of the most stable structure and a Boltzmann-weighted mean/standard deviation electron mobility across all low-energy candidate crystal structures. This latter metric identifies species whose crystal structures have consistently good mobilities, rather than those that happen to have a high mobility for one particular polymorph which may or may not occur experimentally.

In the end, the strong dependence of mobility on crystal packing meant that many candidates exhibited overlapping landscape-averaged mobilities and deviations (Fig. 11). On the other hand, the best candidate identified here proved to have only a single, low-energy structure which also had high mobility, thereby providing confidence that the predicted global minimum structure would likely match experiment. Overall, the evolutionary algorithm search identified candidates with good landscape-average and global minimum mobilities. The species identified by the evolutionary algorithm compared fairly well against several species that had previously been identified by human experts[241] who used a mixture of computation and chemical intuition. Moving forward, incorporating solid-state properties directly into the molecular design search, rather than at the end, will likely to prove important for design of materials whose properties are very sensitive to crystal packing.

Separately, a CSP study correctly predicted the experimental crystal structure for chiral [6]helicene and rationalized the preference for the enantiopure crystal over the racemate.[221] After this study demonstrated good but crystal packing-dependent semi-conducting properties, a follow-up CSP study then examined the impact of derivatizing [6]helicene molecules. Dimer screening was used to investigate over 1300 substituted helicenes, after which CSP was performed on the most promising candidates. In the end, they identified a set of derivatives which had predicted electron mobilities three times that of a previously characterized helicene.[242]

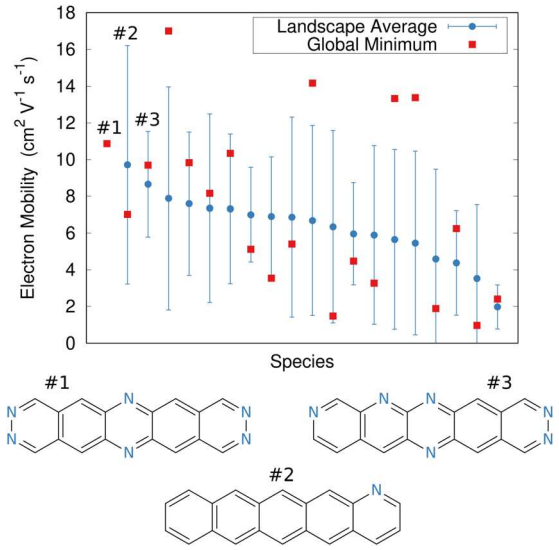

Fig. 11 Predicted electron mobilities for various organic semi-conducting material candidates discovered through an evolutionary search.[220] Red dots indicate the mobility for the most stable predicted structure of each species, while blue dots/error bars indicate the landscape-averaged mean mobility and standard deviation for the ensemble of low-energy structures. Three molecules with the highest landscape-averaged mobilities are shown.

4.7 NMR crystallography

The combination of crystal structure prediction with spectroscopic experiments can be particularly helpful in solving challenging crystal structures. One can frequently identify the experimental crystal structure by predicting candidate crystal structures, simulating the relevant spectroscopic observables for each one, and comparing the results with experiment. While this general idea has been applied to various spectroscopic characterization experiments, including powder X-ray diffraction,[243–247] Raman spectroscopy[153,248,249] and transmission electron microscopy,[250] nuclear magnetic resonance (NMR) crystallography represents the most widely-used combination.[39,251,252]

The combination of CSP, DFT chemical shift prediction, and solid-state NMR has been applied to many different pharmaceuticals and pharmacologically active compounds.[253–262] One appealing feature of these approaches is that the spectroscopic observables provide a second metric for assessing candidate structures that is "orthogonal" to the lattice energy typically used in CSP. NMR crystallography can frequently solve the experimental crystal structure even when that structure is not the most stable one predicted by the CSP. Unsurprisingly, models which predict chemical shifts more accurately[263–266] enhance the discrimination between correct and incorrect structural assignments in NMR crystallography.[267] Further synergies between experiment and computation are also possible in NMR crystallography. For example, structural constraints inferred from the NMR experiments can accelerate the structure determination by reducing the size of the CSP search space.[258]

Machine learning models for predicting chemical shifts[268–284] promise to accelerate NMR crystallography even further. The traditional NMR crystallography protocol involves first predicting a set of candidate crystal structures, computing the NMR spectra for each one, and comparing them against experiment,

all of which can require substantial computational effort. Recently, however, Balodis *et al.* demonstrated how NMR crystallography can directly refine the crystal structure against the experimental solid-state NMR spectrum.[261] The authors used Monte Carlo techniques to optimize an objective function which combined a weighted mixture of the structure's lattice energy and the error between its computed chemical shifts and experiment. These quantities can be evaluated inexpensively by using density functional tight binding (DFTB) for the energies and the machine-learning ShiftML model for the chemical shifts.[272,273] Minimizing this objective function successfully produced the correct structures for several difficult crystals, despite the "moderate" accuracies of DFTB and ShiftML compared to DFT.

The ability to refine crystal structures directly from NMR will be especially beneficial for large, highly flexible molecules, for which traditional crystal structure prediction can be very expensive. For example, Balodis *et al.* solved the crystal structure of ampicillin (Fig. 12),[261] despite the molecule adopting a high-energy intramolecular conformation in the solid-state that could easily be missed in a typical CSP search protocol. Separately, the rapid chemical shift prediction enabled by ShiftML also facilitated the structural characterization of amorphous drug candidate AZD5718.[260] Advances like these are likely to significantly increase the effectiveness of NMR crystallography over the next few years.

4.8 Predicting solubility and mechanical properties

As crystal structure prediction becomes more routine, the focus will increasingly shift to computing the chemical and physical properties of the predicted crystals. In many cases, these properties will depend sensitively on the 3-D structure, making it important to account for the impacts of finite temperature on crystal structure and stability.[40] Examples of computing properties related to gas separation and storage, photomechanical response, semi-conducting, and spectroscopy have already been discussed. Here, two more relevant properties are briefly discussed.

First, the solubility of organic species matters in many chemical applications, but it is especially significant in the pharmaceutical industry, where a large fraction of drugs in development suffer from low solubility. Because experimental measurements of solubility are time-consuming and resource-intensive, there has been considerable interest in predicting solubilities computationally. The field has frequently relied on informatics-driven approaches, including recent machine-learning efforts.[286–289] Although such models can be very effective, they typically omit explicit solid-state contributions and therefore cannot account for how changes in crystal packing will impact solubility, for example.

There are on-going efforts to develop accurate physics-based models which could overcome this limitation. While some approaches simulate the solid–liquid coexistence directly,[290] the more common approach employs a thermodynamic cycle that expresses the Gibbs free energy of dissolution from the free energies of sublimation and solvation,

$$\Delta G_{\mathrm{diss}} = \Delta G_{\mathrm{sub}} + \Delta G_{\mathrm{solv}} = -RT\ln(S_0 V_{\mathrm{m}}) \quad (4)$$

The dissolution free energy can then be related to the solubility S_0 *via* the temperature T, molar volume of the crystal V_{m}, and the ideal gas constant R. The sublimation free energy can be computed from periodic DFT calculations on the crystal lattice (and either approximating[291,292] or explicitly computing the phonon contributions[293]) or MD simulations.[294] The solvation free energy can be computed inexpensively *via* an implicit solvent model[292] or more elaborately with explicit MD simulations. While developments in this space continue apace, increasing reliance on higher-quality quantum mechanical models in computing the free energy contributions is bringing the errors to accuracies that are already approaching the best informatics models.

Predicting relative solubility difference between two crystal forms is arguably easier, since that requires only the free energy difference between the two solid forms, avoiding the need to compute solvation free energies. For the drug rotigotine, for example, Mortazavi *et al.* predicted an 8.3-fold difference in solubility between the two polymorphs with DFT-D, in excellent agreement with the 8.1-fold difference measured experimentally.[172]

Second, knowledge of the mechanical properties of a material provides valuable insights into its durability and potential applications.[295] Predicting the elastic constants enables one to screen materials *in silico* or to link features of the crystal packing to its mechanical response properties. For example, elastic constant predictions can be used to rationalize differences in how pharmaceuticals behave under tableting conditions. They helped explain the better tableting properties of paracetamol form II[296] and several co-crystals[297] compared to form I, of oxalic acid dihydrate over anhydrous oxalic acid,[298] and of co-crystals of celecoxib.[299] They showed that form II aspirin is mechanically stable,[300] contrary to an earlier suggestion based on nano-indentation experiments that did not fully characterize the anisotropy of the crystal. Elastic calculations also helped confirm and rationalize the surprisingly large Young's moduli of amino acid crystals[301] and a nucleic acid-based supramolecular assembly,[302] and they gave insights into the negative linear compressibility in several organic acids.[303] See a recent review for additional details.[295]

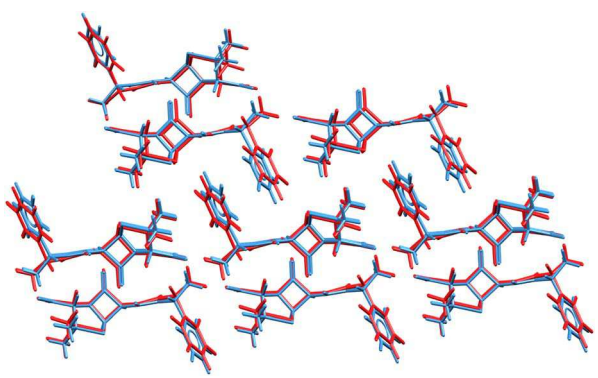

Fig. 12 Overlay of the ampicillin crystal structures determined from X-ray diffraction experiments[285] (blue) and the direct NMR crystallography solution (red) which employed a combination of DFTB energies and ShiftML chemical shifts.[261]

Early force-field predictions demonstrated the ability to compute elastic constants within ~40–50% of experiment,[304] with the neglect of thermal expansion of the crystal being a significant source of error. The accuracy with which these properties is predicted can be improved using quantum mechanical treatments and by accounting for thermal expansion (since mechanical properties are sensitive to molar volume). For example, the combination of accurate electronic structure methods and the quasi-harmonic approximation[143,305] enables quantitative prediction of the mechanical properties of simple compounds such as carbon dioxide[147] or deutero-ammonia.[93] Such techniques are also quite effective in organic compounds such as urea,[145] organic semi-conductors,[306] and energetic materials.[307] While DFT-D has become the most commonly-used approach, good-quality elastic constants can be obtained at lower computational cost. Spackman *et al.*[308] curated a large data set of experimental elastic constants, demonstrated good consistency between experiment and the semi-empirical s-HF-3c model, and they even identified several suspicious experimental measurements based on large discrepancies between the experimental and computed results.

4.9 Can the predicted structures actually be crystallized?

The computational prediction of a new polymorph can sometimes drive its subsequent experimental discovery, as exemplified in the case of the cholesteryl ester transfer protein inhibitor drug Dalcetrapib.[27] CSP on this drug correctly predicted the experimentally known form A and B polymorphs, which are closely related *via* a reversible temperature-dependent order-disorder phase transition. Interestingly, however, it also predicted another experimentally-unknown packing motif that lay very close in energy to form B. Motivated by lattice energy calculations that predicted this new structure to become more stable at pressures above ~0.2 GPa, high-pressure crystallization experiments led to the discovery of a new form C. In the end, this new polymorph proved unstable at ambient conditions and is therefore unlikely to impact the pharmaceutical formulation. In this manner, CSP played an important role in managing the solid-form risks for this drug. Similarly, analysis of the CSP landscape and calculations of the free energies as a function of temperature and pressure led to the experimental discovery of a high-pressure polymorph of iproniazid.[171]

In a third example of CSP driving discovery, a study of the crystal energy landscapes of structurally-similar tolfenamic acid, mefanamic acid, and flufenamic acid identified many thermodynamically plausible isostructural crystal forms. Based on this analysis, the authors successfully templated two new thermodynamically metastable polymorphs of tolfenamic acid using crystals and solid solutions of the other species.[309] Templating experiments informed by knowledge of predicted polymorphs similarly led to the discovery of new polymorphs of carbamazepine[310] and cyheptamide,[311] along with a new co-crystal of caffeine and benzoic acid.[312]

These cases all represent success stories for CSP, but there are counter-examples. The CSP-predicted global minimum energy crystal structure of galunisertib has never been found experimentally, despite years of effort.[26] This is probably due to a combination of poor crystallization kinetics stemming from its very unfavorable intramolecular conformation[26] and the fact that it is not actually the most stable structure (rather, it was artificially stabilized in the CSP by DFT delocalization error).[126]

More generally, identifying which putative CSP structures are likely to be crystallizable proves a major challenge. CSP usually predicts far more candidate crystal structures than are ever realized experimentally.[51] For some of these structures, that simply means that the proper crystallization experiment has not yet been performed. More often, however, it reflects the limitations of CSP approaches, such as the focus on thermodynamic stability instead if kinetic crystallizability, the fact that many distinct lattice energy minima coalesce into a single free-energy basin at finite temperatures, and errors in the predicted energies (*e.g.* due to delocalization error biases). See ref. 51 for further discussion. Additional challenges in assessing synthesizablity and the challenges associated with theory-driven discovery of materials are discussed extensively in two recent reviews by Jelfs and co-workers.[32,33]

To improve our understanding of which predicted polymorphs will be experimentally relevant, on-going research efforts focus on reducing the number of crystal structures on the crystal energy landscape, on strategies for identifying crystals that are likely to be crystallizable, and on predicting the thermodynamic conditions under which different polymorphs are most likely to form. In 2005, Raiteri *et al.* demonstrated that metadynamics simulations dramatically simplified a benzene crystal energy landscape containing tens of lattice energy minima down to just a handful of structures on the free energy surface, most of which have been observed experimentally.[313] The discarded structures are either labile, converting to different forms at finite temperature, or they correspond to different static structure representations from the same dynamic ensemble. Metadynamics has similarly been applied to reduce the crystal energy landscape of pigment red 179 (ref. 314) and, with less success, to 5-fluorouracil.[315] Metadynamics and other enhanced sampling methods have proved effective for searching crystal energy landscapes directly,[313–322] avoiding the need for post-hoc landscape reduction.

Recent efforts to systematize landscape reduction *via* a combination of MD, structure clustering, and metadynamics reduced the numbers of structures on the urea, succinic acid, and ibuprofen landscapes by ~65–90%.[323,324] The ibuprofen study is particularly impressive (Fig. 13), as its landscape containing 555 crystal structures is representative of the complexity of many real-world systems. Achieving this reduction has required significant efforts toward developing automated approaches for fingerprinting, simulating, and clustering the large numbers of structures involved.

The combination of CSP and metadynamics also helped rationalize the discrepant crystallization behaviors of two "sulflowers." Experimentally, the original sulflower molecule crystallizes readily, while the structurally-related persulferated coronene forms only an amorphous solid.[325] Using molecular dynamics and metadynamics, Sugden *et al.* demonstrated that while sulflower has a number of stable low-energy crystal forms

(including the experimental crystal), all 20 of the lowest structures of persulferated coronene became disordered in the dynamics simulations, consistent with its amorphous behavior experimentally.[326]

A promising new, even simpler approach for landscape reduction based on the threshold method was demonstrated by Butler and Day.[327] They coarsely estimate the energy barriers and structural relationships between predicted polymorphs *via* Monte Carlo moves that translate or rotate molecules and/or deform the unit cell parameter, accepting only moves that stay below a given energy threshold. In this manner, the approach identifies structures that can interconvert within a chosen energy threshold.[328] A 5 kJ mol^{-1} threshold reduced the number of structures on the crystal landscape by ~65–99% for several small organics.[327]

Landscape reduction does not guarantee that the remaining predicted crystal structures will be crystallizable. The understanding and modeling of nucleation and growth kinetics that leads to the crystallization of specific polymorphs remains difficult,[329] though progress is being made.[52–55] Instead, heuristic models are often used to identify crystal forms that are likely to crystallize. As noted in discussing galunisertib, conformational strain of the molecules in either the gas phase or appropriate solvents is sometimes considered as a factor in the crystallizabiliy of predicted polymorphs.[26,46,330,331] Montis *et al.* estabilished a relationship between low surface roughness (rugosity) and crystallizability that can be used to infer the relative likelihood of crystallizing various polymorphs on the crystal energy landscape.[332] In systems such as ROY, some of the less-stable polymorphs observed experimentally are among the smoothest, suggesting that kinetics favors their formation. In contrast, several other pharmaceutical polymorphs that have been difficult to produce have both higher energies and high rugocities, suggesting that both thermodynamics and kinetics hinder crystallization. The data-driven generalized convex hull approach of Anelli *et al.* is another promising strategy for exploring which crystal structures might be experimentally synthesizable.[333]

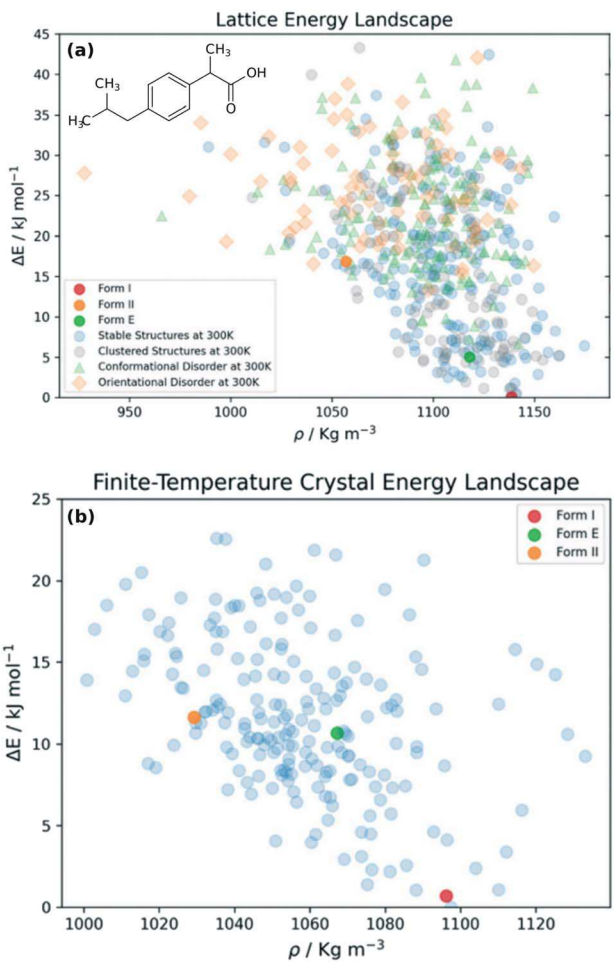

Fig. 13 Elimination of (a) lattice energy minima that are labile or effectively equivalent at (b) finite temperatures using molecular dynamics, clustering, and enhanced sampling techniques reduces the number of predicted ibuprofen crystal structures by 65%. Figure adapted from ref. 323.

Finally, while predicting crystallization kinetics remains difficult, there has been progress in predicting the thermodynamic conditions under which a given polymorph will be preferred—*i.e.* polymorph phase diagrams. At pressures greater than ~10–20 GPa, where factors such as thermal expansion become less significant, phase diagrams can be often be predicted with good accuracy.[152,334–338] However, the situation is more difficult closer to ambient conditions, since the predicted phase transition temperatures can be extremely sensitive to small errors in the computed free energies.[339] Despite these challenges, quite accurate temperature-dependent phase diagrams have been predicted for systems such as ice,[338] carbon dioxide,[336,337] methanol,[155] and resorcinol.[156] Polymorph phase transition temperatures have also been predicted for more complicated drug species.[30,48] However, one should bear in mind how important fortuitous error cancellation is in predicting phase boundaries. For example, a 1 kJ mol^{-1} error in the relative free energies between α and β methanol alters the predicted ambient-pressure phase-transition temperature by more than 200 K![155]

In the end, even if current CSP techniques cannot perfectly determine which crystal forms will be realized experimentally, they remain useful for assessing the polymorphic "risk" for a given species.[28,171] This is particularly valuable for the pharmaceutical industry, as exemplified in the studies on galunisertib,[26] gandotinib,[25] hydrates,[340] and salts[177] discussed earlier.

5 Future outlook

Given the rapid developments in CSP over the last 5–10 years, it is interesting to speculate where new advances will occur over the next several years. First, it is likely that there will be increasing emphasis on using finite-temperature free energies instead of 0 K lattice energies for the final rankings. This is already routinely being done in the pharmaceutical industry, where compute budgets are typically larger than in academia, and it is sometimes done in academic studies as well. Rapid improvements in machine learning potentials will likely also

increase the accuracy with which those free energy simulations can be performed by enabling dynamics-based approaches to be used on potentials that approach quantum mechanical accuracy.[44,101] Improved understanding of how the strengths and weaknesses of widely-used DFT-D methods (*e.g.* delocalization and other systematic errors) impact crystal energetics will make it easier to identify when crystal energy rankings are likely to be problematic.

Second, interpretation of the crystal energy landscape will continue to gain in importance. Rapid developments in crystal energy landscape reduction are likely continue apace. Methods such as meta-dynamics or the threshold algorithm hopefully become much more widespread and routine. Once again, accurate, inexpensive potential energy models and structure clustering strategies based on ML should further improve the performance of these techniques.[44] Improved uncertainty quantification for the computed structure energetics will also help users better assess the risks of predicted polymorphs on the landscape.

Third, it is not unusual for a high-accuracy crystal structure prediction to cost one million CPU hours per species at present. Entering the era of CSP-driven rational design will place a greater emphasis on performing "reasonably reliable" CSPs that have orders of magnitude lower computational cost, such that candidate materials can be screened en masse. Such approaches could mean learning to extract useful information from imperfect crystal energy landscapes (as in the organic semi-conductor design study discussed in Section 4.6), developing new intermediate ranking and refinement models (a.k.a. surrogate models[43]) that more effectively filter structures to reduce the number of final structures for which DFT-D calculations are needed, or even adopting entirely new data-driven topological approaches for generating short-lists of candidate structures quickly, without extensive hierarchical filtering algorithms.[44,341,342]

Fourth, beyond merely predicting structures, rational design efforts will increase the emphasis on computing functional properties of the putative crystals. Examples for gas storage and separations, organic semi-conducting properties, and photomechanical responses were already mentioned above. Feng and co-workers recently computed the photoluminescence properties of ROY and co-crystals of 9-acetylanthracene to understand the interplay of intra- and intermolecular interactions.[208] Improved ability to predict pharmaceutical solubilties (Section 4.8) or to assess the photostability of candidate formulations (Section 4.5) would be very useful as well.

Fifth, as the applicability of CSP expands, there is also a clear need for the development of more user-friendly software tools to democratize access to CSP. Current CSP is still almost exclusively performed by specialists. In academic research environments, CSP often relies on a disjointed collection of software packages and home-built scripts for passing structures between them and processing the results. CSP tools developed in industry are more user-friendly, though those companies often cater more to larger budgets and computing capabilities of pharmaceutical companies than to smaller-scale academic research groups. Moreover, new method developments from different groups are not always widely/publicly available in the short-term. Of course, radical reductions in computational cost will also be needed to enable truly widespread use of CSP by non-expert practitioners.

6 Conclusions

In conclusion, crystal structure prediction has advanced dramatically to the point where experimental crystal structures can be predicted successfully much more often than not. Applications of CSP have moved on from small-molecule benchmarks to real-world pharmaceutical formulations and functional organic materials. New frontiers are opening in areas such as the ability to use CSP to rationally design new materials with targetted properties or to model solid state chemical transformations. Identifying which predicted crystal structures can be made experimentally has been challenging, though good progress is being made there as well. Heinlein's dream of theory-driven materials design is quickly becoming reality, even if it is a couple decades late.

Author contributions

G. J. O. B. conceived and wrote this article.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Support from the National Science Foundation (CHE-1955554) is gratefully acknowledged.

References

1 A. Y. Lee, D. Erdemir and A. S. Myerson, *Annu. Rev. Chem. Biomol. Eng.*, 2011, **2**, 259–280.
2 B. Sandhu, A. S. Sinha, J. Desper and C. B. Aakeröy, *Chem. Commun.*, 2018, **54**, 4657–4660.
3 L. Casali, L. Mazzei, O. Shemchuk, L. Sharma, K. Honer, F. Grepioni, S. Ciurli, D. Braga and J. Baltrusaitis, *ACS Sustainable Chem. Eng.*, 2019, **7**, 2852–2859.
4 B. Olenik, B. Keil and P. Jeschke, *Pest Manage. Sci.*, 2022, **78**, 2746–2758.
5 X. Niu, R. Yang, H. Zhang and J. Yang, *Advanced Agrochem*, 2022, **1**, 39–60.
6 Z. Hao and A. Iqbal, *Chem. Soc. Rev.*, 1997, **26**, 203.
7 C. Wang, H. Dong, L. Jiang and W. Hu, *Chem. Soc. Rev.*, 2018, **47**, 422–500.
8 J. D. Evans, K. E. Jelfs, G. M. Day and C. J. Doonan, *Chem. Soc. Rev.*, 2017, **46**, 3286–3301.
9 T. Friščić and L. R. MacGillivray, *Z. Kristallogr.*, 2005, **220**, 351–363.
10 P. Naumov, S. Chizhik, M. K. Panda, N. K. Nath and E. Boldyreva, *Chem. Rev.*, 2015, **115**, 12440–12490.

11 P. Naumov, D. P. Karothu, E. Ahmed, L. Catalano, P. Commins, J. Mahmoud Halabi, M. B. Al-Handawi and L. Li, *J. Am. Chem. Soc.*, 2020, **142**, 13256–13272.

12 W. M. Awad, D. W. Davies, D. Kitagawa, J. Mahmoud Halabi, M. B. Al-Handawi, I. Tahir, F. Tong, G. Campillo-Alvarado, A. G. Shtukenberg, T. Alkhidir, Y. Hagiwara, M. Almehairbi, L. Lan, S. Hasebe, D. P. Karothu, S. Mohamed, H. Koshima, S. Kobatake, Y. Diao, R. Chandrasekar, H. Zhang, C. C. Sun, C. Bardeen, R. O. Al-Kaysi, B. Kahr and P. Naumov, *Chem. Soc. Rev.*, 2023, **52**, 3098–3169.

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

14 R. A. Heinlein, Where To?, in *Expanded Universe, Ace Science Fiction*, New York, 1980, p. 348.

15 J. Maddox, *Nature*, 1988, **335**, 201.

16 J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta. Crystallogr. B.*, 2000, **56**, 697–714.

17 W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. A. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta. Crystallogr. B.*, 2002, **58**, 647–661.

18 G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta. Crystallogr. B.*, 2005, **61**, 511–527.

19 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. B.*, 2009, **65**, 107–125.

20 D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta. Crystallogr. B.*, 2011, **67**, 535–551.

21 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. B.*, 2016, **72**, 439–459.

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

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

24 J. Nyman and S. M. Reutzel-Edens, *Faraday Discuss.*, 2018, **211**, 459–476.

25 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.

26 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.

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

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

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

30 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.

31 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.

32 F. T. Szczypiński, S. Bennett and K. E. Jelfs, *Chem. Sci.*, 2021, **12**, 830–840.

33 A. M. Mroz, V. Posligua, A. Tarzia, E. H. Wolpert and K. E. Jelfs, *J. Am. Chem. Soc.*, 2022, **144**, 18730–18743.

34 Q. Zhu and S. Hattori, *J. Mater. Res.*, 2023, **38**, 19–36.

35 S. L. Price, *Int. Rev. Phys. Chem.*, 2008, **27**, 541–568.
36 S. L. Price, *Acc. Chem. Res.*, 2009, **42**, 117–126.
37 S. L. Price and L. S. Price, *Solid State Characterization of Pharmaceuticals*, Blackwell, 1st edn, 2011, pp. 427–450.
38 S. L. Price, *Chem. Soc. Rev.*, 2014, **43**, 2098–2111.
39 G. J. O. Beran, *Chem. Rev.*, 2016, **116**, 5567–5613.
40 S. L. Price, *Faraday Discuss.*, 2018, **211**, 9–30.
41 A. R. Oganov, *Faraday Discuss.*, 2018, **211**, 643–660.
42 A. R. Oganov, C. J. Pickard, Q. Zhu and R. J. Needs, *Nat. Rev. Mater.*, 2019, **4**, 331–348.
43 D. H. Bowskill, I. J. Sugden, S. Konstantinopoulos, C. S. Adjiman and C. C. Pantelides, *Annu. Rev. Chem. Biomol. Eng.*, 2021, **12**, 593–623.
44 R. J. Clements, J. Dickman, J. Johal, J. Martin, J. Glover and G. M. Day, *MRS Bull.*, 2022, **47**, 1054–1062.
45 J. Nyman and G. M. Day, *CrystEngComm*, 2015, **17**, 5154–5165.
46 A. J. Cruz-Cabeza and J. Bernstein, *Chem. Rev.*, 2014, **114**, 2170–2191.
47 J. Nyman and G. M. Day, *Phys. Chem. Chem. Phys.*, 2016, **18**, 31132–31143.
48 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.
49 E. C. Dybeck, N. S. Abraham, N. P. Schieber and M. R. Shirts, *Cryst. Growth Des.*, 2017, **17**, 1775–1787.
50 G. R. Woollam, M. A. Neumann, T. Wagner and R. J. Davey, *Faraday Discuss.*, 2018, **211**, 209–234.
51 S. L. Price, *Acta. Crystallogr. B.*, 2013, **69**, 313–328.
52 R. J. Davey, S. L. Schroeder and J. H. Ter Horst, *Angew. Chem., Int. Ed.*, 2013, **52**, 2166–2179.
53 F. Giberti, M. Salvalaglio and M. Parrinello, *IUCrJ*, 2015, **2**, 256–266.
54 J. E. Carpenter and M. Grünwald, *J. Am. Chem. Soc.*, 2021, **143**, 21580–21593.
55 J. Zhou, Y. Zhou and W. Tang, *Crystals*, 2022, **12**, 980.
56 B. P. van Eijck and J. Kroon, *J. Phys. Chem. B*, 1997, **101**, 1096–1100.
57 J. Pillardy, R. J. Wawak, Y. A. Arnautova, C. Czaplewski and H. A. Scheraga, *J. Am. Chem. Soc.*, 2000, 907–921.
58 B. P. van Eijck and J. Kroon, *J. Phys. Chem. B*, 2000, **104**, 8089.
59 W. Wedemeyer, Y. Arnautova, J. Pillardy, R. J. Wawak, C. Czaplewski and H. A. Scheraga, *J. Phys. Chem. B*, 2000, **104**, 8090–8092.
60 R. B. Hammond, K. Pencheva and K. J. Roberts, *J. Phys. Chem. B*, 2005, **109**, 19550–19552.
61 R. Hesterberg, P. Macchi and J. Hulliger, *Cryst. Growth Des.*, 2018, **18**, 7460–7469.
62 A. D. Mighell, V. L. Himes and J. R. Rodgers, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 1983, **39**, 737–740.
63 G. M. Day, *Crystallogr. Rev.*, 2011, **17**, 3–52.
64 C. J. Pickard and R. J. Needs, *J. Phys.: Condens.Matter*, 2011, **23**, 053201.
65 S. Fredericks, K. Parrish, D. Sayre and Q. Zhu, *Comput. Phys. Commun.*, 2021, **261**, 107810.
66 P. G. Karamertzanis and C. C. Pantelides, *J. Comput. Chem.*, 2005, **26**, 304–324.
67 P. G. Karamertzanis and C. C. Pantelides, *Mol. Phys.*, 2007, **105**, 273–291.
68 D. H. Case, J. E. Campbell, P. J. Bygrave and G. M. Day, *J. Chem. Theory Comput.*, 2016, **12**, 910–924.
69 P. Verwer and F. J. J. Leusen, *Rev. Comput. Chem.*, 1997, **12**, 327–365.
70 Y. Wang, J. Lv, L. Zhu and Y. Ma, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **82**, 094116.
71 Y. Wang and Y. Ma, *J. Chem. Phys.*, 2014, **140**, 040901.
72 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.
73 S. Yang and G. M. Day, *J. Chem. Theory Comput.*, 2021, **17**, 1988–1999.
74 Q. Zhu, A. R. Oganov, C. W. Glass and H. T. Stokes, *Acta. Crystallogr. B.*, 2012, **68**, 215–226.
75 F. Curtis, T. Rose and N. Marom, *Faraday Discuss.*, 2018, **211**, 61–77.
76 X. Li, F. S. Curtis, T. Rose, C. Schober, A. Vazquez-Mayagoitia, K. Reuter, H. Oberhofer and N. Marom, *J. Chem. Phys.*, 2018, **148**, 241701.
77 E. Zurek, *Rev. Comp. Chem.*, Wiley, New York, 2016, vol. 29, pp. 274–326.
78 M. A. Neumann, F. J. J. Leusen and J. Kendrick, *Angew. Chem., Int. Ed.*, 2008, **47**, 2427–2430.
79 J. Hermann, R. A. DiStasio and A. Tkatchenko, *Chem. Rev.*, 2017, **117**, 4714–4758.
80 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
81 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
82 E. Caldeweyher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2017, **147**, 034112.
83 R. A. DiStasio, O. A. von Lilienfeld and A. Tkatchenko, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14791–14795.
84 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.
85 A. Ambrosetti, A. M. Reilly, R. A. DiStasio and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.
86 A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **136**, 174109.
87 A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **137**, 054103.
88 A. M. Reilly and A. Tkatchenko, *J. Chem. Phys.*, 2013, **139**, 024705.
89 J. Moellmann and S. Grimme, *J. Phys. Chem. C*, 2014, **118**, 7615–7621.
90 J. G. Brandenburg, T. Maas and S. Grimme, *J. Chem. Phys.*, 2015, **142**, 124104.
91 S. R. Whittleton, A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 441–450.
92 S. R. Whittleton, A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Theory Comput.*, 2017, **13**, 5332–5342.
93 J. Hoja, A. M. Reilly and A. Tkatchenko, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2017, **7**, e1294.

94 J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio and A. Tkatchenko, *Sci. Adv.*, 2019, **5**, eaau3338.
95 J. Hoja and A. Tkatchenko, *Faraday Discuss.*, 2018, **211**, 253–274.
96 A. J. Price, A. Otero-De-La-Roza and E. R. Johnson, *Chem. Sci.*, 2022, **14**, 1252–1262.
97 A. J. Price, R. A. Mayo, A. Otero-De-la Roza and E. R. Johnson, *CrystEngComm*, 2023, **25**, 953–960.
98 O. A. Loboda, G. A. Dolgonos and A. D. Boese, *J. Chem. Phys.*, 2018, **149**, 124104.
99 M. A. Neumann, *J. Phys. Chem. B*, 2008, **112**, 9810–9829.
100 F. Musil, S. De, J. Yang, J. E. Campbell, G. M. Day and M. Ceriotti, *Chem. Sci.*, 2018, **9**, 1289–1300.
101 V. Kapil and E. A. Engel, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, 1–8.
102 L. Iuzzolino, P. McCabe, S. L. Price and J. G. Brandenburg, *Faraday Discuss.*, 2018, **211**, 275–296.
103 E. Caldeweyher and J. G. Brandenburg, *J. Phys.: Condens.Matter*, 2018, **30**, 213001.
104 L. Doná, J. G. Brandenburg and B. Civalleri, *J. Chem. Phys.*, 2019, **151**, 121101.
105 L. Donà, J. G. Brandenburg, I. J. Bush and B. Civalleri, *Faraday Discuss.*, 2020, **224**, 292–308.
106 S. Grimme, A. Hansen, S. Ehlert and J.-M. Mewes, *J. Chem. Phys.*, 2021, **154**, 064103.
107 E. Tuca, G. DiLabio and A. Otero-de-la Roza, *J. Chem. Inf. Model.*, 2022, **62**, 4107–4121.
108 G. A. Dolgonos, O. A. Loboda and A. D. Boese, *J. Phys. Chem. A*, 2018, **122**, 708–713.
109 C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.
110 S. Wengert, G. Csányi, K. Reuter and J. T. Margraf, *Chem. Sci.*, 2021, **12**, 4536–4546.
111 S. Wengert, G. Csányi, K. Reuter and J. T. Margraf, *J. Chem. Theory Comput.*, 2022, **18**, 4586–4593.
112 D. McDonagh, C.-K. Skylaris and G. M. Day, *J. Chem. Theory Comput.*, 2019, **15**, 2743–2758.
113 O. Egorova, R. Hafizi, D. C. Woods and G. M. Day, *J. Phys. Chem. A*, 2020, **124**, 8065–8078.
114 A. J. Stone and A. J. Misquitta, *Int. Rev. Phys. Chem.*, 2007, **26**, 193–222.
115 K. Szalewicz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 254–272.
116 R. Podeszwa, B. M. Rice and K. Szalewicz, *Phys. Rev. Lett.*, 2008, **101**, 115503.
117 R. Podeszwa, B. M. Rice and K. Szalewicz, *Phys. Chem. Chem. Phys.*, 2009, **11**, 5512–5518.
118 K. Szalewicz, *Acc. Chem. Res.*, 2014, **47**, 3266–3274.
119 A. J. Misquitta, G. W. Welch, A. J. Stone and S. L. Price, *Chem. Phys. Lett.*, 2008, **456**, 105–109.
120 G. J. O. Beran, S. Wen, K. Nanda, Y. Huang and Y. Heit, *Top. Curr. Chem.*, 2014, **345**, 59–93.
121 M. P. Metz, K. Piszczatowski and K. Szalewicz, *J. Chem. Theory Comput.*, 2016, **12**, 5895–5919.
122 M. P. Metz and K. Szalewicz, *J. Chem. Theory Comput.*, 2020, **16**, 2317–2339.
123 R. Nikhar and K. Szalewicz, *Nat. Commun.*, 2022, **13**, 1–9.
124 K. R. Bryenton, A. A. Adeleke, S. G. Dale and E. R. Johnson, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, 29–32.
125 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.*, 2012, **112**, 289–320.
126 C. Greenwell and G. J. O. Beran, *Cryst. Growth Des.*, 2020, **20**, 4875–4881.
127 A. Otero-de-la Roza, L. M. LeBlanc and E. R. Johnson, *J. Chem. Theory Comput.*, 2019, **15**, 4933–4944.
128 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.
129 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.
130 G. J. O. Beran, S. E. Wright, C. Greenwell and A. J. Cruz-Cabeza, *J. Chem. Phys.*, 2022, **156**, 104112.
131 C. Greenwell and G. J. O. Beran, *J. Mater. Chem. C*, 2021, **9**, 2848–2857.
132 G. J. O. Beran, *CrystEngComm*, 2019, **21**, 758–764.
133 T. J. Gately, W. Sontising, C. J. Easley, I. Islam, R. O. Al-Kaysi, G. J. O. Beran and C. J. Bardeen, *CrystEngComm*, 2021, **23**, 5931–5943.
134 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.
135 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.
136 C. J. Perry and G. J. O. Beran, *Cryst. Growth Des.*, 2023, **23**, 8352–8360.
137 C. J. Cook, C. J. Perry and G. J. O. Beran, *J. Phys. Chem. Lett.*, 2023, **14**, 6823–6831.
138 S. P. Thomas and M. A. Spackman, *Aust. J. Chem.*, 2018, **71**, 279.
139 J. Nyman, L. Yu and S. M. Reutzel-Edens, *CrystEngComm*, 2019, **21**, 2080–2088.
140 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.
141 B. Rana, G. J. O. Beran and J. M. Herbert, *Mol. Phys.*, 2022, e2138789.
142 R. P. Stoffel, C. Wessel, M.-W. Lumey and R. Dronskowski, *Angew. Chem., Int. Ed.*, 2010, **49**, 5242–5266.
143 A. Erba, *J. Chem. Phys.*, 2014, **141**, 124115.
144 N. Raimbault, V. Athavale and M. Rossi, *Phys. Rev. Mater.*, 2019, **3**, 1–11.
145 A. Erba, J. Maul and B. Civalleri, *Chem. Commun.*, 2016, **52**, 1820–1823.
146 Y. N. Heit and G. J. O. Beran, *Acta. Crystallogr. B.*, 2016, **72**, 514–529.
147 Y. N. Heit, K. D. Nanda and G. J. O. Beran, *Chem. Sci.*, 2016, **7**, 246–255.
148 C. Červinka, M. Fulem, R. P. Stoffel and R. Dronskowski, *J. Phys. Chem. A*, 2016, **120**, 2022–2034.
149 C. Červinka and M. Fulem, *J. Chem. Theory Comput.*, 2017, **13**, 2840–2850.

150 C. Červinka and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2017, **19**, 29940–29953.
151 C. Červinka and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2019, **21**, 14799–14810.
152 S. Hirata, K. Gilliard, X. He, J. Li and O. Sode, *Acc. Chem. Res.*, 2014, **47**, 2721–2730.
153 W. Sontising, Y. N. Heit, J. L. McKinley and G. J. O. Beran, *Chem. Sci.*, 2017, **8**, 7374–7382.
154 J. L. McKinley and G. J. O. Beran, *J. Chem. Theory Comput.*, 2019, **15**, 5259–5274.
155 C. Červinka and G. J. O. Beran, *Chem. Sci.*, 2018, **9**, 4622–4629.
156 C. Cook, J. L. McKinley and G. J. O. Beran, *J. Chem. Phys.*, 2021, **154**, 134109.
157 C. Cook and G. J. O. Beran, *J. Chem. Phys.*, 2020, **153**, 224105.
158 J. A. Weatherby, A. F. Rumson, A. J. A. Price, A. Otero de la Roza and E. R. Johnson, *J. Chem. Phys.*, 2022, **156**, 114108.
159 B. Monserrat, N. D. Drummond and R. J. Needs, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 144302.
160 E. A. Engel, B. Monserrat and R. J. Needs, *Phys. Rev. X*, 2015, **5**, 021033.
161 E. C. Dybeck, D. P. McMahon, G. M. Day and M. R. Shirts, *Cryst. Growth Des.*, 2019, **19**, 5568–5580.
162 N. S. Abraham and M. R. Shirts, *Cryst. Growth Des.*, 2019, **19**, 6911–6924.
163 E. C. Dybeck, N. P. Schieber and M. R. Shirts, *J. Chem. Theory Comput.*, 2016, **12**, 3491–3505.
164 N. P. Schieber, E. C. Dybeck and M. R. Shirts, *J. Chem. Phys.*, 2018, **148**, 144104.
165 M. Rossi, P. Gasparotto and M. Ceriotti, *Phys. Rev. Lett.*, 2016, **117**, 1–6.
166 B. Cheng, E. A. Engel, J. Behler, C. Dellago and M. Ceriotti, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 1110–1115.
167 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.
168 J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter and J. Morris, *Pharm. Res.*, 2001, **18**, 859–866.
169 D. A. McAfee, J. Hadgraft and M. E. Lane, *Eur. J. Pharm. Biopharm.*, 2014, **88**, 586–593.
170 I. B. Rietveld and R. Céolin, *J. Pharm. Sci.*, 2015, **104**, 4117–4122.
171 C. R. Taylor, M. T. Mulvee, D. S. Perenyi, M. R. Probert, G. M. Day and J. W. Steed, *J. Am. Chem. Soc.*, 2020, **142**, 16668–16680.
172 M. Mortazavi, J. Hoja, L. Aerts, L. Quéré, J. van de Streek, M. A. Neumann and A. Tkatchenko, *Commun. Chem.*, 2019, **2**, 70.
173 M. Lahn, S. Herbertz, J. S. Sawyer, A. J. Stauber, I. Gueorguieva, K. E. Driscoll, S. T. Estrem, A. L. Cleverly, D. Desaiah, S. C. Guba, K. A. Benhadji and C. A. Slapak, *Drug Des., Dev. Ther.*, 2015, 4479.
174 A. J. Cruz-Cabeza, G. M. Day and W. Jones, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12808.
175 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.
176 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.
177 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.
178 M. Habgood, R. Grau-Crespo and S. L. Price, *Phys. Chem. Chem. Phys.*, 2011, **13**, 9590–9600.
179 H. C. S. Chan, J. Kendrick, M. A. Neumann and F. J. J. Leusen, *CrystEngComm*, 2013, **15**, 3799.
180 C. R. Taylor and G. M. Day, *Cryst. Growth Des.*, 2018, **18**, 892–904.
181 N. Issa, P. G. Karamertzanis, G. W. A. Welch and S. L. Price, *Cryst. Growth Des.*, 2009, **9**, 442–453.
182 P. G. Karamertzanis, A. V. Kazantsev, N. Issa, G. W. A. Welch, C. S. Adjiman, C. C. Pantelides and S. L. Price, *J. Chem. Theory Comput.*, 2009, **5**, 1432–1448.
183 A. J. Cruz-Cabeza, G. M. Day and W. Jones, *Chem.–Eur. J.*, 2008, **14**, 8830–8836.
184 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.
185 D. E. Braun, P. G. Karamertzanis and S. L. Price, *Chem. Commun.*, 2011, **47**, 5443–5445.
186 E. C. Dybeck, A. Thiel, M. J. Schnieders, F. C. Pickard, G. P. Wood, J. F. Krzyzaniak and B. C. Hancock, *Cryst. Growth Des.*, 2023, **23**, 142–167.
187 D. E. Braun, R. M. Bhardwaj, A. J. Florence, D. A. Tocher and S. L. Price, *Cryst. Growth Des.*, 2013, **13**, 19–23.
188 D. E. Braun, H. Oberacher, K. Arnhard, M. Orlova and U. J. Griesser, *CrystEngComm*, 2016, **18**, 4053–4067.
189 R. S. Hong, A. Mattei, A. Y. Sheikh and M. E. Tuckerman, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, 1–8.
190 D. E. Braun, J. A. McMahon, L. H. Koztecki, S. L. Price and S. M. Reutzel-Edens, *Cryst. Growth Des.*, 2014, **14**, 2056–2072.
191 I. J. Sugden, D. E. Braun, D. H. Bowskill, C. S. Adjiman and C. C. Pantelides, *Cryst. Growth Des.*, 2022, **22**, 4513–4527.
192 M. Habgood, I. J. Sugden, A. V. Kazantsev, C. S. Adjiman and C. C. Pantelides, *J. Chem. Theory Comput.*, 2015, **11**, 1957–1969.
193 I. Sugden, C. S. Adjiman and C. C. Pantelides, *Acta. Crystallogr. B.*, 2016, **72**, 864–874.
194 I. J. Sugden, C. S. Adjiman and C. C. Pantelides, *Acta. Crystallogr. B.*, 2019, **75**, 423–433.
195 L. Yu, *Acc. Chem. Res.*, 2010, **43**, 1257–1266.
196 G. Stephenson, T. Borchardt, S. Byrn, J. Bowyer, C. Bunnell, S. Snorek and L. Yu, *J. Pharm. Sci.*, 1995, **84**, 1385–1386.
197 L. Yu, G. A. Stephenson, C. A. Mitchell, C. A. Bunnell, S. V. Snorek, J. J. Bowyer, T. B. Borchardt, J. G. Stowell and S. R. Byrn, *J. Am. Chem. Soc.*, 2000, **122**, 585–591.

198 S. Chen, I. A. Guzei and L. Yu, *J. Am. Chem. Soc.*, 2005, **127**, 9881–9885.
199 S. Chen, H. Xi and L. Yu, *J. Am. Chem. Soc.*, 2005, **127**, 17439–17444.
200 M. Tan, A. G. Shtukenberg, S. Zhu, W. Xu, E. Dooryhee, S. M. Nichols, M. D. Ward, B. Kahr and Q. Zhu, *Faraday Discuss.*, 2018, **211**, 477–491.
201 K. S. Gushurst, J. Nyman and S. X. M. Boerrigter, *CrystEngComm*, 2019, **21**, 1363–1368.
202 X. Li, X. Ou, H. Rong, S. Huang, J. Nyman, L. Yu and M. Lu, *Cryst. Growth Des.*, 2020, **20**, 7093–7097.
203 A. Lévesque, T. Maris and J. D. Wuest, *J. Am. Chem. Soc.*, 2020, **142**, 11873–11883.
204 A. R. Tyler, R. Ragbirsingh, C. J. McMonagle, P. G. Waddell, S. E. Heaps, J. W. Steed, P. Thaw, M. J. Hall and M. R. Probert, *Chem*, 2020, **6**, 1755–1765.
205 C. A. Mitchell, L. Yu and M. D. Ward, *J. Am. Chem. Soc.*, 2001, **123**, 10830–10839.
206 M. Vasileiadis, A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman and C. C. Pantelides, *Acta. Crystallogr. B.*, 2012, **68**, 677–685.
207 B. A. Nogueira, C. Castiglioni and R. Fausto, *Commun. Chem.*, 2020, **3**, 34.
208 X. Feng, A. D. Becke and E. R. Johnson, *CrystEngComm*, 2021, **23**, 4264–4271.
209 L. R. Warren, E. McGowan, M. Renton, C. A. Morrison and N. P. Funnell, *Chem. Sci.*, 2021, **12**, 12711–12718.
210 E. L. Harty, A. R. Ha, M. R. Warren, A. L. Thompson, D. R. Allan, A. L. Goodwin and N. P. Funnell, *Chem. Commun.*, 2015, **51**, 10608–10611.
211 C. Greenwell, J. Řezáč and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2022, **24**, 3695–3712.
212 N. P. Funnell, C. L. Bull, C. J. Ridley and S. Capelli, *CrystEngComm*, 2019, **21**, 4473–4483.
213 A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper and G. M. Day, *Nature*, 2017, **543**, 657–664.
214 G. J. O. Beran, *Nat. Mater.*, 2017, **16**, 602–604.
215 G. M. Day and A. I. Cooper, *Adv. Mater.*, 2018, **30**, 1704944.
216 A. G. Slater, P. S. Reiss, A. Pulido, M. A. Little, D. L. Holden, L. Chen, S. Y. Chong, B. M. Alston, R. Clowes, M. Haranczyk, M. E. Briggs, T. Hasell, G. M. Day and A. I. Cooper, *ACS Cent. Sci.*, 2017, **3**, 734–742.
217 C. Zhao, L. Chen, Y. Che, Z. Pang, X. Wu, Y. Lu, H. Liu, G. M. Day and A. I. Cooper, *Nat. Commun.*, 2021, **12**, 1–11.
218 J. E. Campbell, J. Yang and G. M. Day, *J. Mater. Chem. C*, 2017, **5**, 7574–7584.
219 J. Yang, S. De, J. E. Campbell, S. Li, M. Ceriotti and G. M. Day, *Chem. Mater.*, 2018, **30**, 4361–4371.
220 C. Y. Cheng, J. E. Campbell and G. M. Day, *Chem. Sci.*, 2020, **11**, 4922–4933.
221 B. Rice, L. M. LeBlanc, A. Otero-de-la Roza, M. J. Fuchter, E. R. Johnson, J. Nelson and K. E. Jelfs, *Nanoscale*, 2018, **10**, 1865–1876.
222 R. L. Greenaway, V. Santolini, M. J. Bennison, B. M. Alston, C. J. Pugh, M. A. Little, M. Miklitz, E. G. B. Eden-Rump, R. Clowes, A. Shakil, H. J. Cuthbertson, H. Armstrong, M. E. Briggs, K. E. Jelfs and A. I. Cooper, *Nat. Commun.*, 2018, **9**, 2849.
223 R. L. Greenaway, V. Santolini, A. Pulido, M. A. Little, B. M. Alston, M. E. Briggs, G. M. Day, A. I. Cooper and K. E. Jelfs, *Angew. Chem., Int. Ed.*, 2019, **58**, 16275–16281.
224 E. H. Wolpert and K. E. Jelfs, *Chem. Sci.*, 2022, **13**, 13588–13599.
225 R. O. Al-Kaysi, A. M. Müller and C. J. Bardeen, *J. Am. Chem. Soc.*, 2006, **128**, 15938–15939.
226 S. Kobatake, S. Takami, H. Muto, T. Ishikawa and M. Irie, *Nature*, 2007, **446**, 778–781.
227 M. Morimoto and M. Irie, *J. Am. Chem. Soc.*, 2010, **132**, 14172–14178.
228 T. Kim, L. Zhu, R. O. Al-Kaysi and C. J. Bardeen, *ChemPhysChem*, 2014, **15**, 400–414.
229 X. Dong, F. Tong, K. M. Hanson, R. O. Al-Kaysi, D. Kitagawa, S. Kobatake and C. J. Bardeen, *Chem. Mater.*, 2019, **31**, 1016–1022.
230 F. Tong, W. Xu, T. Guo, B. F. Lui, R. C. Hayward, P. Palffy-Muhoray, R. O. Al-Kaysi and C. J. Bardeen, *J. Mater. Chem. C*, 2020, **8**, 5036–5044.
231 E. Uchida, R. Azumi and Y. Norikane, *Nat. Commun.*, 2015, **6**, 7310.
232 F. Tong, D. Kitagawa, I. Bushnak, R. O. Al-Kaysi, C. J. Bardeen, R. O. Al-Kaysi and C. J. Bardeen, *Angew. Chem., Int. Ed.*, 2021, **60**, 2414–2423.
233 P. J. Bygrave, D. H. Case and G. M. Day, *Faraday Discuss.*, 2014, **170**, 41–57.
234 S. SeethaLekshmi, T. S. Thakur and S. Varughese, *J. Photochem. Photobiol., C*, 2021, **49**, 100455.
235 L. R. MacGillivray, G. S. Papaefstathiou, T. Friscić, T. D. Hamilton, D.-K. Bucar, Q. Chu, D. B. Varshney and I. G. Georgiev, *Acc. Chem. Res.*, 2008, **41**, 280–291.
236 A. J. Duncan, R. L. Dudovitz, S. J. Dudovitz, J. Stojaković, S. V. Mariappan and L. R. MacGillivray, *Chem. Commun.*, 2016, **52**, 13109–13111.
237 Q. Chu, A. J. Duncan, G. S. Papaefstathiou, T. D. Hamilton, M. B. Atkinson, S. V. Mariappan and L. R. Macgillivray, *J. Am. Chem. Soc.*, 2018, **140**, 4940–4944.
238 G. Kaupp, *Organic Solid-State Reactions*, 2016, DOI: **10.1002/9781118468586.epoc2005**.
239 J. G. Hernández and C. Bolm, *J. Org. Chem.*, 2017, **82**, 4007–4019.
240 D. Tan and T. Friščić, *Eur. J. Org Chem.*, 2018, **2018**, 18–33.
241 M. Winkler and K. N. Houk, *J. Am. Chem. Soc.*, 2007, **129**, 1805–1815.
242 J. A. Schmidt, J. A. Weatherby, I. J. Sugden, A. Santana-Bonilla, F. Salerno, M. J. Fuchter, E. R. Johnson, J. Nelson and K. E. Jelfs, *Cryst. Growth Des.*, 2021, **21**, 5036–5049.
243 M.-A. Perrin, M. A. Neumann, H. Elmaleh and L. Zaske, *Chem. Commun.*, 2009, 3181–3183.
244 Q. Zhu, A. G. Shtukenberg, D. J. Carter, T.-Q. Yu, J. Yang, M. Chen, P. Raiteri, A. R. Oganov, B. Pokroy, I. Polishchuk, P. J. Bygrave, G. M. Day, A. L. Rohl,

M. E. Tuckerman and B. Kahr, *J. Am. Chem. Soc.*, 2016, **138**, 4881–4889.
245 A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, J. Hoja, E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman and B. Kahr, *Chem. Sci.*, 2017, **8**, 4926–4940.
246 A. G. Shtukenberg, C. T. Hu, Q. Zhu, M. U. Schmidt, W. Xu, M. Tan and B. Kahr, *Cryst. Growth Des.*, 2017, **17**, 3562–3566.
247 A. G. Shtukenberg, M. Tan, L. Vogt-Maranto, E. J. Chan, W. Xu, J. Yang, M. E. Tuckerman, C. T. Hu and B. Kahr, *Cryst. Growth Des.*, 2019, **19**, 4070–4080.
248 W. Sontising and G. J. O. Beran, *Phys. Rev. Mater.*, 2019, **3**, 095002.
249 W. Sontising and G. J. O. Beran, *Phys. Rev. Mater.*, 2020, **4**, 063601.
250 M. D. Eddleston, K. E. Hejczyk, E. G. Bithell, G. M. Day and W. Jones, *Chem. – Eur. J.*, 2013, **19**, 7874–7882.
251 S. Datta and D. J. W. Grant, *Nat. Rev. Drug Discovery*, 2004, **3**, 42–57.
252 S. E. Ashbrook and D. McKay, *Chem. Commun.*, 2016, **52**, 7186–7204.
253 E. Salager, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2610.
254 E. Salager, G. M. Day, R. S. Stein, C. J. Pickard, B. Elena and L. Emsley, *J. Am. Chem. Soc.*, 2010, **132**, 2564–2566.
255 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.
256 M. Baias, J.-N. Dumez, P. H. Svensson, S. Schantz, G. M. Day and L. Emsley, *J. Am. Chem. Soc.*, 2013, **135**, 17501–17507.
257 X. Li, M. A. Neumann and J. van de Streek, *IUCrJ*, 2017, **4**, 175–184.
258 A. Hofstetter, M. Balodis, F. M. Paruzzo, C. M. Widdifield, G. Stevanato, A. C. Pinon, P. J. Bygrave, G. M. Day and L. Emsley, *J. Am. Chem. Soc.*, 2019, **141**, 16624–16634.
259 M. K. Dudek, P. Paluch, J. Śniechowska, K. P. Nartowski, G. M. Day and M. J. Potrzebowski, *CrystEngComm*, 2020, **22**, 4969–4981.
260 M. Cordova, M. Balodis, A. Hofstetter, F. Paruzzo, S. O. Nilsson Lill, E. S. E. Eriksson, P. Berruyer, B. Simões de Almeida, M. J. Quayle, S. T. Norberg, A. Svensk Ankarberg, S. Schantz and L. Emsley, *Nat. Commun.*, 2021, **12**, 2964.
261 M. Balodis, M. Cordova, A. Hofstetter, G. M. Day and L. Emsley, *J. Am. Chem. Soc.*, 2022, **144**, 7215–7223.
262 F. Bravetti, S. Bordignon, E. Alig, D. Eisenbeil, L. Fink, C. Nervi, R. Gobetto, M. U. Schmidt and M. R. Chierotti, *Chem.– Eur. J.*, 2022, **28**, e202103589.
263 J. D. Hartman, R. A. Kudla, G. M. Day, L. J. Mueller and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21686–21709.
264 J. D. Hartman, A. Balaji and G. J. O. Beran, *J. Chem. Theory Comput.*, 2017, **13**, 6043–6051.
265 M. Dračínský, P. Unzueta and G. J. O. Beran, *Phys. Chem. Chem. Phys.*, 2019, **21**, 14992–15000.
266 R. J. Iuliucci, J. D. Hartman and G. J. O. Beran, *J. Phys. Chem. A*, 2023, **127**, 2846–2858.
267 J. D. Hartman, G. M. Day and G. J. O. Beran, *Cryst. Growth Des.*, 2016, **16**, 6479–6493.
268 D. Chen, Z. Wang, D. Guo, V. Orekhov and X. Qu, *Chem.–Eur. J.*, 2020, **26**, 10391–10401.
269 E. Jonas, S. Kuhn and N. Schlörer, *Magn. Reson. Chem.*, 2022, **60**, 1021–1031.
270 I. Cortés, C. Cuadrado, A. Hernández Daranas and A. M. Sarotti, *Front. Nat. Prod. Res.*, 2023, **2**, 1–11.
271 M. Rupp, R. Ramakrishnan and O. A. von Lilienfeld, *J. Phys. Chem. Lett.*, 2015, **6**, 3309–3313.
272 F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti and L. Emsley, *Nat. Commun.*, 2018, **9**, 4501.
273 M. Cordova, E. A. Engel, A. Stefaniuk, F. Paruzzo, A. Hofstetter, M. Ceriotti and L. Emsley, *J. Phys. Chem. C*, 2022, **126**, 16710–16720.
274 S. Liu, J. Li, K. C. Bennett, B. Ganoe, T. Stauch, M. Head-Gordon, A. Hexemer, D. Ushizima and T. Head-Gordon, *J. Phys. Chem. Lett.*, 2019, **10**, 4558–4565.
275 E. Jonas and S. Kuhn, *J. Cheminf.*, 2019, **11**, 50.
276 W. Gerrard, L. A. Bratholm, M. J. Packer, A. J. Mulholland, D. R. Glowacki and C. P. Butts, *Chem. Sci.*, 2020, **11**, 508–515.
277 W. Gerrard, C. Yiu and C. P. Butts, *Magn. Reson. Chem.*, 2022, **60**, 1087–1092.
278 P. Gao, J. J. J. J. Zhang, Q. Peng, J. J. J. J. Zhang and V.-A. Glezakou, *J. Chem. Inf. Model.*, 2020, **60**, 3746–3754.
279 H. Han and S. Choi, *J. Phys. Chem. Lett.*, 2021, **12**, 3662–3668.
280 Z. Yang, M. Chakraborty and A. D. White, *Chem. Sci.*, 2021, **12**, 10802–10809.
281 Y. Guan, S. V. Shree Sowndarya, L. C. Gallegos, P. C. S. John and R. S. Paton, *Chem. Sci.*, 2021, **12**, 12012–12026.
282 P. A. Unzueta, C. Greenwell and G. J. O. Beran, *J. Chem. Theory Comput.*, 2021, **17**, 826–840.
283 A. Gupta, S. Chakraborty and R. Ramakrishnan, *Mach. Learn.: Sci. Technol.*, 2021, **2**, 035010.
284 J. B. Kleine Büning and S. Grimme, *J. Chem. Theory Comput.*, 2023, **19**, 3601–3615.
285 H. Suzuki, D. Matsubara, Y. Nakata, M. Ito and S. Noguchi, *Chem. Pharm. Bull.*, 2022, **70**, 731–734.
286 S. Boobier, D. R. Hose, A. J. Blacker and B. N. Nguyen, *Nat. Commun.*, 2020, **11**, 1–10.
287 Z. Ye and D. Ouyang, *J. Cheminf.*, 2021, **13**, 1–13.
288 F. H. Vermeire, Y. Chung and W. H. Green, *J. Am. Chem. Soc.*, 2022, **144**, 10785–10797.
289 A. D. Vassileiou, M. N. Robertson, B. G. Wareham, M. Soundaranathan, S. Ottoboni, A. J. Florence, T. Hartwig and B. F. Johnston, *Digital Discovery*, 2023, **2**, 356–367.
290 Z. Bjelobrk, D. Mendels, T. Karmakar, M. Parrinello and M. Mazzotti, *Cryst. Growth Des.*, 2021, **21**, 5198–5205.
291 D. S. Palmer, J. L. McDonagh, J. B. Mitchell, T. Van Mourik and M. V. Fedorov, *J. Chem. Theory Comput.*, 2012, **8**, 3322–3337.

292 Y. A. Abramov, G. Sun, Q. Q. Q. Zeng, Q. Q. Q. Zeng and M. Yang, *Mol. Pharmaceutics*, 2020, **17**, 666–673.
293 D. J. Fowles, D. S. Palmer, R. Guo, S. L. Price and J. B. O. Mitchell, *J. Chem. Theory Comput.*, 2021, **17**, 3700–3709.
294 J. Park, I. Nessler, B. Mcclain, D. Macikenas, J. Baltrusaitis and M. J. Schnieders, *J. Chem. Theory Comput.*, 2014, **10**, 2781–2791.
295 E. Kiely, R. Zwane, R. Fox, A. M. Reilly and S. Guerin, *CrystEngComm*, 2021, **23**, 5697–5710.
296 T. Beyer, G. M. Day and S. L. Price, *J. Am. Chem. Soc.*, 2001, **123**, 5086–5094.
297 S. Karki, T. Friščić, L. Fábián, P. R. Laity, G. M. Day and W. Jones, *Adv. Mater.*, 2009, **21**, 3905–3909.
298 C. Wang and C. C. Sun, *Mol. Pharmaceutics*, 2019, **16**, 1732–1741.
299 C. Wang, S. Paul, D. J. Sun, S. O. Nilsson Lill and C. C. Sun, *Cryst. Growth Des.*, 2020, **20**, 4217–4223.
300 A. M. Reilly and A. Tkatchenko, *Phys. Rev. Lett.*, 2014, **113**, 055701.
301 I. Azuri, E. Meirzadeh, D. Ehre, S. R. Cohen, A. M. Rappe, M. Lahav, I. Lubomirsky and L. Kronik, *Angew. Chem., Int. Ed.*, 2015, **54**, 13566–13570.
302 V. Basavalingappa, S. Bera, B. Xue, I. Azuri, Y. Tang, K. Tao, L. J. Shimon, M. R. Sawaya, S. Kolusheva, D. S. Eisenberg, L. Kronik, Y. Cao, G. Wei and E. Gazit, *Nat. Commun.*, 2019, **10**, 1–11.
303 F. Colmenero, *Mater. Adv.*, 2020, **1**, 1399–1426.
304 G. M. Day, S. L. Price and M. Leslie, *Cryst. Growth Des.*, 2001, **1**, 13–27.
305 M. Destefanis, C. Ravoux, A. Cossard and A. Erba, *Minerals*, 2019, **9**, 1–17.
306 P. A. Banks, J. Maul, M. T. Mancini, A. C. Whalley, A. Erba and M. T. Ruggiero, *J. Mater. Chem. C*, 2020, **8**, 10917–10925.
307 X. Bidault and S. Chaudhuri, *J. Chem. Phys.*, 2021, **154**, 164105.
308 P. R. Spackman, A. Grosjean, S. P. Thomas, D. P. Karothu, P. Naumov and M. A. Spackman, *Angew. Chem., Int. Ed.*, 2022, **61**, e202110716.
309 D. H. Case, V. K. Srirambhatla, R. Guo, R. E. Watson, L. S. Price, H. Polyzois, J. K. Cockcroft, A. J. Florence, D. A. Tocher and S. L. Price, *Cryst. Growth Des.*, 2018, **18**, 5322–5331.
310 J.-B. Arlin, L. S. Price, S. L. Price and A. J. Florence, *Chem. Commun.*, 2011, **47**, 7074.
311 V. K. Srirambhatla, R. Guo, S. L. Price and A. J. Florence, *Chem. Commun.*, 2016, **52**, 7384–7386.
312 D. K. Bučar, G. M. Day, I. Halasz, G. G. Zhang, J. R. Sander, D. G. Reid, L. R. MacGillivray, M. J. Duer and W. Jones, *Chem. Sci.*, 2013, **4**, 4417–4425.
313 P. Raiteri, R. Martonák and M. Parrinello, *Angew. Chem., Int. Ed.*, 2005, **44**, 3769–3773.
314 T. Zykova-Timan, P. Raiteri and M. Parrinello, *J. Phys. Chem. B*, 2008, **112**, 13231–13237.
315 P. G. Karamertzanis, P. Raiteri, M. Parrinello, M. Leslie and S. L. Price, *J. Phys. Chem. B*, 2008, **112**, 4298–4308.
316 V. Buch, R. Martonák and M. Parrinello, *J. Chem. Phys.*, 2005, **123**, 051108.
317 V. Buch, R. Martonák and M. Parrinello, *J. Chem. Phys.*, 2006, **124**, 204705.
318 P. M. Piaggi and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 10251–10256.
319 T.-Q. Yu and M. Tuckerman, *Phys. Rev. Lett.*, 2011, **107**, 015701.
320 T.-Q. Yu, P.-Y. Chen, M. Chen, A. Samanta, E. Vanden-Eijnden and M. Tuckerman, *J. Chem. Phys.*, 2014, **140**, 214109.
321 E. Schneider, L. Vogt and M. E. Tuckerman, *Acta. Crystallogr. B.*, 2016, **72**, 542–550.
322 H. Song, L. Vogt-Maranto, R. Wiscons, A. J. Matzger and M. E. Tuckerman, *J. Phys. Chem. Lett.*, 2020, **11**, 9751–9758.
323 N. F. Francia, L. S. Price and M. Salvalaglio, *CrystEngComm*, 2021, **23**, 5575–5584.
324 N. F. Francia, L. S. Price, J. Nyman, S. L. Price and M. Salvalaglio, *Cryst. Growth Des.*, 2020, **20**, 6847–6862.
325 T. T. Jensen, C. L. Hall, J. Potticary, I. Andrusenko, M. Gemmi and S. R. Hall, *Chem. Commun.*, 2019, **55**, 14586–14589.
326 I. J. Sugden, N. F. Francia, T. Jensen, C. S. Adjiman and M. Salvalaglio, *CrystEngComm*, 2022, **24**, 6830–6838.
327 P. W. Butler and G. M. Day, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, 1–8.
328 S. Yang and G. M. Day, *Commun. Chem.*, 2022, **5**, 1–13.
329 A. J. Cruz-Cabeza, N. Feeder and R. J. Davey, *Commun. Chem.*, 2020, **3**, 10–13.
330 H. P. G. Thompson and G. M. Day, *Chem. Sci.*, 2014, **5**, 3173–3182.
331 Y. A. Abramov, P. Zhang, Q. Zeng, M. Yang, Y. Liu and S. Sekharan, *Cryst. Growth Des.*, 2020, **20**, 1512–1525.
332 R. Montis, R. J. Davey, S. E. Wright, G. R. Woollam and A. J. Cruz-Cabeza, *Angew. Chem., Int. Ed.*, 2020, **59**, 20357–20360.
333 A. Anelli, E. A. Engel, C. J. Pickard and M. Ceriotti, *Phys. Rev. Mater.*, 2018, **2**, 1–8.
334 S. A. Bonev, F. Gygi, T. Ogitsu and G. Galli, *Phys. Rev. Lett.*, 2003, **91**, 065501.
335 X.-D. Wen, R. Hoffmann and N. W. Ashcroft, *J. Am. Chem. Soc.*, 2011, **133**, 9023–9035.
336 J. Li, O. Sode, G. A. Voth and S. Hirata, *Nat. Commun.*, 2013, **4**, 2647.
337 L. Huang, Y. Han, X. He and J. Li, *RSC Adv.*, 2020, **10**, 236–243.
338 Q. Lu, I. Ali and J. Li, *New J. Chem.*, 2020, **44**, 21012–21020.
339 Y. Abramov, G. Sun, Y. Zhou, M. Yang, Q. Zeng and Z. Shen, *Cryst. Growth Des.*, 2019, **19**, 7132–7137.
340 D. E. Braun, L. H. Koztecki, J. A. McMahon, S. L. Price and S. M. Reutzel-Edens, *Mol. Pharmaceutics*, 2015, **12**, 3069–3088.
341 M. Kilgour, J. Rogal and M. Tuckerman, *J. Chem. Theory Comput.*, 2023, **19**, 4743–4756.
342 R. K. Cersonsky, M. Pakhnova, E. A. Engel and M. Ceriotti, *Chem. Sci.*, 2023, **14**, 1272–1285.