JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. XXXX, XXX, XXX−XXX
pubs.acs.org/JACS

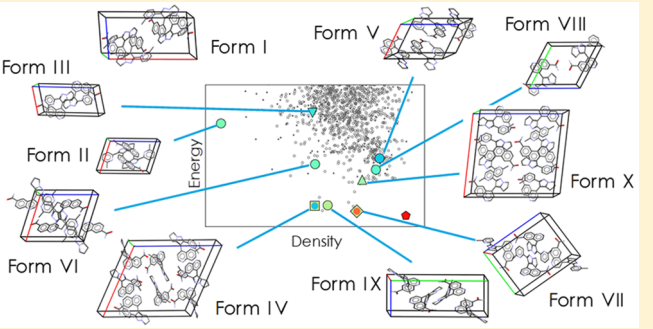
A Prolific Solvate Former, Galunisertib, under the Pressure of Crystal Structure Prediction, Produces Ten Diverse Polymorphs

Rajni M. Bhardwaj,[†] Jennifer A. McMahon,[†] Jonas Nyman,[§,†] Louise S. Price,[‡] Sumit Konar,[⊥] Iain D. H. Oswald,[∥] Colin R. Pulham,[⊥] Sarah L. Price,[‡] and Susan M. Reutzel-Edens*[,†]

[†]Small Molecule Design & Development, Eli Lilly and Company, Indianapolis, Indiana 46285, United States
[‡]Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.
[§]School of Pharmacy, University of Wisconsin—Madison, 777 Highland Avenue, Madison, Wisconsin 53705, United States
[∥]Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral St, Glasgow G4 0RE, U.K.
[⊥]School of Chemistry, University of Edinburgh, David Brewster Road, Edinburgh EH9 3FJ, U.K.

Supporting Information

ABSTRACT: The solid form screening of galunisertib produced many solvates, prompting an extensive investigation into possible risks to the development of the favored monohydrate form. Inspired by crystal structure prediction, the search for neat polymorphs was expanded to an unusual range of experiments, including melt crystallization under pressure, to work around solvate formation and the thermal instability of the molecule. Ten polymorphs of galunisertib were found; however, the structure predicted to be the most stable has yet to be obtained. We present the crystal structures of all ten unsolvated polymorphs of galunisertib, showing how state-of-the-art characterization methods can be combined with emerging computational modeling techniques to produce a complete structure landscape and assess the risk of late-appearing, more stable polymorphs. The exceptional conformational polymorphism of this prolific solvate former invites further development of methods, computational and experimental, that are applicable to larger, flexible molecules with complex solid form landscapes.



INTRODUCTION

Crystalline drugs are almost exclusively delivered to patients as either neat (nonsolvated) forms or, alternatively, as hydrates.[1] Which crystal form is ultimately developed into a drug product depends on how much is known about how the drug crystallizes and the properties of its neat and hydrated forms.[2−5] Experimental solid form screening, generally the first step taken to transform a molecule to a medicine, aims to identify as many solid forms as possible with the hope that one will be commercially viable, as well as to avoid the potentially catastrophic consequences of late-appearing forms.[6−9] Identifying important polymorphs and developing crystallization processes that target them can be especially complicated, however, for molecules which show a high propensity to form solvates.[10] Some neat forms will be missed altogether because they do not nucleate and grow from solvate-forming solvents, while the appearance of others may rely exclusively on crystallizing a solvated precursor. In fact, it is not uncommon for a desolvate to be kinetically trapped upon removal of the product solvate from the crystallizing solution or during drying.

The appearance of solvates during solid form screening will usually alter the strategy taken to find other crystal forms, expanding the search in some areas and potentially contracting it in others. Desolvation and solvent exchange, for example, are experimental approaches that cannot be anticipated in the initial design of experiments but which will need to be exploited to respectively target neat and hydrated crystal forms originating from solvates discovered during the screen. This strategy has proven successful in generating many new forms[11] and may be the only path to a suitable crystalline form. The appearance of a solvate may also prompt screening efforts to be redirected away from that solvent, effectively limiting the search space.[12] However, for prolific solvate-forming compounds, solvents cannot simply be avoided for having formed a solvate; other approaches must be considered to ensure that the search for solid forms is comprehensive and rigorous. Among these, high temperatures are commonly used to increase the probability of crystallizing nonsolvated forms from solvate-forming solvents. It has been proposed for some systems that conformational energy barriers preventing nucleation of stable forms may be more easily overcome at elevated temperatures.[13] Unfortunately, important forms are not always attainable at high temperatures, and many

Received: June 22, 2019
Published: August 9, 2019

compounds are not sufficiently stable for this approach to be feasible.

Given the limitations on experimental screening imposed by solvate formation, having the ability to calculate how a prolific solvate-forming molecule might crystallize on its own would be especially useful for assessing the risk of polymorphism. In recent years, computational crystal structure prediction (CSP) has been explored as a complement to experimental solid form screening, helping to structurally characterize observed solid forms,[14,15] to understand crystallization behaviors at a molecular level, and to decide when it might be safe to stop screening.[16−18] CSP has previously been performed on drug molecules known to form many solvates, such as olanzapine[19] and axitinib,[20] where it proved to be of great utility in explaining experimental observations.

In this work, we explore the solid form landscape of a prolific solvate former, galunisertib (**GAL**, 4-[2-(6-methylpyridin-2-yl)-5,6-dihydro-4*H*-pyrrolo[1,2-*b*]pyrazol-3-yl]quinoline-6-carboxamide, Figure 1). **GAL**, a small-molecule inhibitor of transforming growth factor beta (TGF-β) receptor I that specifically downregulates pSMAD2, was recently investigated for use in the oral treatment of advanced metastatic malignancies, including glioblastoma, pancreatic cancer, and hepatocellular carcinoma.[21] Solid form screening quickly revealed a particularly strong tendency of the TGF-β inhibitor to form solvates with a wide range of solvents. Neat crystal forms were initially found by desolvation, however, conditions were later identified to crystallize many of the nonsolvated polymorphs directly from solution. As shown in this paper, **GAL** was eventually crystallized in more than 50 solvates, one of which was a monohydrate, and 10 neat polymorphs. The monohydrate was chosen for commercial development, based on its ease of crystallization, apparent thermodynamic stability, and solid-state properties.

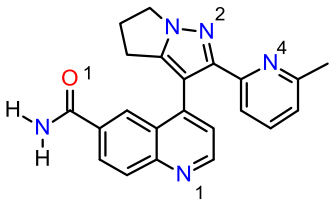


Figure 1. Molecular structure of galunisertib. The hydrogen-bonding acceptors are numbered, and the hydrogens shown are the hydrogen-bond donors.

Having encountered many roadblocks with prolific solvate formation and chemical instability at higher temperatures in the attempts to experimentally find its neat forms, **GAL** was an ideal candidate to explore the use of CSP as a complement to solid form screening. Could CSP be successful in suggesting structures of observed forms for which single crystals suitable for X-ray diffraction could not be obtained? Could computational modeling and solid form informatics aid in understanding the properties of the solid forms? Could such methods help to rationalize desolvation pathways from solvates to neat polymorphs? Are structures of greater stability predicted, and if so, what assurance is there that no such polymorph exists? In this paper, we describe the herculean efforts to isolate and characterize the many neat forms of galunisertib over the course of more than 10 years. The comprehensive determination of ten polymorph crystal structures required going beyond conventional single-crystal X-ray diffraction to leverage both state-of-the-art powder characterization of carefully prepared polycrystalline solids and computational chemistry, including CSP. The CSP study also suggested that specialist experiments at higher pressures should be performed. In setting a new standard for interdisciplinary solid form screening and characterization that can be applied to all organic materials, we have uncovered in **GAL** a truly remarkable system of conformational polymorphs.

RESULTS

Experimental Solid-Form Landscape. **GAL** was recrystallized in search of neat polymorphs and solvates using industry standard solvent-based techniques,[22] and following the discovery of numerous solvates, by thermal and relative humidity (RH) annealing to promote desolvation and/or solvent exchange. The conventional solid form screen used the crystalline monohydrate as the starting material and encompassed more than 800 experiments. In this phase of the solid form screen, nine neat polymorphs (I−IX), along with more than 50 solvates with diverse solvents, were found. Many solvated forms were crystallized from solution as the stable form, while others crystallized as kinetic forms. With such a large number of solvates having readily appeared in the solid form screen, including solvate polymorphs and solvates of different stoichiometry, the experimental effort was ultimately directed to confirming the identity of the forms that were found, characterizing their solid form properties and establishing their transformation (desolvation) pathways. Thus, the solid form screen, while extensive, was known to be far from all-encompassing, with the most obvious deficiency being an inability to recrystallize at high temperatures, owing to the chemical instability of the drug substance. Details of the experimental screen are provided in the Supporting Information (SI).

The thermodynamically most stable form of **GAL** at ambient temperature over a wide RH range is the monohydrate. This form preferentially crystallizes in the presence of even small amounts of water. The monohydrate is, however, susceptible to dehydration at higher temperatures, with conversion to different neat polymorphs depending on the drying conditions (SI Section 1.3). At low relative humidity (<5% RH) and with modest increases in temperature (up to 50 °C), an isomorphic dehydrate, form I, is obtained. This neat polymorph is highly unstable, rapidly reverting to the monohydrate above 5% RH at ambient temperature. Drying the monohydrate at higher temperatures (60 °C and above) sees a change in dehydration pathway, kinetically trapping **GAL** in forms III, IV, and VI, usually as phase mixtures.

Forms IV and VI, the most stable of the neat polymorphs, were crystallized directly, and frequently concomitantly, from a variety of solvents. Desolvation also afforded a viable route to these polymorphic forms, as virtually every solvate could be desolvated under aggressive drying conditions to one or both of these forms (SI Section 1.3). By contrast, forms II, III, V, VII, and VIII were produced by drying specific solvates and only under comparatively mild conditions. While form II was obtained from a tetrahydrofuran solvate, one of a large family of isostructural solvates, forms III and V were only obtained with mild drying of ethanol and methanol solvates, respectively. Likewise, forms VII and VIII were produced by desolvating acetonitrile and acetone solvates, again under mild conditions. Form IX is a highly elusive polymorph, crystallized

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

from solution in only milligram quantities on the few occasions it was observed. That form IX was only ever observed in phase-pure form suggests this metastable form may be a casualty of Ostwald's rule of stages, rapidly disappearing with the nucleation of other more stable forms, usually forms IV and VI.

The search for neat **GAL** polymorphs was expanded when a first CSP study predicted as the global minimum (GM) a substantially more stable, high density structure than forms I–IX (SI Figure S25). This suggested crystallizing **GAL** under pressure on the basis that higher density forms have been found at elevated pressures (cf. dalcetrapib[23] and γ-aminobutyric acid[24]).[25,26] By specifically introducing pressure as an experimental variable, the search for neat polymorphs at higher temperatures was also reconsidered, with past experience having shown that thermal decomposition of molecular compounds can often be inhibited at elevated pressures.[27] No new phases were obtained from numerous attempts to grow **GAL** crystals from solution at elevated pressures (1–6 GPa) (SI Section 1.4); however, a new polymorph (form X) was crystallized by slowly cooling a melt (in the absence of any pressure-transmitting medium or solvent) from 240 °C to ambient temperature under modest pressure (0.4 GPa) in a diamond anvil cell (SI Figure S5). Success in crystallizing form X likely hinged on using form V, one of the lowest melting **GAL** polymorphs, to generate the melt at the lowest possible temperature so as to minimize chemical decomposition. On decompression to ambient pressure, the crystal of form X survived and so it was possible to obtain SCRXD data at ambient pressure. It is worth noting that form X was not the GM structure. In fact, it was not even a high density form.

Figure 2 shows the powder X-ray diffraction (PXRD) patterns of the neat **GAL** polymorphs, as well as the monohydrate. All of the crystal forms, except the isostructural monohydrate and form I, give characteristic PXRD patterns, and these two exceptions are readily distinguished by their solid-state NMR (ssNMR) spectra (SI Section 2.3). Powder X-ray diffraction was used to identify crystal forms generated throughout the solid form screen, and eventually, we were able to establish (to within the detection limit of PXRD) that most of the neat polymorphs had been generated in phase-pure form by comparison with the simulated powder patterns. The experimental solid form landscape, showing the confirmed neat and hydrated forms, routes to their production and selected interconversion pathways, is compiled in Figure 3.

Crystal Structure Prediction. A Z′ = 1 crystal energy landscape (CSP1) was commissioned for **GAL** to complement the experimental solid form screen at the point when only forms I–VI had been detected. The CrystalPredictor and CrystalOptimizer[28] programs were used in CSP1 to survey crystal packing possibilities in 59 space groups for a range of **GAL** conformations, defined primarily by the three internal (amide–quinoline, quinolone–pyrazole, and pyrazole–pyridine) torsion angles (SI Section 3.1).[29] Puckering of the 5-membered dihydropyrrole ring was not included as a search variable in the CSP1 study; only later was this seemingly small molecular change found to have a surprisingly large effect on the calculated lattice energy of some crystal structures.

Plausible structure candidates were found for forms II, III, V, and VI, and later forms VIII–X, in the CSP1 crystal energy landscape (SI Figure S25) by comparison of PXRD patterns, albeit some structures had the 5-membered ring as planar or in the opposite orientation. Highly metastable form I was not among the low energy CSP1 structures, and being Z′ = 2 polymorphs, forms IV and VII were out of the scope of this search. In retrospect, it should not have been surprising that the structures corresponding to the desolvate forms II, III, and V were comparatively high in energy. The other experimentally observed polymorphs (VI, VIII, and X) were also not among the lowest energy structures on the CSP1 landscape, where instead a substantially more stable, high density crystal packing arrangement (GM) was seen as the global lattice energy minimum. When **GAL** forms I–VI were shown by several commonly used DFT+D energy methods (Figure 6) to be appreciably less stable (higher lattice energy) than the GM structure, the pressure was on, literally and figuratively, to produce this structure in the laboratory.

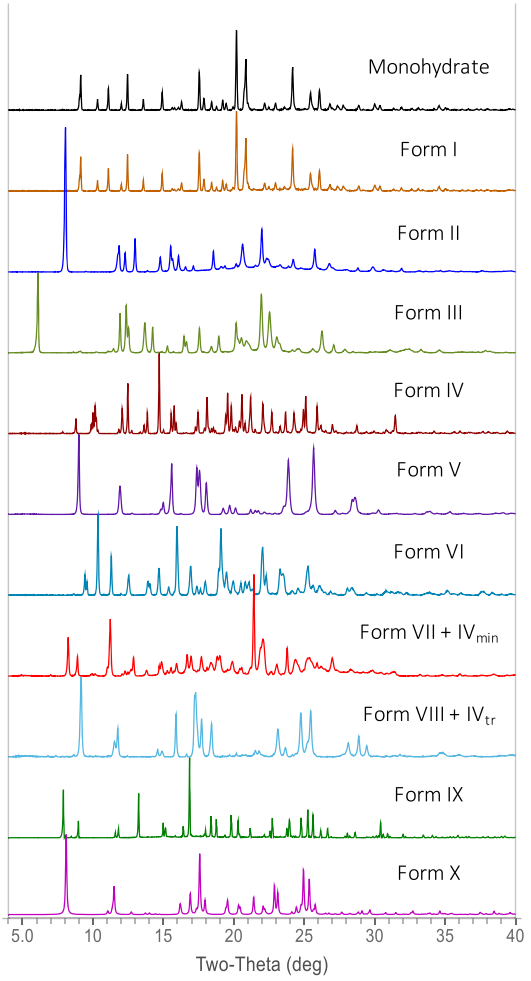


Figure 2. Powder X-ray diffractograms (λ = 1.5406 Å, y axis = intensity in arbitrary units) of galunisertib monohydrate and forms I–X. The form X powder pattern was simulated from SCXRD data. All patterns represent pure phases, except for forms VII and VIII, which are contaminated with form IV.

A second CSP study (CSP2) was commissioned following the discovery of forms VII–IX and subsequent determination by ssNMR spectroscopy that form VII is a Z′ = 2 polymorph (SI Figure S14). The CSP2 study used GRACE[30] to sample **GAL** crystal packing in space groups with Z′ = 1 or Z′ = 2 that account for 99.9% of the observed distribution in the Cambridge Structural Database (SI Section 3.2).[31] The parallel tempering approach implemented in GRACE does not place any restrictions on the molecular conformation when

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

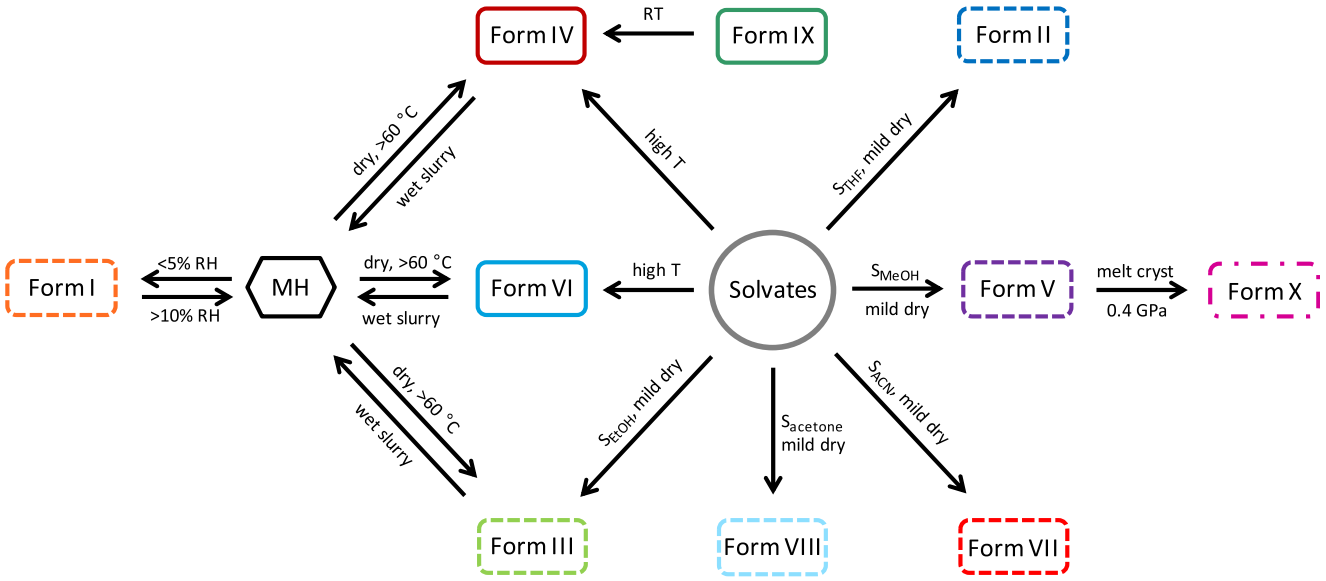


Figure 3. Experimental solid form landscape of galunisertib, summarizing the forms discovered during solid form screening and best routes to their production and interconversion pathways. Solid boxes show forms which were crystallized from solution; dotted boxes show forms obtained by desolvation; the dash-dotted box is the form grown from the melt at modest pressure. MH is the monohydrate, $S_{solvent}$ is a solvate.

generating trial crystal structures, which meant that both dihydropyrrole ring orientations were sampled in the CSP2 search. The CSP2 Z′ = 1 and Z′ = 2 crystal energy landscape calculated using the Perdew–Burke–Ernzerhof (PBE) functional and an empirical pairwise atom–atom dispersion model (NP)[32] is shown in Figure 4.

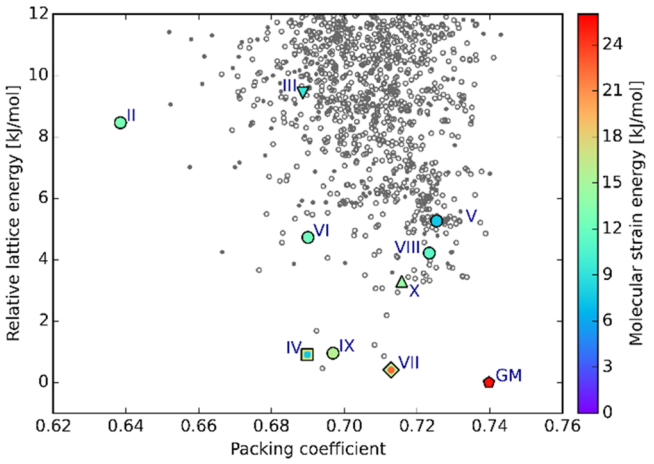


Figure 4. Crystal energy landscape of galunisertib calculated using GRACE. Unobserved crystal packings are shown in gray for Z′ = 1 (●) and Z′ = 2 (○) structures. Experimentally observed neat polymorphs (II–X) and the global minimum (GM) structure are highlighted with symbols, colored by their calculated molecular strain energy and in shapes that represent the nearest conformational energy minimum.

The inclusion of Z′ = 2 structures in the search dramatically increased the number of lattice energy minima found on the crystal energy landscape of **GAL**. All of the neat polymorphs, except form I, were found within 10 kJ mol^{-1} of the same global minimum structure as found in CSP1. However, the PBE+NP energy model used in CSP2 ranked both Z′ = 2 polymorphs (IV and VII) and form IX (in the correct hydropyrrole ring puckering) closer in energy, and within the estimated margin of error, to the GM and a few low energy Z′ = 2 structures.

Crystal Structure Analysis of Neat Polymorphs. Solving the crystal structures of all ten **GAL** polymorphs was an enormous task, not only requiring a concerted effort to grow single crystals or generate highly crystalline, phase pure polycrystalline samples of each of the forms, but also leveraging every experimental tool and computational method at our disposal. The crystal structures of forms II, III, IV, VI, and X were solved from SCXRD data using single crystals grown by a variety of methods. Whereas single crystals of forms IV and VI were grown directly from solution, form X grew from the melt as a large single crystal inside a diamond anvil cell, permitting the structure to be solved at pressure. Having survived decompression, the form X structure could also be solved from full sets of diffraction data collected at ambient pressure. Translucent single crystals of forms II and III suitable for SCXRD were, respectively, produced upon air drying toluene and ethanol solvate crystals. CSP, in finding excellent matches ($RMSD_{20}$ < 0.3 Å) to the form II, III, IV, VI, and X SCXRD structures on the computed crystal energy landscape(s), confirmed that these structures are lattice energy minima.

PXRD was used to solve the crystal structures of the remaining **GAL** polymorphs for which single crystals could not be grown. For forms I, V, and IX, the crystal structures were solved from transmission PXRD patterns collected for phase-pure samples (Figure 2, SI Section 2.2). However, with neither form VII nor VIII having been produced as a pure phase, the PXRD patterns of these polymorphs could not be indexed, let alone used to solve their structures without the help of CSP. For these polymorphs, a powder pattern matching algorithm[33] in GRACE was used to identify one distinct match to the form VII PXRD pattern and several pseudosymmetric matches to the form VIII PXRD pattern (SI Figure S28). The form VII structure model derived by CSP2 helped to identify form IV as the main phase impurity in the form VII material, allowing the structure of form VII to be confirmed through mixed-phase

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

Rietveld refinement (SI Figure S11). The best candidate for form VIII was ultimately identified by combining Pawley-type refinement of the unit cell parameters of the calculated structures against the best available RT PXRD pattern (SI Figure S12), with a comparison of the calculated ssNMR spectrum of each candidate to the experimental ssNMR spectrum of form VIII (SI Figure S29). This structure candidate was used as the starting model for a Rietveld refinement to obtain the form VIII crystal structure. Details of all crystal structure determinations, along with overlays showing the agreement between the experimental and computer-generated structures of the ten neat polymorphs, are provided in the SI (Sections 2.1 and 2.2, Figure S30).

The crystal structures of the **GAL** polymorphs show just how diverse the crystal packing possibilities can be for a flexible pharmaceutical,[34] capable of adopting a wide range of conformations with very different relative positions of the hydrogen-bonding donor and acceptor groups, Figure 5. **GAL**,

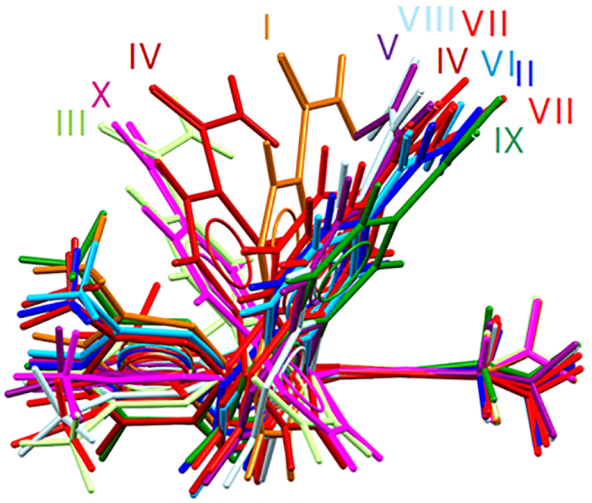


Figure 5. Overlay of molecular conformations in experimentally observed galunisertib forms I–X. Of the two mirror-related conformations present in each structure, only those with the same handedness are shown.

which itself is achiral, adopts chiral (enantiomeric) conformations that together can crystallize in centrosymmetric space groups. In fact, inversion-related amide dimers were observed in all of the structures, except form VI, and with additional hydrogen bonding, a variety of topologies, including 1D ribbons and 2D layers, formed (SI Figure S33). Despite the awkward molecular shape of **GAL**, the 1D ribbons and 2D layers that it forms are reasonably close packed in all of the neat polymorphs, except form II, with packing indices ranging from 0.68 to 0.73 (Figure 4). Form II is a uniquely porous open framework structure (SI Figure S32).

To better understand how efficiently the hydrogen bonding sites of **GAL** are used in the neat polymorphs, the hydrogen bond donor–acceptor pairings in each crystal structure were analyzed by the logit hydrogen-bonding propensity method (SI Section 4.2).[35,36] The hydrogen bond acceptors in the **GAL** molecule could be ranked as follows: quinoline N1 > pyridine N4 ≈ amide O1 > pyrazoline N2 (Figure 1). Interestingly, all of the donor–acceptor pairings gave reasonably high propensity values, yet except for form VII, only three of the acceptors (quinoline N1, pyrazoline N2, amide O1) are used in any of the structures. The fourth acceptor, pyridine N4, despite being the second most basic site, is only involved in hydrogen bonding for one of the symmetry-independent molecules in form VII. The reason why such a strong acceptor as pyridine N4 does not participate in hydrogen bonding is apparent from the full interaction maps[37] (FIMs) generated for **GAL** in the neat forms. FIMs, which are intermolecular interaction probability maps used to analyze the degree to which hydrogen-bonding interactions are satisfied in crystal structures, were generated using uncharged NH and carbonyl O acceptor probes (SI Figure S34). The FIMs show that the very different conformations of **GAL** all have the amide donor–acceptor pair, quinolone N1, and pyrazoline N2 around the periphery of the molecule and available for hydrogen bonding. Importantly, the FIMs show that hydrogen-bonding to the pyridine N4, which is inward-facing and sterically hindered in all but one of the observed crystal conformations, is not statistically favorable.

Thermodynamic Stability. The thermal properties of **GAL** forms I–VII were examined by DSC using fast heating (50 °C min^{-1}) to minimize decomposition. The melting points of these seven polymorphs spanned a range of more than 125 °C, with form VI appearing to be the most stable (highest melting) form at high temperatures (SI Figure S17). Melting enthalpies were reliably measured for forms II–VII, for which the heats of fusion varied by up to 4 kJ mol^{-1} (Figure 6). The measured enthalpies of desolvate forms II and III were consistently lower than those of forms IV–VII, which were within the margin of error of one another (SI Table S12).

The DSC measurements were complemented with calculations of relative lattice energies using various periodic dispersion corrected DFT methods. Figure 6 shows how different dispersion corrections lead to different energy

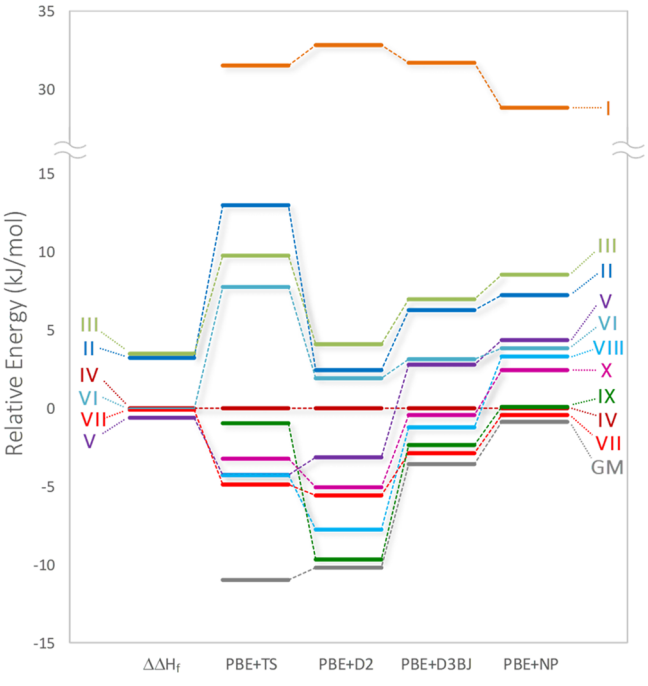


Figure 6. Experimental relative enthalpies of the **GAL** polymorphs contrasted with relative lattice energies of the crystal forms and the CSP-generated GM structure, calculated using various periodic DFT+D methods. All polymorphs are ranked in energy relative to the room temperature stable polymorph, form IV.

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

rankings of the experimentally observed and GM hypothetical structures. The calculations neglect zero point energies and heat capacity contributions to the measured heats of fusion. Form I is much higher in energy than the other neat polymorphs, while forms II and III are, by all methods, consistently higher in energy, in agreement with experiment. With the PBE+NP model used in CSP2, the GM structure, at least at 0 K, is comparable in energy to form IV, the thermodynamically most stable polymorph at ambient temperature.

DISCUSSION

Our exploration of the solid form landscape of **GAL** has uncovered a spectacular polymorphic system. With crystal structures having been determined for all ten of the neat forms, **GAL** is unsurpassed in having the largest number of known polymorph structures. The previous record holders, flufenamic acid,[38] aripiprazole,[39] and ROY,[40,41] each have evidence of further structures and are hence more polymorphic than **GAL** but to date have nine solved structures each. Solving the crystal structures of ten polymorphs was, in this case, no small feat, as single crystals or clearly phase pure samples were simply not attainable for many of the forms. However, with patience, persistence, and a little luck to generate the best samples (single crystals or powders) possible, along with access to state-of-the-art analytical tools and recently developed computational methods, we were able to determine the crystal structure of every experimentally observed polymorph. This unprecedented access to crystallographic information has paved the way for a detailed analysis of the structures in relation to their properties, as well as an assessment of the completeness of the experimental search for neat polymorphs by comparison to structures on computed crystal energy landscapes.

Role of Solvents in Polymorph Appearance. McCrone postulated many years ago that the number of polymorphic forms found is generally proportional to the effort expended in finding them.[42] Still, seeing ten true polymorphs of any molecule is rare, even for those which have been intensively screened.[12,43] A substantially greater effort had gone into finding different polymorphs of **GAL** than, for example, ROY, where as many as 7 of 11 polymorphs have crystallized concomitantly from the same liquid.[41,44,45] Yet, we did not resort to exotic techniques, such as heteronucleation onto isostructural templates[46,47] or polymers,[48] crystallization from complex mixtures,[49,50] or nucleation under confinement[51,52] to produce ten different polymorphs. All of the neat forms of **GAL**, with the notable exception of form X, were produced by conventional methods adapted to the formation of multiple solvates.

What sets **GAL** apart from other highly polymorphic molecules is its ability to form many structurally diverse packings with a wide range of solvents used in crystallization and the relative ease with which the smaller solvents can be liberated from the crystal structures upon harvesting and/or drying the crystals. This allowed for a variety of nonsolvated forms (I–III, V, VII–VIII) to be kinetically trapped by desolvation.[53] The structural relationship between some parent solvates and daughter polymorphs, which might be expected to lower the kinetic barriers to transformation along desolvation pathways that conserve common features,[54,55] was apparent by XPac (Table S17) and Crystal Packing Similarity analysis; see Figure 7. Obvious relationships are seen in forms I and II and

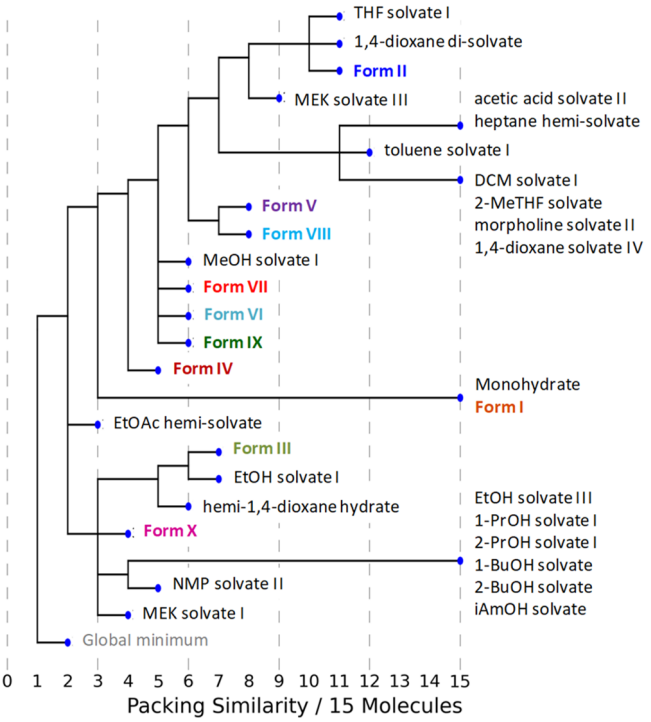


Figure 7. Packing similarity dendrogram showing structural relationships (or lack thereof) between galunisertib solvates, the neat polymorphs, and the GM structure.

their isostructural solvate precursors, each having at least 10 overlapping molecules within 15 molecule clusters. Forms III and V, which are exclusively derived from ethanol solvate I and methanol solvate I, respectively, have fewer molecules in common in their coordination shells as quantified by the $RMSD_{15}$ criterion. However, visual inspection shows aspects of their crystal packing that appear likely to facilitate nondestructive pathways from these parent solvates to the respective neat forms (SI Figure S32).

Assessing the Risk of Late-Appearing Polymorphs. Finding 10 neat polymorphs of a highly solvating molecule is remarkable, even more so considering the constraints imposed on the experimental solid form screen by the compound's chemical instability and poor solubility in some solvents (SI Table S1). Unlike many prolific solvate formers,[56] **GAL** is capable of packing well in three dimensions with itself (Figure 4), and by working around competitive solvate formation, we were able to crystallize forms IV, VI, and occasionally IX, a disappearing polymorph,[7] from solution. However, it was difficult to find conditions where **GAL** crystallized on its own, i.e., without interference from solvates. The possibility of neat polymorphs, particularly more stable forms, having been missed was worrisome. Our modest exploration of crystallization from the melt (requiring the specialist use of pressure) having yielded form X, raised the possibility of finding more polymorphs upon expanding the experimental search conditions.

We calculated the crystal energy landscape of **GAL** to assess the completeness of our experimental search for neat polymorphs. Many energetically competitive structures were predicted as lattice energy minima (Figure 4), although only some are likely to be free energy minima at crystallization process relevant temperatures.[57] Unfortunately, CSP calculations cannot anticipate which "putative" structures are likely

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

to crystallize as polymorphs or, for that matter, the competition a neat form will face from energetically competitive, potentially more stable, solvates. This means that out of all of the unobserved crystal packings, attention will inevitably be drawn to the global minimum structure and whether it has been found in the solid form screen or poses a risk of being a late-appearing polymorph. For **GAL**, the high density GM structure predicted by two different CSP methods, CSP1 requiring electronic structure calculations on all its molecular conformations and CSP2 using periodic DFT+D calculations on all competitive structures, was *not* found.

Although the basic assumption of CSP methods has been that the global minimum energy structure should crystallize, Figure 6 suggests that we cannot be certain that it is the most stable static structure, let alone the most thermodynamically stable at crystallization conditions. The PBE functional suffers from delocalization error[58] overstabilizing charge separation,[59] which can affect conformational energies.[60,61] PBE is also unable to reproduce the experimental stability order for polymorphic systems of small molecules.[62,63] As Figure 6 shows, the relative energies are sensitive to the dispersion correction.[64] Since thermodynamic stability has been approximated by the lattice energy rather than the free energy at crystallization temperatures, it is plausible that the GM structure will be relatively destabilized by thermal motion, as denser rigid aromatic hydrocarbon crystal structures tend to be destabilized by the lattice modes.[65] The quality of the computational results for the observed polymorphs suggests, nonetheless, that the GM structure must be thermodynamically competitive with the observed forms.

The prediction of a high density structure as the global minimum justified widening the experimental search to include crystallization under pressure, although it should be recognized that specific high density forms cannot be targeted under such conditions and the results of a high pressure crystallization may depend on the pressure involved, as well as the rate of compression.[66,67] Unfortunately, our attempts to crystallize **GAL** under pressure, although successful in producing form X, failed to produce the GM structure. This prompted a closer look into possible reasons why the global minimum structure might be kinetically inaccessible. Our "health check"[68] of the GM crystal structure compared and contrasted it to structures that have crystallized, not only the 10 **GAL** polymorphs, but also the nearly one million structures in the Cambridge Structural Database, to identify factors or anomalies which might hinder the first nucleation of this form.

GAL adopts a wide range of conformations in the 10 neat polymorphs, all of which are distorted by crystal packing forces to lower the lattice (free) energy. The molecular strain energy[69] is higher for **GAL** in the GM structure than in any of the known polymorphs but only marginally so relative to one of the independent molecules in form VII (Figure 4, SI Table S20). Given that form VII has only ever been trapped by desolvation, it is possible that the highly strained conformation in the GM structure will be difficult to access in solution or the melt. However, the strain energy is not so high as to rule out the GM structure from alternate approaches, e.g., desolvation. Desolvation would, of course, require both a structurally related solvate and a kinetically favorable transformation pathway, neither of which is assured. Similarly, whether there is a related molecule that crystallizes in a sufficiently similar structure to the GM to be used as a template, either intentionally or accidentally (as an impurity), is far from guaranteed.[70] How can we be confident that the GM could never appear?

In addition to the conformational bias against the GM structure, the hydrogen bonding in this structure is far from optimal. Pyridine N4, which is rarely used in the **GAL** polymorphs but has a high propensity to accept hydrogen bonds, is used in the GM structure. However, the geometry of the hydrogen-bonding interaction to pyridine N4 is highly distorted (Figure 8). The absence of a hot spot near N4 in the full interaction map of **GAL** suggests that the pyridine acceptor is sterically hindered in the GM crystal conformer and therefore not readily accessible to hydrogen-bond donors. This may hinder the formation of the GM as molecules are likely to get trapped in a conformation allowing stronger, directional hydrogen bonds (cf. 3-chloromandelic acid[71]). Collectively, the liabilities of the GM structure revealed in the health check help to explain why it has yet to crystallize and, more importantly, why it may be so kinetically hindered that its appearance would not pose a significant threat to the development of **GAL**.

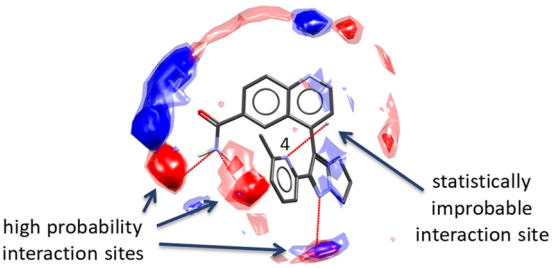


Figure 8. Full interaction map of the global energy minimum structure (GM) of galunisertib showing a statistically improbable hydrogen-bonding interaction to N4. Hydrogen-bond donor and acceptor hotspots were surveyed using carbonyl O acceptor (red) and uncharged NH donor (blue) probes. Hydrogen bonds are shown as red lines.

CONCLUSIONS

Galunisertib now holds the distinction of being the system with the most structurally characterized polymorphs. Observing 10 neat forms, and moreover determining their crystal structures, is particularly remarkable considering that **GAL** is also one of the most prolific solvate forming molecules on record. However, crystal structure prediction suggests there are more polymorphs to be found, a reminder that we still have much to learn about the polymorphism and crystallization of conformationally flexible molecules. With 10 fully characterized crystal structures, experimentally determined energy relationships, and the real possibility of finding more forms, **GAL** is a perfect model system for developing computational methods for relative thermodynamic stability as well as exploring novel experimental approaches for polymorph discovery and structural characterization.

MATERIALS AND METHODS

Materials. **GAL** monohydrate (purity >99%) was obtained from Lilly Research Laboratories. All solvents, purchased from different suppliers and used for crystallization screening, were reagent grade.

Solid form screening encompassed a range of industry standard techniques, including slurry equilibration, solvent evaporation, cooling, standard and inverse antisolvent addition, and vapor diffusion. Unless specified otherwise, galunisertib monohydrate was

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

used as the starting material. The monohydrate and all newly discovered neat and solvated forms were exposed to a variety of temperature and humidity conditions in search of additional forms produced by solid state conversion, desolvation, and solvent exchange. Details of the crystallization conditions surveyed in the solid form screen and the preparation of the neat galunisertib polymorphs, numbered in the order observed, are reported in the SI.

Characterization. *Single-Crystal X-ray Diffraction (SCXRD).* X-ray diffraction data were collected on a Bruker three-circle diffractometer, equipped with a graphite monochromator, Bruker SMART-6000 CCD, or Photon-1000 CMOS detector and using either Cu Kα (λ = 1.54178 Å) radiation from a microfocus sealed tube or Mo Kα (λ = 0.71073 Å) radiation from an IμS microsource. All data were collected at 100 K using an Oxford Cryosystems Cryostream. Single-crystal X-ray diffraction data were also collected on a Rigaku Oxford Diffraction SuperNova diffractometer with Cu Kα radiation.

High-pressure diffraction data were collected using synchrotron radiation of wavelength λ = 0.4859 Å at room temperature on a Newport IS4CCD (four circle) diffractometer with a Pilatus 300 K detector at beamline I19 at the Diamond Light Source, Harwell Science and Innovation Campus.[72]

Details of the crystal structure solution and refinement are provided in SI Section 2.1.

Powder X-ray Diffraction (PXRD). X-ray diffraction patterns were collected at room temperature for lightly ground samples loaded into 0.7 or 1 mm borosilicate capillaries. A PANalytical Empyrean diffractometer, equipped with a θ/θ coupled goniometer in transmission geometry, Cu K$\alpha_{1,2}$ radiation source with a focusing mirror, 0.5° divergence slit, 20 mm mask, 0.25° antiscatter slit and 0.02° soller slit on the incident beam side, a 7.5 mm antiscattering slit and 0.02° soller slit on the diffracted beam side, and a solid state PIXcel 3D detector, was used. A repeated scan scheme[73] was used to collect data for samples in rotating capillaries, and the diffractometer was operated at a tube voltage of 45 kV and tube current of 40 mA using a 2θ step size of 0.013° with 350/800 s per step with 5/15 repeats in the range of 2−70° 2θ. Data analysis was carried out using HighScore Plus 4.0.

The powder diffraction patterns collected for forms I, V, and IX were used to solve their crystal structures. Each pattern was indexed using the 20 lowest angle peaks with the DICVOL04 program as implemented in DASH[74] (v3.4.0), and space groups were determined based on volume considerations, a statistical assessment of the systematic absences[75] as implemented in the DASH structure solution package, and a check using the ADDSYM function in PLATON.[76] Forms VII and VIII candidate structures were identified among the CSP structures. Pawley-type[77] and Rietveld[78] refinement were performed with Topas Academic V5.[79] The background was modeled with Chebyshev polynomials, and the modified Thompson−Cox−Hastings pseudo-Voigt function was used for peak shape fitting. Details of the powder structure solutions are given in SI Section 2.2.

Thermal Analysis. Differential scanning calorimetry (DSC) was conducted using a TA Instruments Modulated DSC Q1000. Samples were equilibrated at 25 °C in inverted hermetically sealed aluminum pans and then heated to 300 °C at 50 °C min^{-1} with a 50 mL min^{-1} N_2 gas purge. The fast heating rate was required to minimize decomposition during the analysis. The temperature and heat flow were calibrated against indium melting.

Calculation and Analysis of Crystal Energy Landscapes. *CSP1.* CrystalPredictor[80,81] (version 1.6) was used to generate approximately two million Z′ = 1 structures of **GAL** in three conformational regions, two of which (A and B) had been observed in the single-crystal structures of solvates and the third (C) with an intramolecular N5−H19···N4 bond. CrystalOptimizer[82−84] was used to refine over 4000 of the lowest energy unique structures, allowing the conformation to respond to the packing forces.

CSP2. The GRACE program (version 2.4[30]) was used to expand the search for **GAL** crystal structures to include both Z′ = 1 and Z′ = 2. For Z′ = 1, the 38 most common space groups were used, and for Z′ = 2, 19 space groups were included (SI Section 3.2). A custom-made force field was parametrized to DFT+D reference data, and crystal structures were generated by parallel tempering with that force field. Finally, the structures were reoptimized with periodic DFT+D as implemented in VASP 5.4.1, using the PBE functional, the Neumann−Perrin (NP) dispersion correction, a plane wave basis set (520 eV, $2\pi \times 0.07$ Å^{-1} k-point grid), and default PAW pseudopotentials (PBE+NP).

Observed polymorphs were matched to CSP-generated structures using two methods. When the experimental crystal structure was available, the packing similarity tool as implemented in COMPACK/Mercury/CCDC Python API was used to attempt to overlay a cluster of 15 or 20 molecules and calculate the root-mean-square distance between heavy atoms in the clusters ($RMSD_{15/20}$).[85] When the experimental structure was not known, the simulated PXRD patterns or indexed cell parameters were manually compared for CSP1. For forms VII and VIII, the PXRD pattern similarity was calculated as a cross-correlation,[33] allowing adjustments to unit cell parameters and the positions of molecules, in order to account for thermal expansion.

Periodic electronic structure calculations were performed using the CASTEP[86,87] plane wave code, as implemented in BIOVIA Materials Studio 2016, using the PBE generalized gradient approximation (GGA) exchange-correlation density functional[88] and ultrasoft pseudopotentials[89] along with a semiempirical Tkatchenko and Scheffler (TS) dispersion correction,[90] a k-point spacing of $2\pi \times 0.05$ Å^{-1}, and a 780 eV cutoff energy for the plane wave basis (PBE+TS). Hydrogen atom positions, all atom positions, and the whole cell were sequentially optimized to find the energy minimum closest to the input structure. The energy at the minimum located using the TS dispersion correction was then re-evaluated with the D2 dispersion correction (PBE+D2). Periodic electronic structure optimizations were also performed using the PBE functional with the D3BJ dispersion correction (PBE+D3BJ) using VASP 5.4.1,[91−94] a 520 eV cutoff energy, and $2\pi \times 0.032$ Å^{-1} k-point distance.

Molecular strain energies were calculated by computationally extracting molecules from the crystal structure, performing a constrained optimization and finally allowing them to relax in vacuum. The strain energy was calculated with Gaussian 16[95] as the difference in B3LYP+D3BJ/6-311G** energy between the constrained and relaxed conformers.

Analysis of the hydrogen-bonding motifs[96] and crystal packing similarity,[97] and generation of full interaction maps[37] were performed using Mercury (version 3.9). The crystal packing similarity between 15 molecule **GAL** clusters in the neat polymorphs, solvates (with solvent molecules removed), and the GM structure was analyzed using a CCDC Python API script.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.9b06634.

Crystal structures for the experimental forms (CIF)

Crystal structures for CSP1 (CIF)

Crystal structures for CSP2 (CIF)

Crystallography (powder, single crystal), solid form screen details, crystal structure prediction study details, solubility, solid form characterization (solid-state NMR spectroscopy, GVS, differential scanning calorimetry, solubility), hydrogen-bond propensity analysis, and full interaction maps (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*reutzel-edens_susan_m@lilly.com

ORCID

Rajni M. Bhardwaj: 0000-0002-7402-9995

Jennifer A. McMahon: 0000-0001-9563-6631

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Jonas Nyman: 0000-0001-9011-3962
Louise S. Price: 0000-0002-7633-1987
Sumit Konar: 0000-0003-3156-6536
Iain D. H. Oswald: 0000-0003-4339-9392
Colin R. Pulham: 0000-0002-3689-9594
Sarah L. Price: 0000-0002-1230-7427
Susan M. Reutzel-Edens: 0000-0003-0806-5565

Notes

The authors declare no competing financial interest.

CCDC deposition numbers 1921238−1921252, 1921488−1921492, 1923244−1923252 1935551−1935552. Crystallographic information files are available free of charge from the Cambridge Crystallographic Data Center (CCDC) upon request (via www.ccdc.cam.ac.uk/data_request/cif, by e-mailing the data_request@ccdc.cam.ac.uk, or contacting The CCDC, 12 Union Road, Cambridge CB2 1EZ, UK; fax: + 44 1223 336033).

■ ACKNOWLEDGMENTS

We wish to dedicate this work to honor the life and legacy of Prof. Joel Bernstein (1941−2019). We thank Lilly scientists Lien Koztecki, Rita Kleemann, Mark Beidelschies, Robert Jones, Suk-fai Lau, Ben Diseroad, David Jackson, Tim Smitka, Chad Hadden, and Jeff Tan for experimental and computational support. We also thank Costas Pantelides and Claire Adjiman (Imperial College London) for sharing the CrystalPredictor and CrystalOptimizer programs used in CSP1. Marcus Neumann is acknowledged for the development, installation and support of the GRACE software. We thank Andrew Craven and Huian Li for support and administration of the Brainiac HPC cluster. Anthony Reilly (CCDC) is acknowledged for the packing similarity dendrogram script. Steve Hunter is thanked for early exploration of high pressure crystallization. We thank the Diamond Light Source for access to Beamline I19, Mark Warren for his input on beamline setup, and the Parsons group (University of Edinburgh) for help during beamtime. This work was funded by Eli Lilly and Company through the Lilly Research Awards Program. The computational infrastructure for CSP1 was developed under EPSRC EP/F03573X/1. The Diamond Light Source is acknowledged for funding under MT16139. IDHO would like to thank the EPSRC for funding his Early Career fellowship EP/N015401/1.

■ REFERENCES

(1) Hilfiker, R.; von Raumer, M. *Polymorphism in the Pharmaceutical Industry: Solid Form and Drug Development*; Wiley-VCH, 2019; pp 1−496.
(2) Huang, L.-F.; Tong, W.-Q. Impact of solid state properties on developability assessment of drug candidates. *Adv. Drug Delivery Rev.* **2004**, *56* (3), 321−334.
(3) Lee, A. Y.; Erdemir, D.; Myerson, A. S. Crystal Polymorphism in Chemical Process Development. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2* (1), 259−280.
(4) Variankaval, N.; Cote, A. S.; Doherty, M. F. From Form to Function: Crystallization of Active Pharmaceutical Ingredients. *AIChE J.* **2008**, *54* (7), 1682−1688.
(5) Sun, C. C. Materials Science Tetrahedron—A useful tool for pharmaceutical research and development. *J. Pharm. Sci.* **2009**, *98* (5), 1671−1687.
(6) Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauer, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K. Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development. *Org. Process Res. Dev.* **2000**, *4* (5), 413−417.
(7) Bučar, D.-K.; Lancaster, R. W.; Bernstein, J. Disappearing Polymorphs Revisited. *Angew. Chem., Int. Ed.* **2015**, *54* (24), 6972−6993.
(8) Desikan, S.; Parsons, R. L.; Davis, W. P.; Ward, J. E.; Marshall, W. J.; Toma, P. H. Process development challenges to accommodate a late-appearing stable polymorph: a case study on the polymorphism and crystallization of a fast-track drug development compound. *Org. Process Res. Dev.* **2005**, *9*, 933−942.
(9) Neumann, M. A.; van de Streek, J. How many ritonavir cases are there still out there? *Faraday Discuss.* **2018**, *211*, 441−458.
(10) Chekal, B. P.; Campeta, A. M.; Abramov, Y. A.; Feeder, N.; Glynn, P. P.; McLaughlin, R. W.; Meenan, P. A.; Singer, R. A. The Challenges of Developing an API Crystallization Process for a Complex Polymorphic and Highly Solvating System. Part I. *Org. Process Res. Dev.* **2009**, *13* (6), 1327−1337.
(11) Braun, D. E.; McMahon, J. A.; Koztecki, L. H.; Price, S. L.; Reutzel-Edens, S. M. Contrasting Polymorphism of Related Small Molecule Drugs Correlated and Guided by the Computed Crystal Energy Landscape. *Cryst. Growth Des.* **2014**, *14* (4), 2056−2072.
(12) Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **2015**, *44* (23), 8619−8635.
(13) Campeta, A. M.; Chekal, B. P.; Abramov, Y. A.; Meenan, P. A.; Henson, M. J.; Shi, B.; Singer, R. A.; Horspool, K. R. Development of a targeted polymorph screening approach for a complex polymorphic and highly solvating API. *J. Pharm. Sci.* **2010**, *99* (9), 3874−3886.
(14) Shtukenberg, A. G.; Zhu, Q.; Carter, D. J.; Vogt, L.; Hoja, J.; Schneider, E.; Song, H.; Pokroy, B.; Polishchuk, I.; Tkatchenko, A.; Oganov, A. R.; Rohl, A. L.; Tuckerman, M. E.; Kahr, B. Powder diffraction and crystal structure prediction identify four new coumarin polymorphs. *Chemical Science* **2017**, *8* (7), 4926−4940.
(15) Shtukenberg, A. G.; Hu, C. T.; Zhu, Q.; Schmidt, M. U.; Xu, W.; Tan, M.; Kahr, B. The Third Ambient Aspirin Polymorph. *Cryst. Growth Des.* **2017**, *17* (6), 3562−3566.
(16) Price, S. L.; Braun, D. E.; Reutzel-Edens, S. M. Can computed crystal energy landscapes help understand pharmaceutical solids? *Chem. Commun.* **2016**, *52* (44), 7065−7077.
(17) Day, G. M. Current approaches to predicting molecular organic crystal structures. *Crystallogr. Rev.* **2011**, *17* (1), 3−52.
(18) Nyman, J.; Reutzel-Edens, S. M. Crystal structure prediction is changing from basic science to applied technology. *Faraday Discuss.* **2018**, *211*, 459−476.
(19) 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. Exploring the Experimental and Computed Crystal Energy Landscape of Olanzapine. *Cryst. Growth Des.* **2013**, *13* (4), 1602−1617.
(20) Vasileiadis, M.; Pantelides, C. C.; Adjiman, C. S. Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule. *Chem. Eng. Sci.* **2015**, *121*, 60−76.
(21) Herbertz, S.; Sawyer, J. S.; Stauber, A. J.; Gueorguieva, I.; Driscoll, K. E.; Estrem, S. T.; Cleverly, A. L.; Desaiah, D.; Guba, S. C.; Benhadji, K. A.; Slapak, C. A.; Lahn, M. M. Clinical development of galunisertib (LY2157299 monohydrate), a small molecule inhibitor of transforming growth factor-beta signaling pathway. *Drug Des., Dev. Ther.* **2015**, *9*, 4479−4499.
(22) Newman, A. Specialized Solid Form Screening Techniques. *Org. Process Res. Dev.* **2013**, *17* (3), 457−471.
(23) Neumann, M. A.; van de Streek, J.; Fabbiani, F. P. A.; Hidber, P.; Grassmann, O. Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nat. Commun.* **2015**, *6*, 7793.
(24) Fabbiani, F. P. A.; Buth, G.; Levendis, D. C.; Cruz-Cabeza, A. J. Pharmaceutical hydrates under ambient conditions from high-pressure seeds: a case study of GABA monohydrate. *Chem. Commun.* **2014**, *50* (15), 1817−1819.

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

(25) Oswald, I. D. H.; Chataigner, I.; Elphick, S.; Fabbiani, F. P. A.; Lennie, A. R.; Maddaluno, J.; Marshall, W. G.; Prior, T. J.; Pulham, C. R.; Smith, R. I. Putting pressure on elusive polymorphs and solvates. *CrystEngComm* **2009**, *11* (2), 359−366.
(26) Fabbiani, F. P. A.; Allan, D. R.; Marshall, W. G.; Parsons, S.; Pulham, C. R.; Smith, R. I. High-pressure recrystallisation—a route to new polymorphs and solvates of acetamide and parabanic acid. *J. Cryst. Growth* **2005**, *275* (1), 185−192.
(27) Connor, L. E.; Morrison, C. A.; Oswald, I. D. H.; Pulham, C. R.; Warren, M. R. Carbon dioxide binary crystals via the thermal decomposition of RDX at high pressure. *Chemical Science* **2017**, *8* (7), 4872−4878.
(28) Pantelides, C. C.; Adjiman, C. S.; Kazantsev, A. V., General Computational Algorithms for Ab Initio Crystal Structure Prediction for Organic Molecules. In *Topics in Current Chemistry*; Atahan-Evrenk, S., Aspuru-Guzik, A., Eds.; Springer: Cham, 2014; Vol. *345*.
(29) Price, S. L. Is zeroth order crystal structure prediction (CSP_0) coming to maturity? What should we aim for in an ideal crystal structure prediction code? *Faraday Discuss.* **2018**, *211*, 9−30.
(30) *GRACE (version 2.4)*, Avant-garde Materials Simulation Deutschland GmbH.
(31) Allen, F. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58* (3), 380−388.
(32) Neumann, M. A.; Perrin, M.-A. Energy Ranking of Molecular Crystals Using Density Functional Theory Calculations and an Empirical van der Waals Correction. *J. Phys. Chem. B* **2005**, *109* (32), 15531−15541.
(33) Habermehl, S.; Mörschel, P.; Eisenbrandt, P.; Hammer, S. M.; Schmidt, M. U. Structure determination from powder data without prior indexing, using a similarity measure based on cross-correlation functions. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2014**, *70* (2), 347−359.
(34) Cruz-Cabeza, A. J.; Bernstein, J. Conformational Polymorphism. *Chem. Rev.* **2014**, *114* (4), 2170−2191.
(35) Galek, P. T. A.; Fábián, L.; Motherwell, W. D. S.; Allen, F. H.; Feeder, N. Knowledge-based model of hydrogen-bonding propensity in organic crystals. *Acta Crystallogr., Sect. B: Struct. Sci.* **2007**, *63* (5), 768−782.
(36) Galek, P. T. A.; Allen, F. H.; Fábián, L.; Feeder, N. Knowledge-based H-bond prediction to aid experimental polymorph screening. *CrystEngComm* **2009**, *11* (12), 2634−2639.
(37) Wood, P. A.; Olsson, T. S. G.; Cole, J. C.; Cottrell, S. J.; Feeder, N.; Galek, P. T. A.; Groom, C. R.; Pidcock, E. Evaluation of molecular crystal structures using Full Interaction Maps. *CrystEngComm* **2013**, *15* (1), 65−72.
(38) López-Mejías, V.; Kampf, J. W.; Matzger, A. J. Nonamorphism in Flufenamic Acid and a New Record for a Polymorphic Compound with Solved Structures. *J. Am. Chem. Soc.* **2012**, *134* (24), 9872−9875.
(39) Zeidan, T. A.; Trotta, J. T.; Tilak, P. A.; Oliveira, M. A.; Chiarella, R. A.; Foxman, B. M.; Almarsson, Ö.; Hickey, M. B. An unprecedented case of dodecamorphism: the twelfth polymorph of aripiprazole formed by seeding with its active metabolite. *CrystEngComm* **2016**, *18* (9), 1486−1488.
(40) Gushurst, K. S.; Nyman, J.; Boerrigter, S. X. M. The PO13 crystal structure of ROY. *CrystEngComm* **2019**, *21*, 1363−1368.
(41) Yu, L. Polymorphism in Molecular Solids: An Extraordinary System of Red, Orange, and Yellow Crystals. *Acc. Chem. Res.* **2010**, *43* (9), 1257−1266.
(42) Haleblian, J.; McCrone, W. Pharmaceutical applications of polymorphism. *J. Pharm. Sci.* **1969**, *58* (8), 911−929.
(43) Stahly, G. P. Diversity in Single- and Multiple-Component Crystals. The Search for and Prevalence of Polymorphs and Cocrystals. *Cryst. Growth Des.* **2007**, *7* (6), 1007−1026.
(44) Tan, M.; Shtukenberg, A. G.; Zhu, S.; Xu, W.; Dooryhee, E.; Nichols, S. M.; Ward, M. D.; Kahr, B.; Zhu, Q. ROY revisited, again: the eighth solved structure. *Faraday Discuss.* **2018**, *211*, 477−491.
(45) Bernstein, J.; Davey, R. J.; Henck, J. O. Concomitant Polymorphs. *Angew. Chem., Int. Ed.* **1999**, *38* (23), 3440−3461.
(46) Case, D. H.; Srirambhatla, V. K.; Guo, R.; Watson, R. E.; Price, L. S.; Polyzois, H.; Cockcroft, J. K.; Florence, A. J.; Tocher, D. A.; Price, S. L. Successful Computationally Directed Templating of Metastable Pharmaceutical Polymorphs. *Cryst. Growth Des.* **2018**, *18* (9), 5322−5331.
(47) Srirambhatla, V. K.; Guo, R.; Price, S. L.; Florence, A. J. Isomorphous template induced crystallisation: a robust method for the targeted crystallisation of computationally predicted metastable polymorphs. *Chem. Commun.* **2016**, *52* (46), 7384−7386.
(48) Pfund, L. Y.; Price, C. P.; Frick, J. J.; Matzger, A. J. Controlling Pharmaceutical Crystallization with Designed Polymeric Heteronuclei. *J. Am. Chem. Soc.* **2015**, *137* (2), 871−875.
(49) Lucaioli, P.; Nauha, E.; Gimondi, I.; Price, L. S.; Guo, R.; Iuzzolino, L.; Singh, I.; Salvalaglio, M.; Price, S. L.; Blagden, N. Serendipitous isolation of a disappearing conformational polymorph of succinic acid challenges computational polymorph prediction. *CrystEngComm* **2018**, *20* (28), 3971−3977.
(50) Shtukenberg, A. G.; Lee, S. S.; Kahr, B.; Ward, M. D. Manipulating Crystallization with Molecular Additives. *Annu. Rev. Chem. Biomol. Eng.* **2014**, *5* (1), 77−96.
(51) Ward, M. D. Perils of Polymorphism: Size Matters. *Isr. J. Chem.* **2017**, *57* (1−2), 82−92.
(52) Hilden, J. L.; Reyes, C. E.; Kelm, M. J.; Tan, J. S.; Stowell, J. G.; Morris, K. R. Capillary Precipitation of a Highly Polymorphic Organic Compound. *Cryst. Growth Des.* **2003**, *3* (6), 921−926.
(53) Stephenson, G. A.; Groleau, E. G.; Kleemann, R. L.; Xu, W.; Rigsbee, D. R. Formation of isomorphic desolvates: Creating a molecular vacuum. *J. Pharm. Sci.* **1998**, *87* (5), 536−542.
(54) Fours, B.; Cartigny, Y.; Petit, S.; Coquerel, G. Formation of new polymorphs without any nucleation step. Desolvation of the rimonabant monohydrate: directional crystallisation concomitant to smooth dehydration. *Faraday Discuss.* **2015**, *179*, 475−488.
(55) Bond, A. D.; Cornett, C.; Larsen, F. H.; Qu, H.; Raijada, D.; Rantanen, J. Structural basis for the transformation pathways of the sodium naproxen anhydrate-hydrate system. *IUCrJ* **2014**, *1* (5), 328−337.
(56) Bērziņš, A.; Zvaniņa, D.; Trimdale, A. Detailed Analysis of Packing Efficiency Allows Rationalization of Solvate Formation Propensity for Selected Structurally Similar Organic Molecules. *Cryst. Growth Des.* **2018**, *18* (4), 2040−2045.
(57) Price, S. L. Why don't we find more polymorphs? *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2013**, *69* (4), 313−328.
(58) Nyman, J.; Yu, L.; Reutzel-Edens, S. M. Accuracy and reproducibility in crystal structure prediction: the curious case of ROY. *CrystEngComm* **2019**, *21*, 2080−2088.
(59) Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **2014**, *140* (18), 18A301.
(60) LeBlanc, L. M.; Dale, S. G.; Taylor, C. R.; Becke, A. D.; Day, G. M.; Johnson, E. R. Pervasive Delocalisation Error Causes Spurious Proton Transfer in Organic Acid−Base Co-Crystals. *Angew. Chem., Int. Ed.* **2018**, *57* (45), 14906−14910.
(61) Uzoh, O. G.; Galek, P. T. A.; Price, S. L. Analysis of the conformational profiles of fenamates shows route towards novel, higher accuracy, force-fields for pharmaceuticals. *Phys. Chem. Chem. Phys.* **2015**, *17* (12), 7936−7948.
(62) Beran, G. J. O. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.* **2016**, *116* (9), 5567−5613.
(63) Wen, S.; Beran, G. J. O. Crystal Polymorphism in Oxalyl Dihydrazide: Is Empirical DFT-D Accurate Enough? *J. Chem. Theory Comput.* **2012**, *8* (8), 2698−2705.
(64) Hoja, J.; Tkatchenko, A. First-principles stability ranking of molecular crystal polymorphs with the DFT+MBD approach. *Faraday Discuss.* **2018**, *211*, 253−274.
(65) Dunitz, J. D.; Filippini, G.; Gavezzotti, A. Molecular Shape and Crystal Packing: a Study of $C_{12}H_{12}$ Isomers, Real and Imaginary. *Helv. Chim. Acta* **2000**, *83* (9), 2317−2335.
(66) Boldyreva, E. High-Pressure Polymorphs of Molecular Solids: When Are They Formed, and When Are They Not? Some Examples

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

of the Role of Kinetic Control. *Cryst. Growth Des.* **2007**, *7* (9), 1662−1668.
(67) Ridout, J.; Price, L. S.; Howard, J. A. K.; Probert, M. R. Polymorphism Arising from Differing Rates of Compression of Liquids. *Cryst. Growth Des.* **2014**, *14* (7), 3384−3391.
(68) Feeder, N.; Pidcock, E.; Reilly, A. M.; Sadiq, G.; Doherty, C. L.; Back, K. R.; Meenan, P.; Docherty, R. The integration of solid-form informatics into solid-form selection. *J. Pharm. Pharmacol.* **2015**, *67* (6), 857−868.
(69) Thompson, H. P. G.; Day, G. M. Which conformations make stable crystal structures? Mapping crystalline molecular geometries to the conformational energy landscape. *Chemical Science* **2014**, *5* (8), 3173−3182.
(70) Braun, D. E.; Oberacher, H.; Arnhard, K.; Orlova, M.; Griesser, U. J. 4-Aminoquinaldine monohydrate polymorphism: prediction and impurity aided discovery of a difficult to access stable form. *CrystEngComm* **2016**, *18* (22), 4053−4067.
(71) Hylton, R. K.; Tizzard, G. J.; Threlfall, T. L.; Ellis, A. L.; Coles, S. J.; Seaton, C. C.; Schulze, E.; Lorenz, H.; Seidel-Morgenstern, A.; Stein, M.; Price, S. L. Are the Crystal Structures of Enantiopure and Racemic Mandelic Acids Determined by Kinetics or Thermodynamics? *J. Am. Chem. Soc.* **2015**, *137* (34), 11095−11104.
(72) Nowell, H.; Barnett, S. A.; Christensen, K. E.; Teat, S. J.; Allan, D. R. I19, the small-molecule single-crystal diffraction beamline at Diamond Light Source. *J. Synchrotron Radiat.* **2012**, *19* (3), 435−441.
(73) Hill, R. A.; Madsen, I. C., Sample preparation, instrument selection and data collection. In *Structure Determination from Powder Diffraction Data*; David, W. I. F., Ed.; Oxford University Press, 2002; pp 98−117.
(74) David, W. I. F.; Shankland, K.; van de Streek, J.; Pidcock, E.; Motherwell, W. D. S.; Cole, J. C. DASH: a program for crystal structure determination from powder diffraction data. *J. Appl. Crystallogr.* **2006**, *39*, 910−915.
(75) Markvardsen, A. J.; David, W. I. F.; Johnson, J. C.; Shankland, K. A probabilistic approach to space-group determination from powder diffraction data. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2001**, *57*, 47−54.
(76) Spek, A. L. Structure validation in chemical crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65* (2), 148−155.
(77) Pawley, G. S. Unit-Cell Refinement from Powder Diffraction Scans. *J. Appl. Crystallogr.* **1981**, *14* (DEC), 357−361.
(78) Rietveld, H. A profile refinement method for nuclear and magnetic structures. *J. Appl. Crystallogr.* **1969**, *2* (2), 65−71.
(79) Coelho, A. A. *Topas Academic V5*; Coelho Software: Brisbane, 2012.
(80) Karamertzanis, P. G.; Pantelides, C. C. Ab initio crystal structure prediction. II. Flexible molecules. *Mol. Phys.* **2007**, *105* (2−3), 273−291.
(81) Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L.; Galek, P. T. A.; Day, G. M.; Cruz-Cabeza, A. J. Successful prediction of a model pharmaceutical in the fifth blind test of crystal structure prediction. *Int. J. Pharm.* **2011**, *418* (2), 168−178.
(82) Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. Efficient Handling of Molecular Flexibility in Lattice Energy Minimization of Organic Crystals. *J. Chem. Theory Comput.* **2011**, *7* (6), 1998−2016.
(83) Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010**, *12* (30), 8478−8490.
(84) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1* (6), 1128−1132.
(85) Chisholm, J. A.; Motherwell, S. COMPACK: a program for identifying crystal structure similarity using distances. *J. Appl. Crystallogr.* **2005**, *38* (1), 228−231.
(86) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C. First principles methods using CASTEP. *Z. Kristallogr. - Cryst. Mater.* **2005**, *220*, 567.
(87) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C. First principles methods using CASTEP. *Z. Kristallogr. - Cryst. Mater.* **2005**, *220*, 567−570.
(88) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865−3868.
(89) Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1990**, *41* (11), 7892−7895.
(90) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102* (7), No. 073005.
(91) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47* (1), 558−561.
(92) Kresse, G.; Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *49* (20), 14251−14269.
(93) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6* (1), 15−50.
(94) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54* (16), 11169−11186.
(95) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, J. A.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, R.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J., *Gaussian 16, Revision B.01*, Gaussian, Inc.: Wallingford CT, 2016.
(96) Chisholm, J.; Pidcock, E.; van de Streek, J.; Infantes, L.; Motherwell, S.; Allen, F. H. Knowledge-based approaches to crystal design. *CrystEngComm* **2006**, *8* (1), 11−28.
(97) 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. Mercury CSD 2.0− new features for the visualization and investigation of crystal structures. *J. Appl. Crystallogr.* **2008**, *41* (2), 466−470.

DOI: 10.1021/jacs.9b06634
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX