

**Sarah L. Price**

Department of Chemistry, University College  
London, 20 Gordon Street, London WC1H 0AJ,  
England

Correspondence e-mail: s.l.price@ucl.ac.uk

# Why don't we find more polymorphs?

Received 8 May 2013

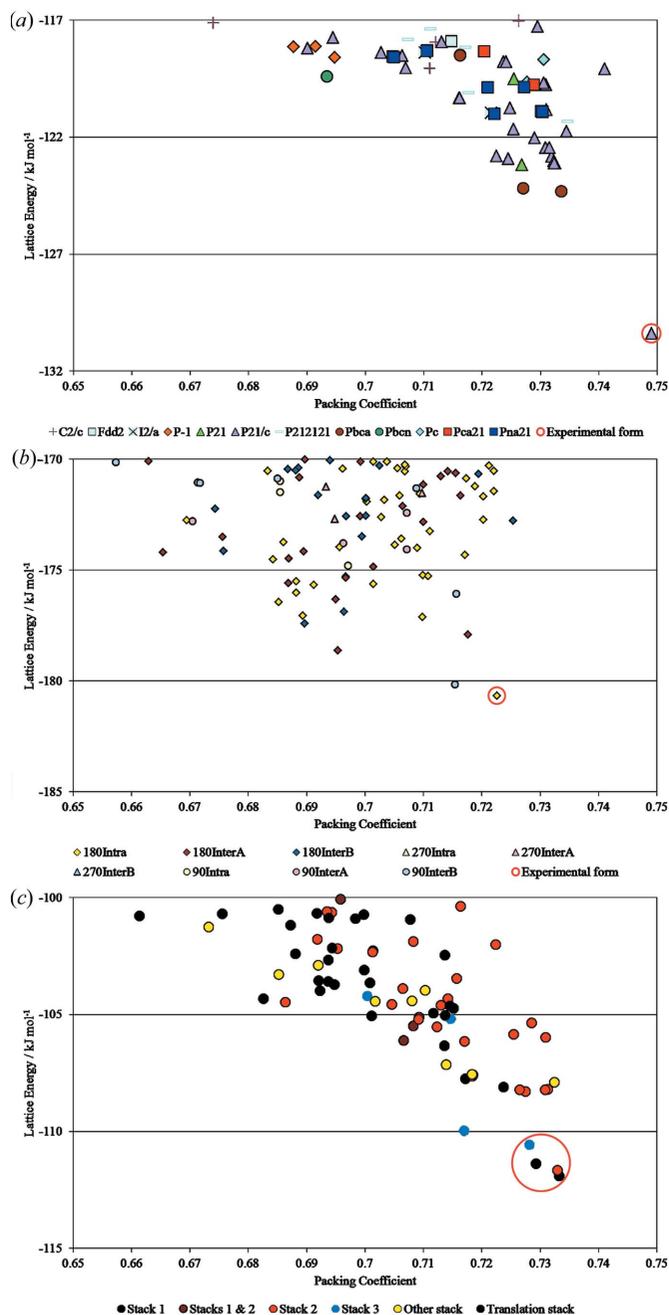
Accepted 8 July 2013

Crystal structure prediction (CSP) studies are not limited to being a search for the most thermodynamically stable crystal structure, but play a valuable role in understanding polymorphism, as shown by interdisciplinary studies where the crystal energy landscape has been explored experimentally and computationally. CSP usually produces more thermodynamically plausible crystal structures than known polymorphs. This article illustrates some reasons why: because (i) of approximations in the calculations, particularly the neglect of thermal effects (see §1.1); (ii) of the molecular rearrangement during nucleation and growth (see §1.2); (iii) the solid-state structures observed show dynamic or static disorder, stacking faults, other defects or are not crystalline and so represent more than one calculated structure (see §1.3); (iv) the structures are metastable relative to other molecular compositions (see §1.4); (v) the right crystallization experiment has not yet been performed (see §1.5) or (vi) cannot be performed (see §1.6) and the possibility (vii) that the polymorphs are not detected or structurally characterized (see §1.7). Thus, we can only aspire to a general predictive theory for polymorphism, as this appears to require a quantitative understanding of the kinetic factors involved in all possible multi-component crystallizations. For a specific molecule, analysis of the crystal energy landscape shows the potential complexity of its crystallization behaviour.

## 1. Introduction

Organic crystal structure prediction (CSP) software has not advanced conceptually since its early days when we were seeking to predict THE crystal structure from the chemical diagram (Day, 2011; Price, 2008*b*). It was based on the hypothesis that the crystal structure was the most thermodynamically stable structure, and at that time the hope was that this structure would be significantly more stable than any other (Fig. 1*a*). Algorithm design concentrated on finding effective methods of searching through the vast multi-dimensional surface for the most stable structure, which is a major task even if the search is restricted to the most common space groups and packing arrangements with one small rigid molecule in the asymmetric unit cell. The unexpected emergence of crystals of a new polymorph of the anti-HIV drug ritonavir a few years into manufacture (Bauer *et al.*, 2001) forcibly reminded us that metastable polymorphs could appear to be kinetically stable, and hence CSP methods could have a valuable practical application in providing reassurance that the most stable polymorph was known.

Polymorphism, the ability of a molecule to crystallize in more than one structure, is important because of the difference in physical properties between the polymorphs. These



**Figure 1**

Contrasting results of CSP studies, showing the lattice energies of the generated structures against the packing coefficient (Gavezzotti, 1983), with the observed structures circled in red. (a) Only one structure is thermodynamically plausible. The experimental structure of isocaffeine (XVb) is clearly predicted. (b) Whilst the observed structure is the most stable, there are other structures on the crystal energy landscape requiring consideration as to whether they could be observed polymorphs. The crystal structures of GSK269984B (XI) are classified by their conformation, a torsion angle that strongly affects the gross conformation and the carboxylic acid conformation. Since the thermodynamically competitive structures have very different gross conformations, these crystal structures would be expected to be long-lived if they crystallized (Ismail *et al.*, 2013). (c) There is a cluster of layer crystal structures around the global minimum, requiring analysis as to whether they could crystallize as distinct polymorphs or might lead to disorder. Both polymorphs of caffeine (XVa) are disordered layer structures. The crystal structures are classified according to their molecular stacking (Habgood, 2011).

differences can affect the uses, processing and characterization of the crystalline material (Bernstein, 2002; Brittain, 2009; Hilfiker, 2006; Storey & Ymén, 2012). The prevalence of polymorphism, as judged from thermal analysis, vibrational spectroscopy or powder X-ray diffraction and with modern screening methods (Stahly, 2007), is an order of magnitude greater than would be estimated from the Cambridge Structural Database (CSD; Allen, 2002; van de Streek & Motherwell, 2005). Despite considerable advances in the ability to solve structures from smaller, more weakly diffracting crystals and powders, the CSD is still dominated by reports of the structure of the first crystal suitable for single-crystal diffraction experiments that could be grown: a bias that is worth noting considering that CSD surveys are often used in CSP to justify reducing the search to common space groups,  $Z' = 1$  (one molecule in the asymmetric unit cell) or limit the range of molecular conformations considered. Determining the crystal structures of metastable polymorphs can be difficult: one of the most highly studied polymorphic systems, nicknamed ROY [(I), for all molecular diagrams see Fig. 2] because of the red–orange–yellow colours of the seven polymorphs whose crystal structures are known, has at least three further polymorphs (Yu, 2010), although the experimental evidence is not sufficient to tell whether they are amongst the unobserved low-energy structures on the crystal energy landscape (Vasileiadis *et al.*, 2012). The increasing desire to explore the range of crystal forms of a given molecule, in the search for new materials with improved properties or for intellectual property considerations, has led to the development of a ‘dazzling panoply’ (Bernstein, 2011) of new techniques for growing crystals with the aim of obtaining new forms (Llinàs & Goodman, 2008) or obtaining better control over the polymorph produced (Chen *et al.*, 2011).

The blind tests of our ability to predict a crystal structure from the chemical diagram, organized by the Cambridge Crystallographic Data Centre (Bardwell *et al.*, 2011), clearly show that the search for the most stable structure almost always generates structures that are mechanically stable, local minima in the lattice energy sufficiently close in energy to the global minimum to be plausible polymorphs. Indeed, it is the exceptions that ‘prove the rule’. For example, the structure of the only polymorph of pigment yellow 74 (II) has such good packing of the bumps into the hollows that this irregular, rigid molecule forms a surprisingly well close packed plane, and the large energy gap (about  $12 \text{ kJ mol}^{-1}$ ) shows that the electrostatic forces strongly favour one mode of stacking (Schmidt, 1999). That it is unusual for a molecule to have one unique, strongly favourable way of packing with itself can also be seen from comparing the CSP generated packings in chiral and non-chiral space groups (D’Oría *et al.*, 2010). An analysis for three molecules that spontaneously resolved showed that the energetic favourability of the observed structures was very small – although a molecule’s shape and interactions can favour a chiral packing in one or two dimensions, it is hard to imagine a molecule which can close pack with strong interactions giving a favourable chiral packing in all three dimensions. For example, a molecule’s structure may favour the

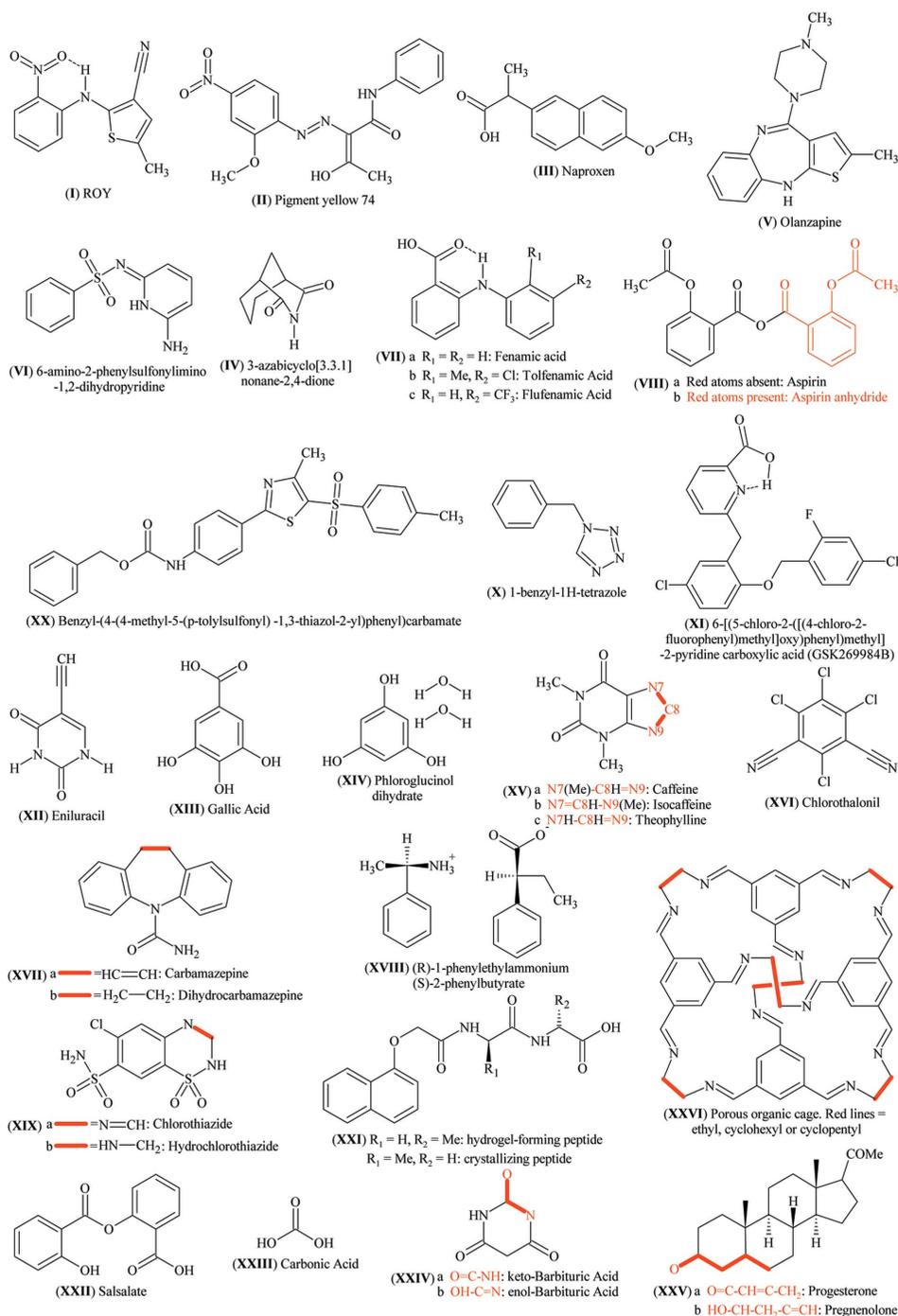
formation of helical columns, but spontaneous resolution requires that the functional groups on the outside of these columns also favour chiral packing in both other dimensions. Crystal engineers (Desiraju, 2007) acknowledge that the problem in designing a one-component organic crystal is to find molecules which have comparable (and strong) interactions to determine all three cell dimensions and close pack. Thus, the fields of crystal engineering and self-assembly are

dominated by multicomponent systems (Ward & Raithby, 2013).

Hence, the crystal energy landscape, the set of crystal structures generated by CSP methods which are sufficiently low in energy to be thermodynamically plausible as potential polymorphs, rarely unambiguously predicts that the molecule can only crystallize in one structure. What the calculation does provide is the set of structures in which different compromises

between all the intermolecular interactions and the conformational flexibility provide approximately the same, most favourable crystal energy. The CSP procedure with an adequately exhaustive search generates all reasonable packing modes for the given molecule, the analysis of which can yield very valuable information about the favoured modes of packing of the molecule in its various solid forms (Carlucci & Gavezzotti, 2005; Gavezzotti, 2002). The variation in the number of structures, their energy separation and most importantly, the nature of the compromises, depends on the specific molecule: some of the molecule–molecule contacts in the first coordination shell are likely to be only weakly stabilizing and hence readily varied without significant change in energy. Indeed, an estimated 9% of crystal structures for non-charged compounds contain a destabilizing molecule–molecule contact (Gavezzotti, 2010). Hence, the majority of crystal energy landscapes contain a variety of crystal packings that are quite close in energy to the most stable. Do these thermodynamically plausible structures represent practically important polymorphs?

Experimental searches for solid forms can produce a large number of crystal structures containing a given molecule, particularly if the molecule is used in co-crystal screening. Whilst there is a strong relationship between the experimental structural landscape of a molecule (Price, 2009), there are usually more thermodynamically feasible crystal struc-



**Figure 2**

Molecular diagrams of some of the compounds whose crystal energy landscapes are discussed in the text. Molecules IV, VI and XX retain the numbering used in the Cambridge Crystallographic Data Centre blind test papers, and disrupt the sequence.

tures computed than known polymorphs. This article outlines some of the reasons why, mainly based on the output of the project 'Control and Prediction of the Organic Solid State' in which experimental studies inspired, developed and validated the calculation and interpretation of crystal energy landscapes. The range of molecules for which CSP studies have proved useful is shown in the key to the compounds discussed (Fig. 2). This article seeks to illustrate the issues that need to be considered before we could hope to predict the range of solid forms of a given molecule, rather than reviews the use of CSP as a complement to experimental solid-form screening. That would be a huge and quickly outdated endeavour as, since it was last attempted (Beyer *et al.*, 2001), the rapid development in the sophistication of both experimental screening and CSP methods has shown the problems in defining a standard search either *in silico* (CSP) or *in silico* (and other containers used in solid-form screening). Hence, this article attempts to answer the title question, or more precisely: Why do crystal structure prediction methods usually generate more thermodynamically feasible structures than known polymorphs?

### 1.1. Because of approximations in the calculations, particularly the neglect of thermal effects

The energy differences between polymorphs are of the order of  $1 \text{ kJ mol}^{-1}$ , which is a real challenge to the accuracy of current computational methods, even expecting considerable cancellation of the errors in the necessary approximations. Hence, success in terms of usually getting the known crystal structures as the most stable, or sufficiently close (within the first three is the arbitrary criterion used for blind tests of CSP) has been shown to require very computationally expensive methods. These are all based on electronic structure calculations on either the crystal structures (Chan *et al.*, 2011a; Neumann & Perrin, 2005) or the molecule in each plausible conformation (Day, 2011; Price *et al.*, 2010). No method in current use was able to reliably calculate the relative lattice energies for all the molecules in the last CSP blind test (Bardwell *et al.*, 2011). This is because the different types of intermolecular and intramolecular forces that need to be balanced in calculating the lattice energy depend considerably on the type of molecule. We can estimate whether the theoretical underpinnings of a given type of crystal energy evaluation is suitable for a given molecule, in terms of its size, functional groups, likely conformational flexibility and intermolecular interactions, so that a CSP study is worthwhile. [A CSP study is a waste of time if the computational method used in the final optimization method does not reproduce the known crystal structures of the molecule (or related molecules) with reasonable energies.] However, it is the specific molecule itself that determines whether it has one good way of packing with itself, or a range of almost equally good compromises, and hence the confidence with which the most stable structure can be predicted. For example, when a simple CSP method was used to calculate the crystal energy landscapes of the isomers of dichloronitrobenzene (Barnett, Johnson *et al.*, 2008), the 2,3-isomer was clearly found (by

$1.75 \text{ kJ mol}^{-1}$  relative to its next best structure) as the structure with the lowest lattice energy. However, the success of this method in finding the observed structures of other dichloronitrobenzenes at their global minimum decreased as it became less stable and also less distinct in energy from other packings of the molecule. Hence, if, as in the cases discussed below, the known structures are amongst the most stable on the crystal energy landscape, and the relative energies of the unobserved types of structures are not too sensitive to the approximations in the calculations, the question as to why the alternative structures are not observed polymorphs remains.

An obvious limitation of the lattice-energy based CSP methods, emphasized by studies of atomic, ionic or molecular crystals of small polyatomics, is that CSP calculations usually are based on static lattice energies (*i.e.* the 0 K energies, neglecting zero-point motions), rather than the free energies. Molecular dynamics (MD) studies to include molecular motions show that this can significantly reduce the number of low-energy structures. A metadynamics study of benzene (Raiteri *et al.*, 2005) showed that the vast number of lattice energy minima reduced to approximately the number of known polymorphs. However, benzene does undergo facile solid-state transformations. In contrast, a study of 5-fluorouracil showed that 75% of the 60 most stable distinct lattice-energy minima were also free-energy minima at 310 K (Karamertzanis, Raiteri *et al.*, 2008). This reluctance to transform is consistent with the two polymorphs not showing any transformation below the melting point (Hulme *et al.*, 2005), as is often observed when a pair of polymorphs have different hydrogen-bonding motifs. Hence, the number of structures which are energetically competitive at 0 K that are artefacts of the neglect of the molecular motions within the crystal structure is very dependent on the ease of solid-state transformations between the (hypothetical) crystal structures.

Many organic solid-state polymorphic phase transformations are difficult, as once the molecule is close packed within a crystal, there is a significant barrier to rearrangement. It has been argued that all organic molecular transitions are first order (Mnyukh, 2001), with the phase change requiring a nucleation and growth mechanism, rather than a mechanism that goes from single-crystal to single-crystal, maintaining translational symmetry throughout. Certainly, even the solid-state phase transformation of tetrachlorobenzene, which occurs reversibly with no apparent loss in sample quality with very subtle structural changes except in terms of crystallographic symmetry, is first order with a sample-dependent transformation temperature and hysteresis (Barnett *et al.*, 2006). The type of polymorphism that can be readily observed by crystal to crystal transformation on changing the temperature can come from the high temperature, higher symmetry phase being a dynamic average over two lower symmetry lattice-energy minima. Even zero-point motion can average lattice-energy minima, giving rise to a higher symmetry structure that is strictly a transition state in the static lattice-energy potential surface. For example, the powder X-ray diffraction pattern for racemic naproxen (III) was a good match for the lowest energy structure found on the

$Z' = 1$  crystal energy landscape (Braun, Ardid-Candel *et al.*, 2011), but this was a transition state structure. Lowering the symmetry gave a  $Z' = 2$  structure as a true lattice-energy minimum, but the  $1 \text{ kJ mol}^{-1}$  energy lowering only corresponded to a minor shifting of layers and minor changes in conformation, and so this structure was an equally good match for the powder X-ray diffraction pattern. Solid-state NMR was used to confirm that structure was indeed *Pbca*  $Z' = 1$ , confirming the estimate that even the zero-point motions would average over the  $Z' = 2$  *Pca* $_2$  lattice-energy minima.

Hence, we often have far more lattice-energy minima than free-energy minima. Short molecular dynamics simulations can reduce the number of minima, and equilibrated simulations can increasingly be applied to study phase transitions, where the phase transitions can be observed. However, CSP is most valuable in predicting polymorphs which are not readily observed. In these cases, the dynamical simulation of the crystal at a specific temperature is not going to reveal anything – the molecular packing was determined during crystallization, and expensive MD calculations probably will only show the thermal expansion produced by the anharmonicity in the lattice vibrations. Changes in which is the most stable polymorph with pressure, can, in principle, be readily predicted by CSP by recalculating the relative energies as a function of pressure (Kendrick *et al.*, 2013).

## 1.2. Because of the molecular rearrangement during nucleation and growth

Unless solid-state transformations are facile, the crystal packing becomes fixed at the temperature and pressure at which it is formed. It is the reorientation and conformational changes of the molecules as they associate during the time allowed for crystallization from the melt or solution that determines whether a molecule crystallizes in the stable or metastable form. If the crystal that is formed the fastest is always the most stable, then it would normally not be possible to observe polymorphism (Desiraju, 2007). However, when the kinetics of crystallization favour a metastable structure, then polymorphs will be observed. Discussions of Ostwald's Law of Stages (Threlfall, 2003), disappearing polymorphs (Dunitz & Bernstein, 1995) and polymorph selectivity (Blagden & Davey, 2003) all show the complexity of the interplay between the kinetics of nucleation (Davey *et al.*, 2013) and growth (Yu, 2007). Although great progress is being made in the computational modelling of nucleation (Anwar & Zahn, 2011) and the relative growth rates of the crystal faces (*i.e.* crystal morphology; Rohl, 2003; Schmidt & Ulrich, 2012) as a complement to experimental investigations, we still know very little about the competition between the kinetics of alternative crystallization routes that lead to different polymorphs. Concomitant crystallization (Bernstein *et al.*, 1999), where different polymorphs crystallize in the same experiment, demonstrates that this competition can be very finely balanced. The search for conditions for producing phase pure samples of different polymorphs is empirical and can be frustrating; attempts to find conditions to obtain form III of

olanzapine (V) separately from form II have not yet been successful, consistent with the likely relationship between the structures deduced from the crystal energy landscape (Bhardwaj *et al.*, 2013).

The first nucleation of a new polymorph is a key event. Difficulty in nucleating accounts for the late appearance of new polymorphs such as form II ritonavir. Nucleation is generally a sufficiently rare event that it is unlikely to be seen in brute-force molecular dynamics simulations (Anwar & Zahn, 2011): most simulation methods for studying nucleation require the definition of a path or end structure. This is problematic when experiment or CSP suggests there are many thermodynamically competitive crystal structures as distinct end points. A more practical approach to nucleation competition may be molecular dynamics simulations on nanocrystallite clusters derived from the CSP competitive structures, as applied to tetrolic acid in the two solvents that lead to different polymorphs (Habgood, 2012). The change in relative stability of polymorphs with size of nucleus has been demonstrated for L-glutamic acid and D-mannitol, where small cluster nanocrystallites of the bulk metastable phase are more stable than clusters derived from the stable form (Hammond *et al.*, 2012).

Whilst our understanding of crystallization kinetics is so limited, we can only qualitatively compare the structures on the crystal energy landscape to assess whether they are so similar that they are likely to correspond to the same structure somewhere along the crystallization route. Are the structures sufficiently different for it to be plausible that different conditions would lead to their crystallization as distinct polymorphs? We are still at the stage of learning, by comparison of the crystal energy landscape with experimental structures of polymorphs, what differences in the computed structures correlate with observable polymorphs. The development of diverse methods of comparing dozens of crystal structures is essential to this task. The analysis tools in *Mercury* (Macrae *et al.*, 2008), such as calculating the optimum overlay of the coordination cluster of the molecules between two crystals, or the complementary analysis in terms of supramolecular constructs by *XPac* (Gelbrich & Hursthouse, 2005; Gelbrich *et al.*, 2012), prove very useful in determining packing features, such as dimers, stacks or layers, that are in common or differ between the low-energy structures. Most recent applications of crystal energy landscapes to rationalize the results of solid-form screening on pharmaceutical-sized molecules rely extensively on a combination of *XPac* and hydrogen-bonding analyses (Bhardwaj *et al.*, 2013; Braun, Ardid-Candel *et al.*, 2011; Ismail *et al.*, 2013; Kendrick *et al.*, 2013; Montis *et al.*, 2012).

Qualitative estimates of whether two structures are so closely related that only the more stable would crystallize, or sufficiently different that there would be a barrier to inter-conversion during nucleation, can be based on dominant strong interactions or synthons (Desiraju, 2007). For example, hydrogen bonds are generally expected to be sufficiently strong that structures with different hydrogen bonding can be expected to be observed as different polymorphs. This

assumption was behind the 2001 blind test discussion (Motherwell *et al.*, 2002), where two target molecules were generally agreed to have structures with alternative hydrogen bonding as competitive in energy to the known structures, and so polymorphs of the target structures might be expected. A new polymorph of 6-amino-2-phenylsulfonylimino-1,2-dihydropyridine (VI) was later found, after 80 cell measurements on single crystals, with the anticipated alternative hydrogen-bonding synthon (Jetty *et al.*, 2003). A third polymorph was found later (Roy & Matzger, 2009) and these three structures proved to be the most stable in a more recent CSP study (Chan *et al.*, 2011*b*). In contrast, it was only after 3-azabicyclo[3.3.1]nonane-2,4-dione (IV) had been resynthesized and thermal analysis found a plastic phase that the strength of the imide hydrogen bonding was considered (Hulme *et al.*, 2007). The spherical shape of the molecule and the weakness of the electrostatic potential around the hydrogen-bonding proton meant that doubly hydrogen-bonded dimers could easily rearrange to give a hydrogen-bonded catemer. The automated solvent crystallization screen did find a novel polymorph, and two solvates, but all were based on catemeric hydrogen bonding. The energy difference between the ordered crystalline polymorphs and the most stable, dimer-based unobserved crystal structure was well within the error of the methods used (0.6–2.1 kJ mol<sup>-1</sup>), so the target structure in the blind test was the thermodynamically stable form. Differences in hydrogen bonding between the low-energy structures are easily spotted by software that uses distance criteria, aided by hydrogen-bond graph sets (Bernstein *et al.*, 1995), but this is not a reliable guide as to whether the hydrogen bonds produce a sufficient barrier to transformation to produce polymorphs.

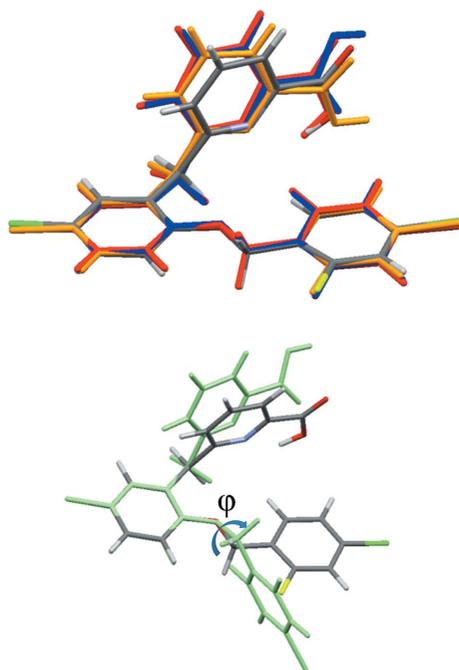
The ability to trap molecules in metastable polymorphs because they cannot rearrange readily can be associated with larger functional groups and conformational flexibility. For example, to discuss whether the fenamate group could be considered a polymorphophore (López-Mejías *et al.*, 2009) we analysed the  $Z' = 1$  crystal energy landscapes of monomorphous fenamic acid (VIIa) and pentamorphous tolfenamic acid (VIIb) (Uzoh *et al.*, 2012). All low-energy crystal structures contained carboxylic acid dimers with the two phenyl rings extending like paddle wheels. In the isolated dimer the phenyl rings can adopt a wide range of angles, limited only by steric clashes between specific non-hydrogenic substituents. However, once the phenyl rings from different dimers interdigitate in forming the nucleus or crystal, the conformational flexibility is drastically reduced. This general type of argument helps explain why most fenamates are polymorphic, with nonamorphous flufenamic acid (VIIc) (López-Mejías *et al.*, 2012) rivalling ROY (I) in the number of polymorphs that have been structurally characterized. The crystal energy landscape of tolfenamic acid [like that of ROY (I); Vasileiadis *et al.*, 2012] shows structures that are competitive in energy and so may yet prove to correspond to further polymorphs. In stark contrast, the  $Z' = 1$  lattice energy landscape of fenamic acid (VIIa) had its most stable structure equi-energetic with the only observed structure (which is  $Z' = 2$  and so could not have been found in the search). Both structures are closely related to the next

most stable computed structure, 2 kJ mol<sup>-1</sup> higher in energy, which can be stabilized either by changing the torsion angles of half the molecules to give the observed  $Z' = 2$  structure, or by sliding the sheets to obtain the CSP-generated most stable structure. Hence, it seems unlikely that fenamic acid will have long-lived polymorphs, as the computationally generated energetically competitive structure is too similar to the observed form. A bromo derivative of ROY appears to be monomorphous, although other derivatives are polymorphic (Lutker *et al.*, 2008). Hence, the common feature of the fenamates (VII) and ROY (I) polymorphophores, of having flexibly linked aromatic groups that can adopt a wide range of conformations only until the molecules have associated, does generally promote polymorphism. However, the crystal energy landscape is needed to show whether or not the specific substituents allow the range of molecular shapes to adopt thermodynamically competitive structures and hence be possible polymorphs.

Conformational flexibility clearly plays a significant role in crystallization. The molecular conformations observed in crystal structures are generally low in energy but how close this is to a local or global minimum energy conformation for the isolated molecule depends on the nature of the molecule. Small molecules such as aspirin (VIIIa) (Ouvrard & Price, 2004) usually crystallize in a conformation that is sufficiently close to a local minimum that there is only a minor difference in the van der Waals surface. For molecules with two or more rigid functional groups linked by flexible torsion angles, like the fenamates (VII), the isolated molecule conformational energy wells can be sufficiently shallow that the conformations in crystal structures can be close in energy to a gas-phase minimum, but the difference in the flexible torsion angles can give rise to a different overall molecular shape (Uzoh *et al.*, 2012). The need to produce a densely packed crystal structure competes with conformational preference, for example, succinic acid adopts a planar conformation in the overwhelming majority of its 56 predominantly cocrystal structures, whereas the most stable conformation is non-planar. The non-planar isolated molecule structure cannot pack to produce any crystal structures that are competitive in energy with the known polymorphs (Issa *et al.*, 2012), and whatever the conformation of succinic acid in solution it can easily become planar during crystallization. However, generally crystal structure conformations are good guides to conformational preferences in other phases (Cruz-Cabeza *et al.*, 2012). Indeed, for many small molecules the conformational flexibility is such that searching with the rigid isolated molecule conformation(s) and then adjusting the flexible torsion angles within the crystal packing (Kazantsev, Karamertzanis, Adjiman & Pantelides, 2011), to gain improvements in close packing or hydrogen bonding (*e.g.* for OH and NH<sub>2</sub> groups), is a very effective strategy. However, when there are a wide range of conformations corresponding to very different shapes of the molecule which are low in energy, a different search strategy is needed. Such strategies have been successful in blind tests of CSP, for example, benzyl-(4-(4-methyl-5-(*p*-tolylsulfonyl)-1,3-thiazol-2-yl)phenyl)carbamate (XX) (Bard-

well *et al.*, 2011; Kazantsev, Karamertzanis, Adjiman, Pantelides, Price *et al.*, 2011) or the contrasting case of a very flat two-dimensional conformational profile for 1-benzyl-1*H*-tetrazole (X) (Spencer *et al.*, 2012). It is challenging to evaluate a reliable conformational energy surface for larger molecules, even in isolation, as the intramolecular dispersion plays an important role. [For example, routine SCF calculations give a broad minimum in the conformational profile of tolfenamic acid (VIIb), in the region which corresponds to a local maximum ( $\sim 5 \text{ kJ mol}^{-1}$ ) for a range of higher quality *ab initio* methods and a low population of experimental structures (Uzoh *et al.*, 2012)]. A particular challenge to evaluating relative lattice energies for conformational polymorphs is when an intramolecular hydrogen bond in one polymorph (and the gas phase) changes to being intermolecular in another (Karamertzanis, Day *et al.*, 2008). The problems in modelling dispersion and polarization and all other contributions to equal accuracy makes modelling even the lattice energy accurately very demanding of current computational chemistry methods (Kendrick *et al.*, 2011; Price, 2008a).

Once the crystal energy landscape has been calculated, a comparison of the conformations within the crystal structures will reveal the range of conformations that can pack densely with favourable intermolecular interactions. For example, for olanzapine (V) there are two low-energy conformational wells, but one conformation only produces one low-energy



**Figure 3**

The conformational issues that can arise in interpreting a crystal energy landscape of a flexible pharmaceutical GSK269984B (XI): (a) overlay of the *ab initio* optimized isolated molecule conformation (red) with the conformations in all the crystal structures obtained in an extensive experimental screen guided by the crystal energy landscape (Ismail *et al.*, 2013); the only single component structure (atomic colours), the *N*-methylpyridone (blue) and a 1:2 dimethylsulphoxide (orange) solvate; (b) overlay of the observed conformation (atomic colours) and that in the next most stable calculated structure (light green) on Fig. 1(b).  $\varphi$  is the angle used to classify the structures.

structure which has such a low packing coefficient (Gavezzotti, 1983) that it is unlikely to be an observed polymorph (Bhardwaj *et al.*, 2013). However, in other cases, such as GSK269984B (XI), the structures that are competitive in energy (Fig. 1b) have such different gross conformations from that observed in the three structurally characterized forms (Fig. 3), that if formed, they would be long-lived polymorphs (Ismail *et al.*, 2013). Hence, the question arises as to how much conformational rearrangement can occur during nucleation.

The degree of rearrangement possible during nucleation and growth will be very dependent on the molecule and has to be contrasted with the timescales of different crystallization methods. Simple molecules which are liquids or gases at room temperature may have plenty of time to rearrange to the most stable form as they are cooled to low temperature, although it is notable that a different polymorph is often produced by crystallization under high pressure (Ridout & Probert, 2013). The rarity of polymorphs in a screen of dichloronitrobenzenes could be attributed to most crystallizations being performed just below the melting points of these weakly interacting molecules (Barnett, Johnson *et al.*, 2008) implying considerable scope for the rearrangement of any metastable polymorphs. At the other extreme, it may be that the effective building block for crystallization is determined at the earliest stages of crystallization. There are established correlations between the prenucleation clusters observed in different solvents, either experimentally or in simulations (Davey *et al.*, 2013), and the polymorph formed. The formation of preferred pre-nucleation aggregates or conformations could be even more pronounced for larger, flexible molecules. Olanzapine (V) forms a centrosymmetric dimer in all 59 solid crystalline forms (three polymorphs, seven groups of isostructural solvates; Bhardwaj *et al.*, 2013), but the crystal energy landscape includes other structures which do not contain the dimer but have molecular sheets in common with the known forms. Since the dimer interaction dominates the lattice energy of the known polymorphs it is clearly tempting to assume that the dimer forms so early in the molecular association, prior to nucleation and growth, possibly dominating the solution structure, that it is not possible to nucleate structures without this motif. We are still very much at the early stages of being able to detect and understand how kinetic factors determine which thermodynamically feasible crystal structures could be kinetically trapped as metastable polymorphs.

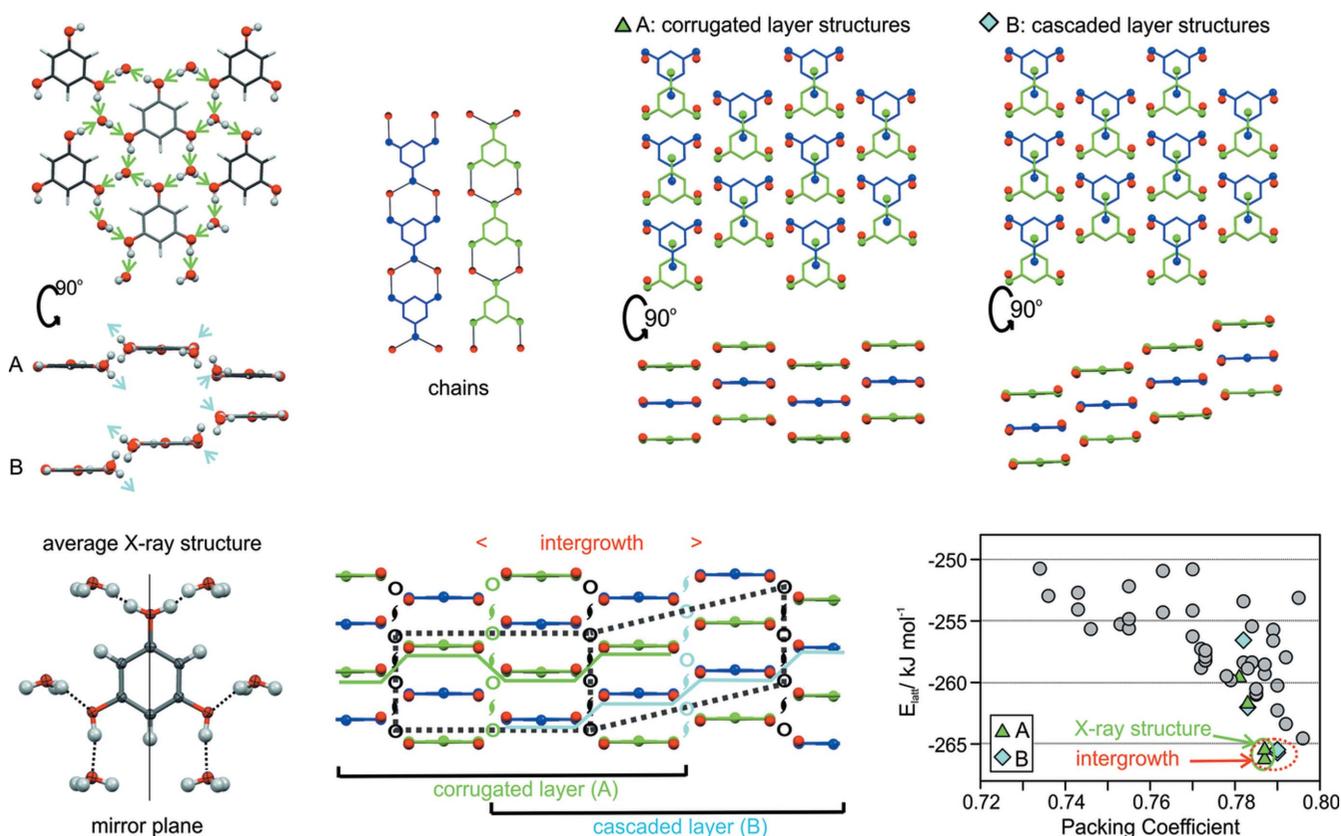
### 1.3. Because the structures can be present in disordered or solvated crystals and other phases

Cyclopentane adopts the lowest lattice-energy structure below 123 K (at low pressure), but then has an intermediate dynamically disordered phase, before transformation above 137 K to the high-temperature cubic phase. This behaviour can be rationalized as the increasing dynamical motion with temperature sampling an increasing number of the multitude of very closely spaced lattice-energy minima (Torrise *et al.*, 2008). Thus, one way of interpreting plastic or dynamically disordered phases is that the molecules are moving between

the large number of almost equi-energetic lattice-energy structures: *i.e.* the low-energy barriers between the different structures produce dynamic disorder giving rise to high-symmetry structures. Thus, distinct lattice-energy minima separated by low-energy barriers can correspond to the same experimental structure, *i.e.* many of the low-energy structures are ‘observed’ within the dynamically disordered high-temperature phases. The distinction between dynamic and static disorder is complex, depending on the temperature and the barriers for correlated motions of the molecules and in practise will depend on the size and defects of individual crystals. Indeed, for all molecules, configurational and thermal entropy will lead to disorder at some temperature before or at the melting temperature. Hence many polymorphs are observed by thermal microscopy (Kuhnert-Brandstatter, 1982), although isomorphous seeding, sublimation and quench cooling can be involved.

Static disorder is likely when the crystal energy landscape contains two or more structures that are so closely related both in structure and energy that it is hard to imagine molecules being able to assemble into perfect ordered crystals of either structure (van Eijck, 2002). This static disorder has been demonstrated in the case of eniluracil (XII) where there are

many low-energy structures (Copley *et al.*, 2008) which are effectively the same if you do not distinguish between the C4=O and C6–H groups. These two groups are not involved in forming the hydrogen-bonded ribbons and barely affect the interdigitation of the ethynyl groups. These crystal structures are sufficiently close in energy that configurational entropy stabilizes a disordered structure (Habgood, Grau-Crespo & Price, 2011). Careful analysis of the diffraction from four different single crystals, grown from two sets of similar conditions, confirmed that the structures were indeed disordered, with the site occupancy factors of the major component varying between 0.70 and 0.84 for the different crystals (Copley *et al.*, 2008). In this case, different individual structures had been proposed by structure solution from powder X-ray diffraction data, and even single-crystal diffraction data could reasonably have been interpreted as polymorphism. It seems more reasonable, given the crystal energy landscape, to view eniluracil crystallization as being intrinsically disordered, with the degree of disorder being dependent on crystallization conditions. This disorder is seen in isostructural 5-chlorouracil and 5-bromouracil, and indeed this symmetric ribbon motif is commonly found in the crystal energy landscape of uracils with a small functional group at the 5 position. However, the



**Figure 4**

Analysis of the structures on the crystal energy landscape of phloroglucinol dihydrate (XIV) (Braun *et al.*, 2012). Group A (green triangle) structures match the proton disordered experimental structure, whereas group B (blue diamond) structures differ in the layer structure and so provide a model for the intergrowths that result in morphological features and diffuse scattering. Top, left to right: the hydrogen-bonding directionality which produces the proton disorder in the average X-ray structure below, protons are otherwise omitted for clarity: the two hydrogen-bonded chains of phloroglucinol and water (red) molecules, whose contrasting layer structures support the intergrowth structure below. The energy differences between these structures support this model for the disorder.

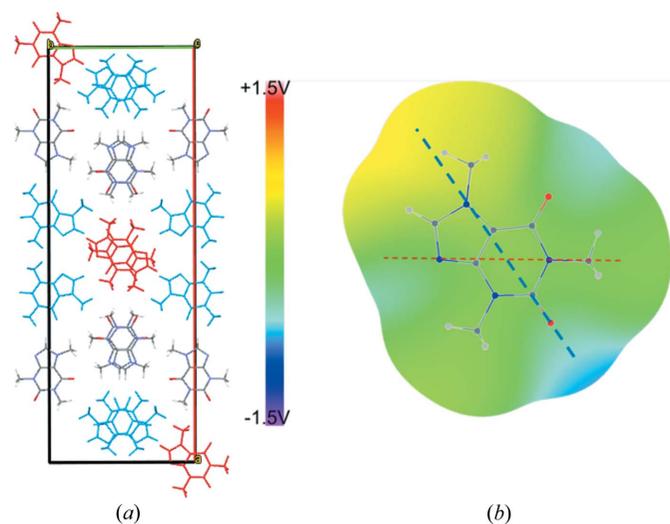
nature of the substituent determines the differences in energy between the various packings of the symmetric ribbon and the many other hydrogen-bonding motifs which can be generated and observed for 5-uracils (Barnett, Hulme *et al.*, 2008).

An important, but arguably small, difference in unique structures on the crystal energy landscape is when they differ only in some proton positions, and yet have very similar energies. Structures which differ only in the directionality of the hydrogen bonds, *e.g.* O—H···O *versus* O···H—O, have frequently been found in the crystal energy landscapes of hydroxybenzoic acids, such as gallic acid (XIII), and their hydrates (Braun *et al.*, 2013). Proton positions are not always readily determined by crystallographic methods, and hence some methods of structural comparison would not consider such structures as different, and cases of proton disorder are likely to be under-reported. Judging from the computed crystal energy landscape, proton disorder seems probable for the metastable monohydrate of 2,4-dihydroxybenzoic acid (Braun, Karamertzanis & Price, 2011). The 50:50 disorder in one hydroxyl and water protons in phloroglucinol dihydrate (XIV) suggested by the crystal energy landscape (Braun *et al.*, 2012) is implicit in the crystallographic symmetry. However, the observed variation in morphology and diffuse scattering of phloroglucinol dihydrate appears to arise from the disorder introduced by the proportion of packing faults represented by the close relationship between corrugated and cascaded layer structures on the crystal energy landscape (Fig. 4).

There are many other cases where the crystal energy landscape shows structures that have the same sheet stacked differently, which are close in energy. This phenomenon was observed on the crystal energy landscape of aspirin (VIIIa), which at the time was believed to be monomorphic, and the high susceptibility to shear was used to rationalize why the structure with the alternative stacking of layers was not observed (Ouvrard & Price, 2004). This structure was later found, although metastable (Vishweshwar *et al.*, 2005), and single crystals of aspirin were found to contain domains of form II and form I as intergrown structures (Bond *et al.*, 2007). Crystallization of aspirin in the presence of aspirin anhydride (VIIIb) has since been found to produce single crystals of form II which are very long-lived (Bond *et al.*, 2011). Hence, when the crystal energy landscape has nearly equi-energetic structures containing the same layer yet stacked in different ways, whether crystallization will lead to only one structure probably with some stacking faults corresponding to the alternative stacking, polytypism, polymorphic domains or polymorphs can be very sensitive to both crystallization conditions and the definition of the distinctions in this continuum of structures. Quantification of stacking disorder through lattice-energy calculations has recently been used to explain the stacking disorder in tris(bicyclo[2.1.1]hexeno)benzene (Schmidt & Glinemann, 2012). This type of calculation for evaluating the configurational entropy from the lattice energies of an ensemble of ordered supercells, generated by a site-occupancy disorder model (Habgood, Graucrespo & Price, 2011), has been applied to eniluracil (XII), the formation of a solid solution of *p*-dichlorobenzene and *p*-

dibromobenzene and the low-temperature structure of caffeine (XVa) (Habgood, 2011). Evaluation of configurational entropy requires the definition of likely disorder components, which can be derived from the crystal energy landscape (Habgood, 2011). Disorder components will correspond to near-symmetry in the intermolecular interactions (*e.g.* there is almost a mirror plane in the electrostatic potential on the van der Waals surface of the molecule but not in the molecular diagram) and results in there being almost equi-energetic low-energy structures on the crystal energy landscape, related by the disorder model (Figs. 5 and 1c). The closely spaced lattice-energy minima can be considered as observed structures since they are members of the symmetry-adapted ensemble of structures that contribute to the entropic stabilization of the observed disordered structure.

Even without quantification of configurational entropy, the relationships between closely spaced structures on the crystal energy landscape quickly give a rationalization of complex crystallization behaviour, and can help characterize structures. For example, a  $Z' = 1$  CSP study of chlorothalonil (XVI) (Tremayne *et al.*, 2004) produced two layer structures that correlated with the high-symmetry disordered layer structure of form 2 and two herringbone structures that had very similar powder patterns, and eventually proved to be components of the  $Z' = 3$  single-crystal structure of form 3. All four 'observed' structures were within  $1.25 \text{ kJ mol}^{-1}$  of the global minimum structure, which corresponded to the stable form 1. A more



**Figure 5**  
Disorder in caffeine (XVa): (a) An ordered representation of the structure of form II caffeine (Enright *et al.*, 2007) with the ordered molecule in atomic colouring, and the two independent disordered components in red and blue. The stacking was used to classify the structures on the crystal energy landscape in Fig. 1(c) (*i.e.* stack 1 atomic, stack 2 red, stack 3 blue). (b) The electrostatic potential on the van der Waals surface plus  $1.2 \text{ \AA}$  for caffeine;  $+1.5 \text{ V}$  corresponds to an interaction energy of  $+1.4 \text{ kJ mol}^{-1}$  with a positive point charge of  $0.01 e$ . The experimental structure has disorder components corresponding to rotation by  $180^\circ$  about the two marked axes (Habgood, 2011), with the heavier dashed line corresponding to the blue disorder component which has greater approximate symmetry in the steric and electrostatic interactions of the molecule than the red component which corresponds to a crystallographic axis.

extensive search, including higher  $Z'$  structures or more space groups, would have increased the number of related structures well within the energy range of plausible polymorphism, and similarly increasing the experimental search would be expected to find more structural variations.

The formation of solvates as disordered intermediates or final products of crystallization also needs considering. The crystal energy landscape of a DMSO solvate of carbamazepine (XVIIa) (Cruz-Cabeza *et al.*, 2011) suggested the solvent disorder seen in the low-temperature crystal structure, and in the dynamically disordered high-temperature phase. Solvent can be undetected due to its mobility within the crystal structure, which, since the solvent stabilizes the structure, may mislead the interpretation of the relative stability of observed polymorphs. Investigations prompted by the relatively high lattice energy of carbamazepine form II (Cruz-Cabeza *et al.*, 2007a) showed that solvents move quite freely in the channels. Only recently have water wires in these hydrophobic nanopores of carbamazepine form II been observed by single-crystal X-ray diffraction (Prohens *et al.*, 2013). On the other hand, no experimental evidence could be found for water in the channels of a diastereomeric salt (*R*)-1-phenylethylammonium-(*S*)-2-phenylbutyrate (XVIII) to account for why the observed structure was quite high in energy on the crystal energy landscape (Antoniadis *et al.*, 2010). Frameworks, layers or other bimolecular packing motifs seen in isostructural and disordered solvates are often on the crystal energy landscape of molecules which are much larger than solvent molecules, such as olanzapine (V) (Bhardwaj *et al.*, 2013) and (hydro)chlorothiazide (XIX) (Johnston *et al.*, 2007, 2011). Indeed, the framework structures of inclusion compounds which can contain a variety of guests have been found as low density, higher-energy structures on the crystal energy landscape of the framework molecules (Cruz-Cabeza *et al.*, 2009). CSP has also been able to predict the assembly of porous organic molecular crystals (XXVI) (Jones *et al.*, 2011), a case where the exact nature as well as disorder in other components present is irrelevant.

An extreme example of disorder is gel formation, where small molecules are dispersed within the solid phase. CSP studies on two chemically similar functionalized dipeptides (XXI) showed that the one which forms a *meta*-stable hydrogel prefers to pack as tightly coiled molecular columns, whereas this mode of packing is not found for the other dipeptide that forms a crystalline solid (Adams *et al.*, 2010). Low-energy CSP structures for melamine uric acid dihydrate gave sufficiently close matches to the PXRD pattern of the dried xerogel to give insight into the gel structure formed by sonication of an aqueous solution of melamine and uric acid (Anderson *et al.*, 2008).

Another extreme form of disorder is the amorphous solid state. Analysis of the crystal energy landscape of salsalate (XXII) reveals a pair of hydrogen-bond motifs that appear to be severely detrimental to the molecule's ability to pack efficiently and stably, indicating an explanation for the formation of a stable amorphous phase (Habgood *et al.*, 2013). Concurrent experimental work suggests that impurities from thermal

degradation may also play a role in frustrating crystallization. Conformational flexibility and size may help, but are not necessary for the formation of an amorphous state. The polymorphous nature of carbonic acid (XXIII) also appears to be rationalized by two different conformations of carbonic acid having a multitude of crystal packings of similar energies (Winkel *et al.*, 2007).

Thus, analysis of the structures on the crystal energy landscape often suggests that there are problems in forming a close-packed ordered solid. If these result in difficulties in growing a crystal suitable for single-crystal diffraction, then it can be difficult to determine the existence and nature of the disorder, and the CSP generated structures can support experimental progress in determining structures of disordered through to amorphous solids. It appears that some of packing relationships in the low-energy structures are observed, but only at the molecular level as disorder within the solid forms. For larger molecules, some of the calculated structures for a given molecule are closely related to the molecular packing observed in solvated structures, where the solvent acts as a space-filler and may be disordered. Indeed, desolvating solvates can be an important route to new polymorphs (Braun *et al.*, 2013; Ismail *et al.*, 2013), emphasizing that the point at which the solvent is expelled during crystallization can vary. Thus, some metastable structures on the single component crystal energy landscape reflect the packing in loosely solvated solid forms. Conversely, the interplay between thermodynamics and kinetics in different multi-component crystallization routes can lead to even the most stable structure on the crystal energy landscape not being observed.

#### 1.4. Because the structures are metastable relative to other molecular compositions

CSP methods will always generate crystal structures with the molecules in the stoichiometry specified: searches with more than one molecule in the asymmetric unit cell are considerably more demanding because of the need to consider the different relative positions of the molecules within the asymmetric unit. The individual search does not tell you whether even the most stable structure generated is going to be more stable than the components or another stoichiometry and hence could be observed. Analysis of the structures generated can be highly suggestive. For example, the lowest energy structures generated for a hypothetical 1,4-dicyanobenzene:succinic acid cocrystal have layers of succinic acid molecules alternating with layers of 1,4-dicyanobenzene molecules, with the layers closely resembling the structures in succinic acid and 1,4-dicyanobenzene polymorphs. Given that this is the most favourable type of structure for the cocrystal, it is not surprising that no cocrystal was found in an extensive experimental search (Issa *et al.*, 2012). Similarly CSP searches for urea:acetic acid (1/1) produced structures with unsatisfied hydrogen-bonding donors; urea:acetic acid (1/3) had an acid molecule which was only filling space and not hydrogen bonded to urea, whereas the urea:acetic acid (1/2) structures had ribbons which satisfied all the hydrogen bonding (Cruz-

Cabeza *et al.*, 2008). The second most stable structure for urea:acetic acid (1/2) corresponded to the observed crystal, although this was only stable in the mother liquor. Comparing the relative lattice energies of urea and three acetic acid molecules in various single and multicomponent structures did give the 1:2 structure as thermodynamically preferred, but the energy difference was small (Cruz-Cabeza *et al.*, 2008). There is considerable variability in the energy differences between different stoichiometries of solvates, and this approach was able to determine the structure of a novel acetic acid solvate of theobromine produced by grinding (Cruz-Cabeza *et al.*, 2010).

The prediction of whether a structure will form, by comparing lattice energies as evaluated using the same computational model for either known or lowest energy structures, is obviously a crude model when it comes to hydrate prediction: the comparison is with the ice polymorphs. Comparison of stoichiometric ordered hydrate lattice energies with those of the anhydrate and ice form XI showed that the stabilization of the hydrate was generally small, and sensitive to small changes in conformation of, for example, the OH and NH<sub>2</sub> groups of otherwise rigid molecules (Hulme & Price, 2007). Later CSP hydrate studies modelling molecular flexibility, polarization within the crystal, and comparing with the energy range of ice polymorphs, was able to rationalize why 2,5-dihydroxybenzoic acid forms no hydrates, despite being subjected to the same screening that found a stable hemihydrate and metastable monohydrate for 2,4-dihydroxybenzoic acid. Despite structural differences in the hemi-, mono- and dihydrates generated, the dominant reason for the hydrate formation was that 2,4-dihydroxybenzoic acid is less stable and less dense in the anhydrous phase than its isomer (Braun, Karamertzanis & Price, 2011).

Cocrystals might be expected to be more readily predictable, and indeed a CSP search found only two carbamazepine:isonicotinamide cocrystal structures with lattice energies lower or comparable with the lattice energies of the most stable polymorphs of the components, corresponding to the two polymorphs of this cocrystal (Habgood *et al.*, 2010). This study also showed that the lack of stable cocrystals of carbamazepine with the isomer picolinamide arose from the difference in intermolecular hydrogen-bonding capability. More complex differences in cocrystallization behaviour, in terms of structure and stoichiometry, of caffeine with 2-, 3- and 4-hydroxybenzoic acid have also been explained by CSP studies (Habgood & Price, 2010). Thus, CSP can be very effective for generating structures of cocrystals (Karamertzanis *et al.*, 2009), but the energy differences between cocrystals and components are generally of the order of polymorphic energy differences (Issa *et al.*, 2009) and so require very accurate evaluation (Chan *et al.*, 2013). Cocrystal formation may not be in kinetic competition with the formation of the most stable polymorphs of both components, and the experimental thermodynamic advantage of cocrystal formation can be small and a balance of many small energy terms (Oliveira *et al.*, 2011).

Proton migration can also mean that a CSP generated structure does not form. The distinction between a salt and a

cocrystal affects regulatory issues for active pharmaceutical ingredients, producing intense debate (Aitipamula *et al.*, 2012). Even crystallization of simple pyridines with carboxylic acids can result in a salt, cocrystal or disordered solid form as defined by the acidic H-atom position, when the difference in p*K*<sub>a</sub>s is small (Mohamed *et al.*, 2009). The proton position is very important for computer modelling of the solid phases. This is emphasized by comparing the crystal energy landscapes of pyridinium carboxylate salts and the corresponding pyridine carboxylic acid co-crystals (Mohamed *et al.*, 2011): the experimental salt or cocrystal was found low in energy on the lattice energy landscape with the correct proton connectivity, but only the proton disordered system had its structure low in energy on both landscapes. In addition, despite the similarity of the ionized and neutral forms of the carboxylic acid...pyridine heterosynthon, there are hydrogen-bonding motifs that are only favourable for the salt or for the cocrystal. The importance of proton position has recently been emphasized by the discovery (Schmidt *et al.*, 2011) of tautomeric polymorph IV of barbituric acid (XXIV), as the thermodynamically stable phase at ambient conditions, whereas the keto tautomer is the most stable in the gas phase, solution and in polymorphs I–III. Crystal energy landscapes can be calculated for different protonation states to provide insight into the different types of crystal packings available for a given chemical diagram. However, comparing the energies of the different global minima to determine which is thermodynamically favoured requires ionization potentials and tautomeric energy differences to too great an accuracy for current computational methods and the approximations made in creating the thermodynamic cycles. The crystal structures found experimentally depend on how protons, solvent or other components present can rearrange themselves during the crystallization process. The permutations in crystallization conditions can seem infinite, reflecting McCrone's point about the number of polymorphs found being proportional to time and money put into research on that compound (McCrone, 1965). Crystal energy landscapes do provide an upper limit on the range of practically important polymorphs, but there is a need to consider competition with alternative phases and isomers.

### 1.5. Because the right crystallization experiment has not yet been performed

Interest in CSP has been aroused by the late discovery of polymorphs corresponding to calculated low-energy structures, for example, form II aspirin in a failed cocrystallization experiment (Vishweshwar *et al.*, 2005), form II 5-fluorouracil by crystallizing from dry nitromethane (Hulme *et al.*, 2005) or form II maleic acid in an attempt to recrystallize a caffeine–maleic acid cocrystal produced by grinding (Day *et al.*, 2006). The serendipitous finding of new polymorphs of substances that have been heavily studied (which can be calamitous if it leads to disappearing polymorphs) emphasizes the impossibility of experimentally covering all types of crystallization conditions that have led to the discovery of a new polymorph

of any molecule (Llinàs & Goodman, 2008). Hence, the desire to use crystal structures to design crystallization experiments to target structures that are ‘predicted’ once the factors §§1.1–1.4 have been considered.

This has been demonstrated in the case of carbamazepine (XVIIa), a generic anti-epileptic that has been extensively used in polymorphism studies. The first four polymorphs all contain the  $R_2^2(8)$  amide hydrogen-bond dimer, with this type of hydrogen-bonding motif occurring in all the 50 ordered solid forms (Childs *et al.*, 2009). The first CSP study on carbamazepine (Florence, Johnston *et al.*, 2006) was therefore surprising in showing a catemeric structure as slightly more stable than the most stable observed polymorph. Other, more accurate calculations revised this stability order (Cruz-Cabeza *et al.*, 2006; Welch *et al.*, 2008), but the catemeric form remained thermodynamically plausible. Extensive searching for this polymorph using an automated system, using 66 solvents and 5 different crystallization protocols failed to find a new polymorph (Florence, Johnston *et al.*, 2006), although this also did not find form IV which had been found by polymer templating (Grzesiak *et al.*, 2003). Cruz-Cabeza *et al.* (2007b) suggested that there was a kinetic disadvantage for carbamazepine to form strained hydrogen-bonding chains. Crystallizing carbamazepine with the closely related molecule dihydrocarbamazepine (XVIIb) led to a solid solution (Florence, Leech *et al.*, 2006), which was isostructural with the catemeric dihydrocarbamazepine form II, showing carbamazepine could form the required hydrogen bonds. Finally, subliming carbamazepine onto a crystal of dihydrocarbamazepine form II led to the formation of the first crystals of carbamazepine form V, the targeted isostructural form (Arlin *et al.*, 2011). Thus, polymorph screening on closely related molecules yielded a crystal that could template a targeted polymorph.

Less specific properties of the low-energy crystals may be useful in designing crystallization strategies. Examination of the hydrogen bonding can suggest that experiments in specific solvents may be worthwhile (Cross *et al.*, 2003), whereas denser structures could be targeted by crystallization under pressure (Fabbiani & Pulham, 2006; Kendrick *et al.*, 2013). However, the range of polymorphs is clearly limited by the ability to vary the crystallization conditions sufficiently to change the mode of self-assembly.

### 1.6. Because the right crystallization experiment cannot be performed

The physical properties of many molecules can severely limit the range of crystallization experiments that are feasible. For example, our extensive searches for polymorphs within the 5-uracil family were severely limited by thermal instability and limited solubility (Barnett, Hulme *et al.*, 2008). There can be difficulty in crystallizing at all (Baird *et al.*, 2010; Hursthouse *et al.*, 2009); indeed, stabilizing amorphous phases or the design of new ionic liquids (Dean *et al.*, 2009) relies on preventing crystallization. Since the main value of CSP is to assess the risk of ‘unexpected’ crystallization outcomes, the challenge is to be

sure that a calculated low-energy crystal structure could never be found, or at least not within the process conditions for an industrial product.

Let us consider an ‘*ad absurdum*’ example. The crystal energy landscape of progesterone (XXVa), considering all common space groups for an organic molecule, predicts that a racemic crystal is more stable than the two known forms. However, progesterone naturally occurs as a single enantiomer: the predicted racemic structure could only be experimentally realised by crystallizing natural progesterone with its synthetically produced mirror image molecule (Lancaster *et al.*, 2006). In this case, we could be confident that no crystallization of natural progesterone could lead to the racemic crystal structure without the deliberate addition of the synthetic *ent*-progesterone. However, other chiral molecules can racemize, indeed some very efficient chiral separation processes have been devised, based on having a racemization equilibrium going on at the same time (Noorduin *et al.*, 2008).

It is a lot easier to exclude racemization than it is to exclude the possibility of a significant conformational change. This becomes a major issue when considering larger, flexible molecules. For example, in the last blind test (Bardwell *et al.*, 2011) some groups included both *cis* and *trans* amide conformations of the model pharmaceutical (XX) in their searches, whereas others made the reasonable assumption that it was *trans*. The question becomes increasingly difficult when the barriers to gross conformational changes are smaller. For example, the only anhydrate polymorph found in extensive screening of GSK269984B (XI) was confirmed by the crystal energy landscape to be the thermodynamically most stable form (Ismail *et al.*, 2013). However, there were thermodynamically competitive structures which differed markedly in conformation, both in gross shape and in having intermolecular instead of intramolecular hydrogen bonds (Fig. 1b). Further experimentation targeting conformational change did find novel solvates, which had intermolecular hydrogen bonds. However, their structures revealed that only the carboxylic acid proton conformation had changed: the anhydrate and solvates all had essentially the same conformation as the isolated molecule (Fig. 3). Other conformations exist in solution, as expected from the conformational analysis. How could we target the conformational polymorphs, or diagnose that they could not nucleate in a given manufacturing process?

The crystal energy landscape can contain ‘disappearing polymorphs’ (Dunitz & Bernstein, 1995), and understanding problems of reproducing recipes for polymorphs (Threlfall, 2000) may help devise experiments to stabilize or exclude low-energy structures from being realisable polymorphs. For example, attempts to crystallize form II progesterone (XXVa) failed until it was found concomitantly with form I and a pregnenolone (XXVb):progesterone (XXVa) cocrystal. This allowed the crystal structure of form II to be redetermined from a single crystal, but all samples converted to form I (Lancaster *et al.*, 2007). Analysis of a 50-year old form II sample showed that it contained 11 impurities totalling 4.8%, whereas a form I sample from the same period had only 3 totalling 1.9% (Lancaster *et al.*, 2011). Since the impurity

profile of modern sources of progesterone is totally different, it appears that the crystallization conditions that in the past would produce very 'long lived' form II progesterone cannot be reproduced. This is in contrast to other cases where an impurity can be specified and used to produce a given polymorph, for example 1 mol% ethamidosulfathiazole is needed to produce form I sulfathiazole (Blagden *et al.*, 1998). Thus, whether or not certain metastable, low-energy structures are observed as practically important polymorphs can be a matter of whether the right additives or impurities are present.

### 1.7. Because the polymorphs are not detected or structurally characterized by current techniques

Our database of calculated crystal energy landscapes currently covers over 150 molecules plus 40 two-component systems of fixed stoichiometry (salts, co-crystals and hydrates). This sample is not statically meaningful, not only in terms of numbers and bias towards small rigid molecules, but also in the reasons why the systems were chosen for study. A further bias (shared by the CSD) is towards systems where at least one crystal suitable for single-crystal X-ray diffraction has been grown. Huge advances in instrumentation are producing a rapid change in the types of crystals whose structures can be determined (Bond, 2012). The use of crystal energy landscapes to help characterize or add confidence to structures derived from powder X-ray diffraction is becoming well established. This varies from providing the first crystal structures of pigments (Schmidt *et al.*, 2010) to determining the structure of a novel polymorph (Perrin *et al.*, 2009) or the racemic form (Braun, Ardid-Candel *et al.*, 2011) when single crystals can only be grown for other packings, to solving ambiguities in proton positions (Wu *et al.*, 2013). Currently less established uses include solving the crystal structures from high-resolution  $^1\text{H}$  solid-state NMR (Baías *et al.*, 2013; Salager *et al.*, 2010) or from electron diffraction patterns (Eddleston *et al.*, 2013*b*). The application of transmission electron microscopy to find crystallites with distinct morphologies and use CSP to help determine the crystal structure of theophylline (XVc) form VI is particularly noteworthy, as the new polymorph occurred in a mixture with form II at a concentration below the limits of detection of analytical methods routinely used for pharmaceutical characterization (Eddleston *et al.*, 2013*a*). Investigations of the relationship between the (disordered) polymorphic structure and the phonon modes are becoming increasingly complemented by calculations (Li *et al.*, 2010). Thus, current limitations in detection and characterization techniques can be a further reason why we do not find more polymorphs. As we improve the methods and approximations used to calculate spectra from increasingly accurate computed structures, there will be an increase in the number of polymorphs characterized because CSP helped provide confidence in the structure.

## 2. Conclusions – where are we now?

CSP calculations are now at a stage where they can be used to answer the question 'what types of alternative packings are

thermodynamically competitive with the known crystal structures?'. This can valuably limit the range of possible structures. For example, this explains why certain molecules defy Etter's rules and crystallize with unused hydrogen-bond acceptors (Lewis *et al.*, 2005). The comparison of the computer-generated structures with all the experimental crystal structures containing the molecule, and each other, is often the most important and human time-consuming step in CSP studies. As the examples discussed above show, the crystal energy landscape frequently raises the question; are there more polymorphs to be discovered?

It is quite unusual that the crystal energy landscape will show that the molecule has a unique way of packing with itself, and so a CSP study can confidently say that there are no possible polymorphs. Our most confident prediction of monomorphism is for isocaffeine (XVb) (Fig. 1*a*). The contrast with the crystal energy landscape of its isomer caffeine (XVa) (Fig. 1*c*), which has only been observed in static and dynamically disordered phases, underlines the difficulty of even qualitative crystal structure prediction without CSP studies. It is thus rare that CSP-generated crystal energy landscapes do not require structural analysis and human interpretation. As experimental capabilities increase, and as crystal energy landscapes are calculated for increasingly complex systems, the rules for interpreting the landscapes are emerging. The above reasons for why the thermodynamically feasible structures may or may not be potential polymorphs give some qualitative guidance on points to consider. The individual studies illustrate how the molecule-specific considerations complement experimental investigations of its solid form diversity, but often these studies suggest that further research, either on the molecule or in developing experimental or modelling techniques is desirable.

The further work required depends on the reasons why the CSP study was undertaken and is often limited by the resources currently available. It depends critically on the range of structures on the molecule's crystal energy landscape and the barriers to transformation between them at some point during the crystallization process. If the aim is solely to determine the relative stability of the structures (or compete in blind tests of CSP), then the range of structures determines which approximations in calculating the crystal energy landscape need assessing to estimate the uncertainty in the ranking. If there is a large spread in densities, free-energy calculations are likely to rerank structures that are close in lattice energy, whereas if there are significant differences in conformation and types of hydrogen bonding then it is more effective to assess the sensitivity to different approaches of lattice energy calculation and quality of wavefunction. It can be desirable to contrast a range of different electronic structure based methods of calculating lattice energies (Gelbrich *et al.*, 2013; Habgood, Price *et al.*, 2011). If the aim is to interpret the experimental evidence for additional polymorphs or forms of disorder, then comparisons of calculated spectra or diffractograms and the link between the structures and observed transformations may give a very strong indication of the type of structure responsible. If the aim is to find an

alternative solid form with a better compromise of properties for manufacture or formulation, such as solubility, mechanical or morphological properties, then estimates of these properties are required. This may then lead to the desire to target the finding of a specific polymorph, using the structure to suggest possible templates or types of experiment. Alternatively, the aim may be to establish that energetically competitive structures are unlikely to emerge in a specific manufacturing process.

Crystallization of organic molecules is often a complex interplay between thermodynamics and kinetics when the thermodynamic driving force barely differentiates between the structures. Theoretically, we can define a path to the prediction of the most stable phase as a function of temperature and pressure, and whether it is ordered or disordered, but doing this reliably for all but the simplest molecules is a huge challenge to the development of computational chemistry techniques (Gavezzotti, 2002). The reliable prediction of practically useful polymorphs, or the crystallization conditions to reliably and reproducibly crystallize even the most stable phase, requires better fundamental understanding of crystallization processes appropriate to the range of structures on the crystal energy landscape. We need much greater molecular understanding of nucleation (Davey *et al.*, 2013) and of how different crystallization conditions, including solvents, impurities and surfaces (Carter & Ward, 1993), can direct the formation of different solid forms before we can routinely design an experiment to lead to a desired polymorph, or conclude that it could never be found.

Hence, the current and future uses of calculated crystal energy landscapes are as a complement to experimental studies. Some of the structures on the lattice energy landscape are artefacts of the neglect of temperature and other inaccuracies in the modelling. However, even when we can calculate realistic free energy landscapes, the question as to whether the structures are sufficiently different to be observed polymorphs will often remain. As the examples cited above illustrate, we do not yet know enough about the kinetic factors that can lead to the crystallization and longevity of metastable solid phases to answer this question, let alone produce software capable of predicting all polymorphs and the experimental conditions for finding them. However, calculated crystal energy landscapes can be used as part of the interdisciplinary work needed to fully understand and characterize the solid state of organic molecules.

I wish to thank the many people from different disciplines who worked on the CPOSS (Control and Prediction of the Organic Solid State) project <http://www.cposs.org.uk>, and the EPSRC for funding some of them under the Basic Technology program GR/S24114 and EP/F03573X/1. This paper has been shaped by discussions with many people at meetings associated with CPOSS, the Cambridge Crystallographic Data Centre blind tests of CSP and the EPSRC Directed Assembly Network (EP/H035052/1 and EP/K014382/1 Nucleation and Crystallization theme). Drs Louise Price, Doris Braun and

Matthew Habgood are particularly thanked for their help in producing this manuscript.

## References

- Adams, D. J., Morris, K., Chen, L., Serpell, L. C., Bacsá, J. & Day, G. M. (2010). *Soft Matter*, **6**, 4144–4156.
- Aitipamula, S. *et al.* (2012). *Cryst. Growth Des.* **12**, 2147–2152.
- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Anderson, K., Day, G., Paterson, M., Byrne, P., Clarke, N. & Steed, J. (2008). *Angew. Chem. Int. Ed.* **47**, 1058–1062.
- Antoniadis, C. D., D’Oria, E., Karamertzanis, P. G., Tocher, D. A., Florence, A. J., Price, S. L. & Jones, A. G. (2010). *Chirality*, **22**, 447–455.
- Anwar, J. & Zahn, D. (2011). *Angew. Chem. Int. Ed.* **50**, 1996–2013.
- Arlin, J., Price, L. S., Price, S. L. & Florence, A. J. (2011). *Chem. Commun.* **47**, 7074–7076.
- Baias, M., Widdifield, C. M., Dumez, J. N., Thompson, H. P., Cooper, T. G., Salager, E., Bassil, S., Stein, R. S., Lesage, A., Day, G. M. & Emsley, L. (2013). *Phys. Chem. Chem. Phys.* **15**, 8069–8080.
- Baird, J. A., Van Eerdenbrugh, B. & Taylor, L. S. (2010). *J. Pharm. Sci.* **99**, 3787–3806.
- Bardwell, D. A. *et al.* (2011). *Acta Cryst.* **B67**, 535–551.
- Barnett, S. A., Broder, C. K., Shankland, K., David, W. I. F., Ibberson, R. M. & Tocher, D. A. (2006). *Acta Cryst.* **B62**, 287–295.
- Barnett, S. A., Hulme, A. T., Issa, N., Lewis, T. C., Price, L. S., Tocher, D. A. & Price, S. L. (2008). *New J. Chem.* **32**, 1761–1775.
- Barnett, S. A., Johnson, A., Florence, A. J., Price, S. L. & Tocher, D. A. (2008). *Cryst. Growth Des.* **8**, 24–36.
- Bauer, J., Spanton, S., Henry, R., Quick, J., Dziki, W., Porter, W. & Morris, J. (2001). *Pharm. Res.* **18**, 859–866.
- Bernstein, J. (2002). *Polymorphism in Molecular Crystals*. Oxford: Clarendon Press.
- Bernstein, J. (2011). *Cryst. Growth Des.* **11**, 632–650.
- Bernstein, J., Davey, R. J. & Henck, J. O. (1999). *Angew. Chem. Int. Ed.* **38**, 3441–3461.
- Bernstein, J., Davis, R. E., Shimon, L. & Chang, N. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555–1573.
- Beyer, T., Lewis, T. & Price, S. L. (2001). *CrystEngComm*, **3**, 178–212.
- Bhardwaj, R. M., Price, L. S., Price, S. L., Reutzel-Edens, S. M., Miller, G. J., Oswald, I. D. H., Johnston, B. F. & Florence, A. J. (2013). *Cryst. Growth Des.* **13**, 1602–1617.
- Blagden, N. & Davey, R. J. (2003). *Cryst. Growth Des.* **3**, 873–885.
- Blagden, N., Davey, R., Rowe, R. & Roberts, R. (1998). *Int. J. Pharm.* **172**, 169–177.
- Bond, A. D. (2012). *CrystEngComm*, **14**, 2363–2366.
- Bond, A. D., Boese, R. & Desiraju, G. (2007). *Angew. Chem. Int. Ed.* **46**, 618–622.
- Bond, A. D., Solanko, K. A., Parsons, S., Redder, S. & Boese, R. (2011). *CrystEngComm*, **13**, 399–401.
- Braun, D. E., Ardid-Candel, M., D’Oria, E., Karamertzanis, P. G., Arlin, J. B., Florence, A. J., Jones, A. G. & Price, S. L. (2011). *Cryst. Growth Des.* **11**, 5659–5669.
- Braun, D. E., Bhardwaj, R. M., Florence, A. J., Tocher, D. A. & Price, S. L. (2013). *Cryst. Growth Des.* **13**, 19–23.
- Braun, D. E., Karamertzanis, P. G. & Price, S. L. (2011). *Chem. Commun.* **47**, 5443–5445.
- Braun, D. E., Tocher, D. A., Price, S. L. & Griesser, U. J. (2012). *J. Phys. Chem. B*, **116**, 3961–3972.
- Brittain, H. G. (2009). *Polymorphism in Pharmaceutical Solids*. New York, London: Informa Healthcare.
- Carlucci, L. & Gavezzotti, A. (2005). *Chem. Eur. J.* **11**, 271–279.
- Carter, P. W. & Ward, M. D. (1993). *J. Am. Chem. Soc.* **115**, 11521–11535.
- Chan, H. C. S., Kendrick, J. & Leusen, F. J. J. (2011a). *Phys. Chem. Chem. Phys.* **13**, 20361–20370.
- Chan, H. C. S., Kendrick, J. & Leusen, F. J. J. (2011b). *Angew. Chem. Int. Ed.* **50**, 2979–2981.

- Chan, H. C. S., Kendrick, J., Neumann, M. A. & Leusen, F. J. J. (2013). *CrystEngComm*, **15**, 3799–3807.
- Chen, J., Sarma, B., Evans, J. M. B. & Myerson, A. S. (2011). *Cryst. Growth Des.* **11**, 887–895.
- Childs, S. L., Wood, P. A., Rodríguez-Hornedo, N., Reddy, L. S. & Hardcastle, K. I. (2009). *Cryst. Growth Des.* **9**, 1869–1888.
- Copley, R. C. B., Barnett, S. A., Karamertzanis, P. G., Harris, K. D. M., Kariuki, B. M., Xu, M., Nickels, E. A., Lancaster, R. W. & Price, S. L. (2008). *Cryst. Growth Des.* **8**, 3474–3481.
- Cross, W. I., Blagden, N., Davey, R. J., Pritchard, R. G., Neumann, M. A., Roberts, R. J. & Rowe, R. C. (2003). *Cryst. Growth Des.* **3**, 151–158.
- Cruz-Cabeza, A. J., Day, G. M. & Jones, W. (2008). *Chem. Eur. J.* **14**, 8830–8836.
- Cruz-Cabeza, A. J., Day, G. M. & Jones, W. (2009). *Chem. Eur. J.* **15**, 13033–13040.
- Cruz-Cabeza, A. J., Day, G. M. & Jones, W. (2011). *Phys. Chem. Chem. Phys.* **13**, 12808–12816.
- Cruz Cabeza, A. J., Day, G. M., Motherwell, W. D. S. & Jones, W. (2006). *Cryst. Growth Des.* **6**, 1858–1866.
- Cruz-Cabeza, A. J., Day, G. M., Motherwell, W. D. S. & Jones, W. (2007a). *Chem. Commun.* pp. 1600–1602.
- Cruz-Cabeza, A. J., Day, G. M., Motherwell, W. D. S. & Jones, W. (2007b). *Cryst. Growth Des.* **7**, 100–107.
- Cruz-Cabeza, A. J., Karki, S., Fábrián, L., Frišćić, T., Day, G. M. & Jones, W. (2010). *Chem. Commun.* **46**, 2224–2226.
- Cruz-Cabeza, A. J., Liebeschuetz, J. W. & Allen, F. H. (2012). *CrystEngComm*, **14**, 6797–6811.
- Davey, R. J., Schroeder, S. L. M. & ter Horst, J. H. (2013). *Angew. Chem. Int. Ed.* **52**, 2166–2179.
- Day, G. M. (2011). *Crystalllogr. Rev.* **17**, 3–52.
- Day, G. M., Trask, A. V., Motherwell, W. D. S. & Jones, W. (2006). *Chem. Commun.* pp. 54–56.
- Dean, P. M., Turanjanin, J., Yoshizawa-Fujita, M., MacFarlane, D. R. & Scott, J. L. (2009). *Cryst. Growth Des.* **9**, 1137–1145.
- Desiraju, G. (2007). *Angew. Chem. Int. Ed.* **46**, 8342–8356.
- D’Oria, E., Karamertzanis, P. G. & Price, S. L. (2010). *Cryst. Growth Des.* **10**, 1749–1756.
- Dunitz, J. D. & Bernstein, J. (1995). *Acc. Chem. Res.* **28**, 193–200.
- Eddleston, M. D., Hejczyk, K. E., Bithell, E. G., Day, G. M. & Jones, W. (2013a). *Chem. Eur. J.* **19**, 7883–7888.
- Eddleston, M. D., Hejczyk, K. E., Bithell, E. G., Day, G. M. & Jones, W. (2013b). *Chem. Eur. J.* **19**, 7883–7888.
- Eijck, B. P. van (2002). *Phys. Chem. Chem. Phys.* **4**, 4789–4794.
- Enright, G. D., Terskikh, V. V., Brouwer, D. H. & Ripmeester, J. A. (2007). *Cryst. Growth Des.* **7**, 1406–1410.
- Fabbiani, F. P. & Pulham, C. R. (2006). *Chem. Soc. Rev.* **35**, 932–942.
- Florence, A. J., Johnston, A., Price, S. L., Nowell, H., Kennedy, A. R. & Shankland, N. (2006). *J. Pharm. Sci.* **95**, 1918–1930.
- Florence, A. J., Leech, C. K., Shankland, N., Shankland, K. & Johnston, A. (2006). *CrystEngComm*, **8**, 746–747.
- Gavezzotti, A. (1983). *J. Am. Chem. Soc.* **105**, 5220–5225.
- Gavezzotti, A. (2002). *CrystEngComm*, **4**, 343–347.
- Gavezzotti, A. (2010). *Acta Cryst.* **B66**, 396–406.
- Gelbrich, T., Braun, D. E., Ellern, A. & Griesser, U. J. (2013). *Cryst. Growth Des.* **13**, 1206–1217.
- Gelbrich, T. & Hursthouse, M. B. (2005). *CrystEngComm*, **7**, 324–326.
- Gelbrich, T., Threlfall, T. L. & Hursthouse, M. B. (2012). *CrystEngComm*, **14**, 5454–5464.
- Grzesiak, A. L., Lang, M., Kim, K. & Matzger, A. J. (2003). *J. Pharm. Sci.* **92**, 2260–2271.
- Habgood, M. (2011). *Cryst. Growth Des.* **11**, 3600–3608.
- Habgood, M. (2012). *Phys. Chem. Chem. Phys.* **14**, 9195–9203.
- Habgood, M., Deij, M. A., Mazurek, J., Price, S. L. & ter Horst, J. H. (2010). *Cryst. Growth Des.* **10**, 903–912.
- Habgood, M., Grau-Crespo, R. & Price, S. L. (2011). *Phys. Chem. Chem. Phys.* **13**, 9590–9600.
- Habgood, M., Lancaster, R. W., Gateshki, M. & Kenwright, A. M. (2013). *Cryst. Growth Des.* **13**, 1771–1779.
- Habgood, M. & Price, S. L. (2010). *Cryst. Growth Des.* **10**, 3263–3272.
- Habgood, M., Price, S. L., Portalone, G. & Irrera, S. (2011). *J. Chem. Theory Comput.* **7**, 2685–2688.
- Hammond, R. B., Pencheva, K. & Roberts, K. J. (2012). *CrystEngComm*, **14**, 1069–1082.
- Hilfiker, R. (2006). *Polymorphism in the Pharmaceutical Industry*. Weinheim: Wiley-VCH.
- Hulme, A. T., Johnston, A., Florence, A. J., Fernandes, P., Shankland, K., Bedford, C. T., Welch, G. W., Sadiq, G., Haynes, D. A., Motherwell, W. D., Tocher, D. A. & Price, S. L. (2007). *J. Am. Chem. Soc.* **129**, 3649–3657.
- Hulme, A. T. & Price, S. L. (2007). *J. Chem. Theory Comput.* **3**, 1597–1608.
- Hulme, A. T., Price, S. L. & Tocher, D. A. (2005). *J. Am. Chem. Soc.* **127**, 1116–1117.
- Hursthouse, M. B., Huth, L. S. & Threlfall, T. L. (2009). *Org. Process Res. Dev.* **13**, 1231–1240.
- Ismail, S. Z., Anderton, C. L., Copley, R. C. B., Price, L. S. & Price, S. L. (2013). *Cryst. Growth Des.* **13**, 2396–2406.
- Issa, N., Barnett, S. A., Mohamed, S., Braun, D. E., Copley, R. C. B., Tocher, D. A. & Price, S. L. (2012). *CrystEngComm*, **14**, 2454–2464.
- Issa, N., Karamertzanis, P. G., Welch, G. W. A. & Price, S. L. (2009). *Cryst. Growth Des.* **9**, 442–453.
- Jetti, R. K. R., Boese, R., Sarma, J. A. R. P., Reddy, L. S., Vishweshwar, P. & Desiraju, G. R. (2003). *Angew. Chem. Int. Ed.* **42**, 1963–1967.
- Johnston, A., Bardin, J., Johnston, B. F., Fernandes, P., Kennedy, A. R., Price, S. L. & Florence, A. J. (2011). *Cryst. Growth Des.* **11**, 405–413.
- Johnston, A., Florence, A. J., Shankland, N., Kennedy, A. R., Shankland, K. & Price, S. L. (2007). *Cryst. Growth Des.* **7**, 705–712.
- Jones, J. T., Hasell, T., Wu, X., Bacsa, J., Jelfs, K. E., Schmidtman, M., Chong, S. Y., Adams, D. J., Trewin, A., Schiffman, F., Cora, F., Slater, B., Steiner, A., Day, G. M. & Cooper, A. I. (2011). *Nature*, **474**, 367–371.
- Karamertzanis, P. G., Day, G. M., Welch, G. W. A., Kendrick, J., Leusen, F. J. J., Neumann, M. A. & Price, S. L. (2008). *J. Chem. Phys.* **128**, 244708–244717.
- Karamertzanis, P. G., Kazantsev, A. V., Issa, N., Welch, G. W., Adjiman, C. S., Pantelides, C. C. & Price, S. L. (2009). *J. Chem. Theory Comput.* **5**, 1432–1448.
- Karamertzanis, P. G., Raiteri, P., Parrinello, M., Leslie, M. & Price, S. L. (2008). *J. Phys. Chem. B.* **112**, 4298–4308.
- Kazantsev, A. V., Karamertzanis, P. G., Adjiman, C. S. & Pantelides, C. C. (2011). *J. Chem. Theory Comput.* **7**, 1998–2016.
- Kazantsev, A. V., Karamertzanis, P. G., Adjiman, C. S., Pantelides, C. C., Price, S. L., Galek, P. T., Day, G. M. & Cruz-Cabeza, A. J. (2011). *Int. J. Pharm.* **418**, 168–178.
- Kendrick, J., Leusen, F. J., Neumann, M. A. & van de Streek, J. (2011). *Chem. Eur. J.* **17**, 10735–10743.
- Kendrick, J., Stephenson, G. A., Neumann, M. A. & Leusen, F. J. J. (2013). *Cryst. Growth Des.* **13**, 581–589.
- Kuhnert-Brandstatter, M. (1982). *Thermomicroscopy of Organic Compounds*, Vol. XVI, *Comprehensive Analytical Chemistry*, edited by G. Svelha, pp. 329–491. Amsterdam: Elsevier Scientific Publishing.
- Lancaster, R. W., Harris, L. D. & Pearson, D. (2011). *CrystEngComm*, **13**, 1775–1777.
- Lancaster, R. W., Karamertzanis, P. G., Hulme, A. T., Tocher, D. A., Covey, D. F. & Price, S. L. (2006). *Chem. Commun.* pp. 4921–4923.
- Lancaster, R. W., Karamertzanis, P. G., Hulme, A. T., Tocher, D. A., Lewis, T. C. & Price, S. L. (2007). *J. Pharm. Sci.* **96**, 3419–3431.
- Lewis, T. C., Tocher, D. A. & Price, S. L. (2005). *Cryst. Growth Des.* **5**, 983–993.
- Li, R., Zeitler, J. A., Tomerini, D., Parrott, E. P., Gladden, L. F. & Day, G. M. (2010). *Phys. Chem. Chem. Phys.* **12**, 5329–5340.

- Llinàs, A. & Goodman, J. M. (2008). *Drug Discov. Today*, **13**, 198–210.
- López-Mejías, V., Kampf, J. W. & Matzger, A. J. (2009). *J. Am. Chem. Soc.* **131**, 4554–4555.
- López-Mejías, V., Kampf, J. W. & Matzger, A. J. (2012). *J. Am. Chem. Soc.* **134**, 9872–9875.
- Lutker, K. M., Tolstyka, Z. P. & Matzger, A. J. (2008). *Cryst. Growth Des.* **8**, 136–139.
- Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
- McCrone, W. C. (1965). *Polymorphism*, Vol. II, *Physics and Chemistry of the Organic Solid State*, edited by D. Fox, M. M. Labes & A. Weissberger, pp. 725–767. New York: Wiley Interscience.
- Mnyukh, Y. (2001). *Fundamentals of Solid-State Phase Transitions, Ferromagnetism and Ferroelectricity*. Bloomington: 1st Books Library.
- Mohamed, S., Tocher, D. A. & Price, S. L. (2011). *Int. J. Pharm.* **418**, 187–198.
- Mohamed, S., Tocher, D. A., Vickers, M., Karamertzanis, P. G. & Price, S. L. (2009). *Cryst. Growth Des.* **9**, 2881–2889.
- Montis, R., Hursthouse, M. B., Chan, H. C. S., Kendrick, J. & Leusen, F. J. J. (2012). *CrystEngComm*, **14**, 1672–1680.
- Motherwell, W. D. S. *et al.* (2002). *Acta Cryst.* **B58**, 647–661.
- Neumann, M. A. & Perrin, M. A. (2005). *J. Phys. Chem. B*, **109**, 15531–15541.
- Noorduyn, W. L., Izumi, T., Millemaggi, A., Leeman, M., Meeke, H., Van Enckevort, W. J., Kellogg, R. M., Kaptein, B., Vlieg, E. & Blackmond, D. G. (2008). *J. Am. Chem. Soc.* **130**, 1158–1159.
- Oliveira, M. A., Peterson, M. L. & Davey, R. J. (2011). *Cryst. Growth Des.* **11**, 449–457.
- Ouvrard, C. & Price, S. L. (2004). *Cryst. Growth Des.* **4**, 1119–1127.
- Perrin, M. A., Neumann, M. A., Elmaleh, H. & Zasko, L. (2009). *Chem. Commun.* pp. 3181–3183.
- Price, S. L. (2008a). *Int. Rev. Phys. Chem.* **27**, 541–568.
- Price, S. L. (2008b). *Phys. Chem. Chem. Phys.* **10**, 1996–2009.
- Price, S. S. (2009). *Acc. Chem. Res.* **42**, 117–126.
- Price, S. L., Leslie, M., Welch, G. W., Habgood, M., Price, L. S., Karamertzanis, P. G. & Day, G. M. (2010). *Phys. Chem. Chem. Phys.* **12**, 8478–8490.
- Prohens, R., Font-Bardia, M. & Barbas, R. (2013). *CrystEngComm*, **15**, 845–847.
- Raiteri, P., Martoňák, R. & Parrinello, M. (2005). *Angew. Chem. Int. Ed.* **44**, 3769–3773.
- Ridout, J. & Probert, M. J. (2013). *Cryst. Growth Des.* **13**, 1943–1948.
- Rohl, A. L. (2003). *Curr. Opin. Solid State Mater. Sci.* **7**, 21–26.
- Roy, S. & Matzger, A. (2009). *Angew. Chem. Int. Ed.* **48**, 8505–8508.
- Salager, E., Day, G. M., Stein, R. S., Pickard, C. J., Elena, B. & Emsley, L. (2010). *J. Am. Chem. Soc.* **132**, 2564–2566.
- Schmidt, M. U. (1999). *Crystal Engineering: From Molecules and Crystals to Materials*. Berlin: Springer.
- Schmidt, M. U., Brüning, J., Glinnemann, J., Hützler, M. W., Mörschel, P., Ivashevskaya, S. N., van de Streek, J., Braga, D., Maini, L., Chierotti, M. R. & Gobetto, R. (2011). *Angew. Chem. Int. Ed.* **50**, 7924–7926.
- Schmidt, M. U. & Glinnemann, J. (2012). *Z. Kristallogr.* **227**, 805–817.
- Schmidt, M. U., Paulus, E. F., Rademacher, N. & Day, G. M. (2010). *Acta Cryst.* **B66**, 515–526.
- Schmidt, C. & Ulrich, J. (2012). *Chem. Eng. Technol.* **35**, 1009–1012.
- Spencer, J., Patel, H., Deadman, J. J., Palmer, R. A., Male, L., Coles, S. J., Uzoh, O. G. & Price, S. L. (2012). *CrystEngComm*, **14**, 6441–6446.
- Stahly, G. P. (2007). *Cryst. Growth Des.* **7**, 1007–1026.
- Storey, R. A. & Ymén, I. (2012). *Solid State Characterization of Pharmaceuticals*. Chichester: Wiley.
- Streek, J. van de & Motherwell, S. (2005). *Acta Cryst.* **B61**, 504–510.
- Threlfall, T. (2000). *Org. Process Res. Dev.* **4**, 384–390.
- Threlfall, T. (2003). *Org. Process Res. Dev.* **7**, 1017–1027.
- Torriani, A., Leech, C. K., Shankland, K., David, W. I., Ibberson, R. M., Benet-Buchholz, J., Boese, R., Leslie, M., Catlow, C. R. & Price, S. L. (2008). *J. Phys. Chem. B*, **112**, 3746–3758.
- Tothadi, S. & Desiraju, G. R. (2012). *Philos. Trans. R. Soc. Lond. A*, **370**, 2900–2915.
- Tremayne, M., Grice, L., Pyatt, J. C., Seaton, C. C., Kariuki, B. M., Tsui, H. H., Price, S. L. & Cherryman, J. C. (2004). *J. Am. Chem. Soc.* **126**, 7071–7081.
- Uzoh, O. G., Cruz-Cabeza, A. J. & Price, S. L. (2012). *Cryst. Growth Des.* **12**, 4230–4239.
- Vasileiadis, M., Kazantsev, A. V., Karamertzanis, P. G., Adjiman, C. S. & Pantelides, C. C. (2012). *Acta Cryst.* **B68**, 677–685.
- Vishweshwar, P., McMahon, J. A., Oliveira, M., Peterson, M. L. & Zaworotko, M. J. (2005). *J. Am. Chem. Soc.* **127**, 16802–16803.
- Ward, M. D. & Raithby, P. R. (2013). *Chem. Soc. Rev.* **42**, 1619–1636.
- Welch, G. W. A., Karamertzanis, P. G., Misquitta, A. J., Stone, A. J. & Price, S. L. (2008). *J. Chem. Theory Comput.* **4**, 522–532.
- Winkel, K., Hage, W., Loerting, T., Price, S. L. & Mayer, E. (2007). *J. Am. Chem. Soc.* **129**, 13863–13871.
- Wu, H., Habgood, M., Parker, J. E., Reeves-McLaren, N., Cockcroft, J. K., Vickers, M., West, A. R. & Jones, A. G. (2013). *CrystEngComm*, **15**, 1853–1859.
- Yu, L. (2007). *CrystEngComm*, **9**, 847–851.
- Yu, L. (2010). *Acc. Chem. Res.* **43**, 1257–1266.