

Cite this: *Chem. Sci.*, 2022, 13, 1288

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

How many more polymorphs of ROY remain undiscovered†

Gregory J. O. Beran,^a Isaac J. Sugden,^b Chandler Greenwell,^a David H. Bowskill,^b Constantinos C. Pantelides^b and Claire S. Adjiman^b

With 12 crystal forms, 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile (a.k.a. ROY) holds the current record for the largest number of fully characterized organic crystal polymorphs. Four of these polymorph structures have been reported since 2019, raising the question of how many more ROY polymorphs await future discovery. Employing crystal structure prediction and accurate energy rankings derived from conformational energy-corrected density functional theory, this study presents the first crystal energy landscape for ROY that agrees well with experiment. The lattice energies suggest that the seven most stable ROY polymorphs (and nine of the twelve lowest-energy forms) on the $Z' = 1$ landscape have already been discovered experimentally. Discovering any new polymorphs at ambient pressure will likely require specialized crystallization techniques capable of trapping metastable forms. At pressures above 10 GPa, however, a new crystal form is predicted to become enthalpically more stable than all known polymorphs, suggesting that further high-pressure experiments on ROY may be warranted. This work highlights the value of high-accuracy crystal structure prediction for solid-form screening and demonstrates how pragmatic conformational energy corrections can overcome the limitations of conventional density functionals for conformational polymorphs.

Received 2nd November 2021
Accepted 10th December 2021

DOI: 10.1039/d1sc06074k

rsc.li/chemical-science

1 Introduction

Estimates indicate that about half of organic molecules exhibit polymorphism in the solid state,¹ and these changes in crystal packing can profoundly alter physical properties such as color, stability, solubility, and carrier mobility. While the discovery of two or three polymorphs for a given species is common, some molecules are prolific polymorph formers that can adopt many more crystal forms. For example, ten fully characterized polymorphs have been reported for anti-cancer drug candidate galunisertib,² nine for flufenamic acid,³ aripiprazole,⁴ and tolafenamic acid,⁵ six for the energetic material triacetone-triperoxide (TATP),⁶ and at least five polymorphs for many other species.^{7–14} Large numbers of crystalline phases are also found for small molecules such as water,¹⁵ nitrogen,¹⁶ and carbon dioxide.^{17,18}

The molecule 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile in Fig. 1, which is often called ROY due

to its vividly-colored red, orange, and yellow crystals, is currently the most prolific organic polymorph former known. Twelve fully characterized ROY crystal structures have been discovered in the past twenty-five years: Y,^{19,20} ON,^{19,20} R,^{19,20} OP,²⁰ YN,²⁰ ORP,²⁰ YT04,²¹ Y04,^{21,22} R05,^{23,24} PO13,^{25,26} R18,²⁷ and Y19.²⁶ The structure for a thirteenth RPL polymorph²⁸ has been proposed,²⁹ but it has not yet been definitively characterized. Structures for four of these forms (PO13, R18, Y19, and Y04) have been reported only since 2019. Whereas Y04 and PO13 were previously known but incompletely characterized, R18 and Y19 are entirely new polymorphs. This rapid increase in the number of solved ROY polymorphs seems to support Walter McCrone's oft-cited conjecture³⁰ that “in general, the number of forms known for a given compound is proportional to the time and money spent

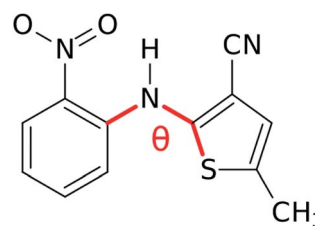


Fig. 1 The ROY molecule. Its different conformational polymorphs and their associated colors generally involve changes in the S–C–N–C dihedral angle θ .

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

^bDepartment of Chemical Engineering, Sargent Centre for Process Systems Engineering, Imperial College London, London, SW7 2AZ, UK

† Electronic supplementary information (ESI) available: Additional computational details, optimized crystal structures, ROY conformational energy scan conformational energies, polymorph stabilities, new analysis of a previously reported ROY structure prediction landscape, and the list of structures considered at higher pressures. See DOI: 10.1039/d1sc06074k

structure prediction and applies our conformational energy correction approach to investigate the landscape containing all experimentally-known polymorphs of ROY and hundreds of additional candidate structures. This effort results in the first ROY crystal energy landscape that is highly consistent with experiment. It suggests that researchers have already discovered most if not all of the highly stable ROY polymorphs. Realizing any new polymorphs at ambient conditions will likely require selective crystallization strategies capable of kinetically trapping more metastable structures. On the other hand, one currently unknown candidate structure is predicted to become thermodynamically more stable than all existing ROY polymorphs at around 10 GPa of pressure, suggesting that high pressure experiments might discover a new polymorph.

2 Theoretical approach

Candidate crystal structures of ROY were generated *via* crystal structure prediction, as described briefly below. See ESI† for further details. The initial stages of the CSP protocol follow the general strategy from Pantelides *et al.*⁷⁶ and were employed in a 2012 study of ROY.⁴⁰

First, relevant ROY conformations and local approximate models⁷⁷ describing the intramolecular conformational energetics for two key conformational degrees of freedom were generated *via* gas-phase studies at the B3LYP/6-31G(d,p) level of theory using Gaussian 09.⁷⁸ See ESI† for details. Second, a global CSP search for low-energy minima on the $Z' = 1$ lattice energy surface was performed using CrystalPredictor II,⁷⁹ using the smooth intramolecular potential algorithm⁸⁰ for intramolecular interactions and parameters from the FIT potential^{81–84} set for intermolecular interactions. Two million minimizations in 61 space groups were carried out. After removal of duplicates, 2869 distinct crystal structures were identified within 20 kJ mol⁻¹ of the global minimum. Third, the lowest 1000 of these structures were locally minimized with additional conformational flexibility in CrystalOptimizer,⁸⁵ using B3LYP/6-31G(d,p) to describe the intramolecular energetics, multipolar electrostatics (up to hexadecapoles),⁸⁶ and repulsion/dispersion parameters from the FIT potential set.

Fourth, all structures within 10 kJ mol⁻¹ of the global minimum energy structure were fully relaxed with dispersion-corrected periodic planewave DFT calculations using the B86bPBE density functional^{87,88} and the exchange-dipole moment (XDM) dispersion correction⁸⁹ in Quantum Espresso.^{90,91} To identify any additional higher-lying structures from the CrystalOptimizer landscape that might also be important, the DFT-optimized crystal geometries and energies from the 50-lowest-energy CrystalOptimizer structures were used to refit the repulsion/dispersion parameters with CrystalEstimator.⁹² Such molecule-specific tailoring has been shown to lead to improved outcomes in CSP.^{48,93–99} After refining all 1000 structures from stage 3 with the modified potential, around a dozen additional structures fell within the 10 kJ mol⁻¹ window and were also relaxed with DFT. Because the search only examined $Z' = 1$, the three experimentally-reported $Z' = 2$ polymorphs (R05, R18, and RPL) were added to the set

manually. For the RPL polymorph, the proposed structure from ref. 29 was used. In total, approximately 300 crystal structures were relaxed with B86bPBE-XDM. After removal of duplicates, the landscape includes 264 crystal structures.

Finally, because delocalization error in generalized gradient approximation (GGA) and hybrid functionals artificially stabilizes more planar conformations of ROY,^{29,70,71} a single-point gas-phase monomer conformational energy correction was applied to each crystal.⁷¹ These corrections replace the B86bPBE-XDM intramolecular conformational energy with one computed from spin-component-scaled dispersion-corrected second-order Møller–Plesset perturbation theory (SCS-MP2D) at the complete-basis-set limit.¹⁰⁰ This model reproduces benchmark coupled cluster theory conformational energies of ROY well (Fig. S3†) at modest computational cost. The conformationally-corrected energy per unit cell is computed as:

$$E_{\text{crystal}}^{\text{corrected}} = E_{\text{crystal}}^{\text{DFT}} + \sum_i \left(E_{\text{mon},i}^{\text{SCS-MP2D}} - E_{\text{mon},i}^{\text{DFT}} \right) \quad (1)$$

where $E_{\text{mon},i}$ is the gas-phase energy of the i -th monomer in the unit cell (in its crystalline geometry) and i runs over all molecules in the cell. Space group symmetry can be exploited to compute corrections only for the symmetrically unique monomers. For ROY, the single-point monomer correction performed here requires 1–2 orders of magnitude less computational effort than the periodic DFT geometry optimization, depending on the unit cell size. SCS-MP2D calculations were performed using PSI4 and a custom implementation¹⁰¹ of the spin-component scaling and dispersion correction.

3 Results and discussion

3.1 The ROY crystal energy landscape

Fig. 2a presents the initial CrystalOptimizer energy landscape from stage 3 of the CSP protocol. This $Z' = 1$ landscape is densely populated, containing 284 structures within 10 kJ mol⁻¹ of the R polymorph which it predicts to be the global minimum energy structure. Similar to an earlier version of this landscape,⁴⁰ it includes all ten experimentally-known $Z' = 1$ polymorphs. However, the energy model incorrectly predicts the R polymorph to be 1.8 kJ mol⁻¹ more stable than form Y. Moreover, while many of the experimental polymorphs appear among the 30 lowest structures, forms ON (rank 115) and Y19 (rank 172) lie considerably higher. As shown in Fig. 3, the relative CrystalOptimizer lattice energies correlate moderately well with the relative experimental enthalpies that Yu reported for the Y, YT04, R, OP, ON, YN, and ORP polymorphs,¹⁰² with a root-mean-square error (RMSE) of 2.5 kJ mol⁻¹. The largest errors of 3–4 kJ mol⁻¹ occur for the R and ORP polymorphs. While perfect agreement between relative 0 K lattice energies and relative finite-temperature enthalpies cannot be expected due to differences in the enthalpy phonon contributions between forms that are absent in the lattice energies, this data suggests that the CrystalOptimizer energy rankings probably exhibit appreciable errors relative to the magnitude of the energy differences between ROY polymorphs.



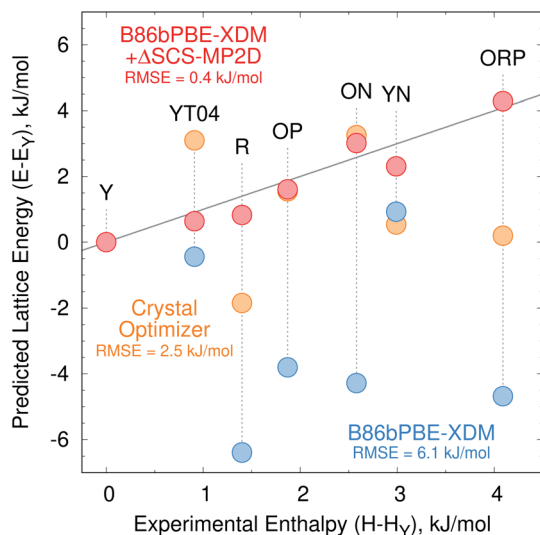


Fig. 3 Relative lattice energies for seven ROY polymorphs predicted with the CrystalOptimizer potential (orange), B86bPBE-XDM (blue) and B86bPBE-XDM + the SCS-MP2D monomer-correction (red) compared against the experimentally measured relative enthalpies (kJ mol^{-1}). Points lying along the solid gray line would represent perfect agreement between the relative theoretical lattice energies and experimental enthalpies.

Next, all structures indicated in gray or red in Fig. 2a were fully relaxed with periodic density functional theory using the dispersion-corrected B86bPBE-XDM density functional. This set includes all 284 structures lying within 10 kJ mol^{-1} of form R on the CrystalOptimizer landscape plus a handful of higher-energy structures that were stabilized upon refitting of the CrystalOptimizer potential, as described in the Theoretical methods section.

In the end, after relaxing approximately 300 crystal structures with DFT and removing duplicates, the landscape contains 264 crystal structures. DFT refinement broadens the energy range of these structures considerably to 22 kJ mol^{-1} , and Fig. 2b plots the lowest 16 kJ mol^{-1} region of this landscape. The DFT-optimized structures are also noticeably more dense compared to the CrystalOptimizer ones, reflecting the differences between an intermolecular potential fitted to finite-temperature crystal structures and 0 K electronic structure calculations. Nevertheless, the DFT-optimized structures exhibit good agreement with their experimental counterparts, with an average 15-molecule cluster root-mean-square deviation (RMSD15)¹⁰³ value of $0.23 \pm 0.08 \text{ \AA}$ (Table S2†).

Unfortunately, the B86bPBE-XDM energy rankings agree poorly with experiment, similar to what has been found in previous DFT studies of ROY.^{24,29,68,70,71} As shown in Fig. 3, the B86bPBE-XDM lattice energies exhibit an RMS error of 6.1 kJ mol^{-1} compared to the relative experimental enthalpies, more than double the error of the CrystalOptimizer potential. Most notably, the red- and orange-colored experimental polymorphs are substantially overstabilized relative to the yellow ones. B86bPBE-XDM predicts 92 structures to be more stable than the experimentally most stable form Y, and 96% of those

lower-energy structures adopt intramolecular conformations associated with red/orange polymorphs instead of yellow ones. Forms R18, R, and R05 are predicted to be among the most stable structures, and an experimentally-unknown structure is predicted to be the global minimum.

The poor B86bPBE-XDM energy rankings reflect the impact of delocalization error in the density functional, which artificially stabilizes structures with more planar ROY conformations.^{29,68,70,71} In contrast, the CrystalOptimizer energy model derives its intramolecular conformational energies from the hybrid B3LYP functional, which is less impacted by delocalization error than the GGA functional B86bPBE-XDM. Nevertheless, B3LYP and other hybrid functionals do still overstabilize the red and orange conformations, albeit to a lesser extent.⁷¹

To address the delocalization error, the intramolecular DFT conformational energies were replaced with wavefunction-based SCS-MP2D ones that do not suffer from delocalization error. This computationally inexpensive correction (*cf.* eqn (1)) dramatically improves the crystal energy landscape (Fig. 2c). Form Y becomes the most stable polymorph, followed by YT04, R, OP, YN, Y04, R05, PO13, ON, ORP, R18, RPL, and Y19. As shown in Fig. 3, these conformational energy-corrected relative lattice energies correlate excellently with the relative experimental enthalpies, reducing the rms error from 6.1 kJ mol^{-1} to only 0.4 kJ mol^{-1} . This represents a remarkable improvement, despite the aforementioned caveat about comparing electronic lattice energies and experimental enthalpies. Moreover, the predicted polymorph stability ordering for these seven forms matches the experimental one almost perfectly. Only the YN and ON energy ordering is reversed, though the predicted energy difference is only 0.7 kJ mol^{-1} .⁷¹ This excellent agreement between the predicted relative lattice energies and relative experimental enthalpies supports the assumption of the monomer correction approach that the B86bPBE-XDM functional is describing the intermolecular interactions well in this system.

The predicted relative stabilities are also consistent with available experimental evidence beyond that shown in Fig. 3. For example, the model predicts Y04 to be less stable than YT04 and R, which is consistent with the experimentally observed spontaneous conversion of Y04 to those forms.^{21,22} Similarly, forms R05 and YN lie above Y and R, which is consistent with experimentally observed transformations over time.^{20,24}

The ORP, R18, RPL, and Y19 polymorphs lie 1–4 kJ mol^{-1} above the other experimentally known polymorphs in Fig. 2c, but this relative instability also seems consistent with available experimental information. For example, ORP is metastable and converts to Y over time,²⁰ and the relative lattice energy agrees well with the experimental enthalpy difference (Fig. 3). RPL is also highly metastable experimentally, converting to lower-energy form YN in a matter of hours or days, and even faster upon even slight mechanical pressure.²⁸ The stabilities of the newly discovered R18 and Y19 forms have not been reported, but unconventional crystallization techniques were required to produce each one: R18 was discovered during extensive screening *via* encapsulated nanodroplet crystallization.²⁷ Y19



was crystallized using seeds of a mixed crystal consisting of a 40 : 60 mixture of ROY and FuROY, the latter of which replaces the sulfur atom with an oxygen atom.²⁶ In other words, the relative instability predicted for these forms is plausible given the difficulties in crystallizing them experimentally.

3.2 Potential for discovering new polymorphs

Having obtained a CSP landscape that reproduces important features of the experimentally known polymorphs, the potential for discovering new polymorphs in the future can be assessed. The crystal energy landscape contains 174 crystal structures within the commonly studied 10 kJ mol⁻¹ lattice energy window, 161 of which have not been observed experimentally. 59 of those unknown structures lie in the 6.6 kJ mol⁻¹ energy window between the most stable (Y, rank #1) and least stable (Y19, rank #72) experimental forms.

Most significantly, the seven lowest-energy structures on the CSP landscape have already been discovered experimentally: Y, YT04, R, OP, YN, Y04, R05. The rank #10 (PO13) and #12 (ON) structures are also already known. Only 13 currently undiscovered structures lie below the metastable and/or difficult-to-crystallize energy window of forms ORP, R18, RPL, and Y19.

The most stable of those candidates occur at ranks #8, #9, #11, and #13, all lying within about 3 kJ mol⁻¹ of form Y. Structures #8, #9 and #11 would be yellow polymorphs. The first two adopt an intramolecular conformation almost identical to that of YT04, while the third nearly matches the Y polymorph conformation. With a conformation similar to forms ON and OP, structure #13 would probably be orange. The intermolecular packing arrangements in #8, #11, and #13 differ from any of the experimentally-known polymorphs. In contrast, structure #9 adopts the same 1-D wine-rack motif as YT04, but with different packing relationships between the 1-D stacks. Given the common packing motif, crystallization might reasonably lead preferentially to the formation of the thermodynamically more stable YT04 polymorph over structure #9. Several of the other lower-energy candidate structures also exhibit varying degrees of packing motif similarity with known forms that could make them difficult to crystallize instead of the already known polymorphs. Finally, we note that the two predicted structures that were exceptionally stable on the DFT landscape in Fig. 2b are destabilized considerably by applying the conformational correction, increasing to ranks #17 (+3.8 kJ mol⁻¹) and #43 (+5.5 kJ mol⁻¹).

Crystal packing similarities between the candidate structures and the polymorphs of FuROY, the ROY-analog which replaces the sulfur with an oxygen, were also investigated. The only notable partial packing similarity occurs between the orange-red polymorph of FuROY and the rank #194 structure. However, at 10.6 kJ mol⁻¹ less stable than form Y, structure #194 seems unlikely to be realized experimentally. The lack of low-energy structures with strong packing similarities to the FuROY polymorphs is consistent with the failure of FuROY homoseeding to produce any new polymorphs of ROY.²⁶

What about crystal structures with more than one molecule in the asymmetric unit ($Z' > 1$)? While the present crystal

structure prediction searched over only $Z' = 1$ structures, the 2019 study by Nyman *et al.* also included $Z' = 2$ structures.²⁹ That crystal landscape is imperfect: it found only 6 of the then 9 known experimental polymorphs, and its structures are heavily biased towards more planar ROY conformations (*i.e.* red/orange-colored polymorphs) due to delocalization error in the density functional used to produce it. Nevertheless, it can provide insight into the potential for important $Z' = 2$ structures not considered in the current search. As discussed in Section S2.4,† a single-point energy re-ranking with B86bPBE-XDM and the SCS-MP2D conformational energy correction of the 500 lowest-energy structures on that landscape found one experimentally-unknown candidate within 3 kJ mol⁻¹ of form Y (*i.e.* the energy window containing the nine most stable experimental polymorphs in Fig. 2c). However, this structure exhibits significant packing similarities to the more stable YN polymorph, suggesting it might be difficult to crystallize experimentally. The landscape also contains nearly 50 unknown $Z' = 2$ structures within 5 kJ mol⁻¹ of form Y. Similar to what was found for the landscape in Fig. 2c, however, it is unclear whether any of those structures are particularly likely to be realizable experimentally.

So while the conformational energy-corrected crystal energy landscapes contain many unknown crystal structures that might conceivably be crystallized experimentally, it seems likely that doing so would prove difficult for the vast majority of the candidate structures in Fig. 2c. The lowest-energy forms have already been discovered. Most of the experimentally unknown predicted structures lie higher in energy than the already difficult-to-produce metastable forms like RPL, R18, and Y19, while several of the lower-energy candidates exhibit partial packing similarities to more experimental polymorphs. Of course, one should never entirely rule out the potential for experimental ingenuity and/or serendipity to discover some of these hypothetical forms.

3.3 ROY polymorphs at high pressure

Given the dearth of promising new low-energy candidate structures to be crystallized at ambient conditions, we turn our attention to the potential for new high-pressure polymorphs of ROY. Experimental studies have investigated the behavior of the Y,¹⁰⁴ OP,¹⁰⁵ and ON¹⁰⁶ polymorphs at pressures up to 5.2 GPa, 9.3 GPa, and 4.2 GPa, respectively. With the exception of form ON,¹⁰⁶ hydrostatic pressure generally flattens the key SCNC intramolecular dihedral angle between the two aromatic rings. In form Y, this dihedral flattening results in a “collapsible wine rack” compression of the crystal structure and an associated piezochromic color change from yellow to red.¹⁰⁴ However, no pressure-induced polymorphic phase transitions have been yet observed in ROY.

Given the number of high-density forms with relatively low lattice energies, we investigate the pressure-dependent relative enthalpies for the 12 fully characterized experimental forms and 14 additional structures exhibiting packing densities similar to or greater than form Y and energies less than 6 kJ mol⁻¹ above Y at ambient pressure. See Fig. S5† for the



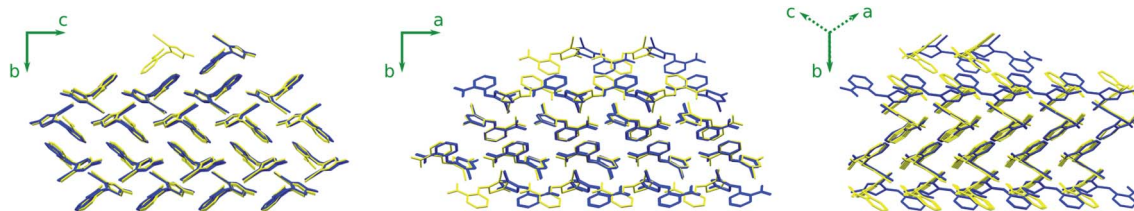


Fig. 5 Comparison of form Y (yellow) and structure #15 (blue) at ambient pressure. Cell axes are indicated for form Y. For the middle and right perspectives in particular, note how the structures share a common packing motif in the central layers but the adjacent layers above and below in the *b* direction differ by mirror image inversion about the *a* axis.

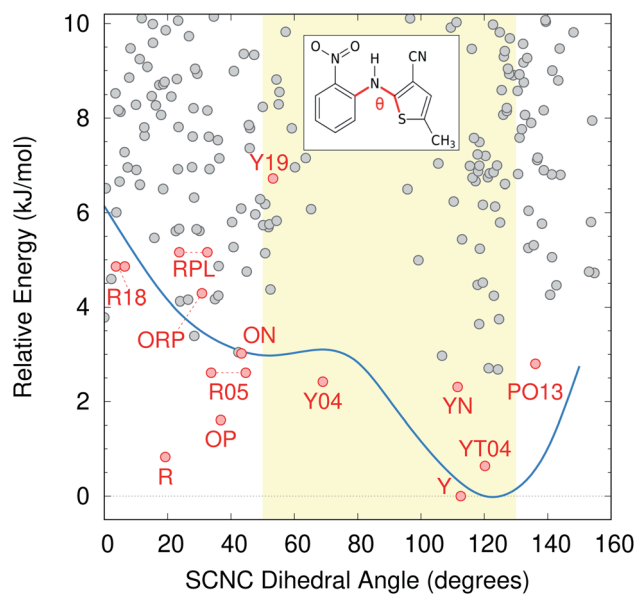


Fig. 6 ROY crystal structure energy landscape as a function of the intramolecular conformation, as defined by the key SCNC dihedral angle. The gas-phase SCS-MP2D conformational energy scan over this dihedral angle is superimposed on the landscape in blue, while the yellow shaded region indicates the approximate dihedral angle range that leads to yellow polymorphs instead of red/orange ones.

4 Conclusions

This study has produced the first crystal structure prediction landscape for ROY that is consistent with available experimental evidence. This was achieved by augmenting the CSP approach that had previously produced the most realistic landscape with conformational energy corrections to DFT lattice energies to overcome the limitations of delocalization error in the approximate density functionals. The relative lattice energies of the known polymorphs agree well with relative experimental enthalpies and other qualitative observations about polymorph stabilities. In the future, one might investigate the accuracy of the DFT intermolecular description further, though the good agreement with experiment provides some confidence that those interactions are described well in the model used here. While the landscape presented here appears accurate, future work could improve it further by examining structures with multiple molecules in the asymmetric unit more

thoroughly and by computing finite-temperature free energies. Both would require considerable further investment of computational resources, however.

The ROY landscapes presented here suggest that the $Z' = 1$ polymorphs and, with a lesser degree of confidence, the $Z' = 2$ polymorphs with the most stable lattice energies have already been discovered. Further polymorph discoveries at ambient pressure would seem to require experimental ingenuity to trap metastable forms, similar to the efforts involved in crystallizing forms R18, Y19, and Y04. On the other hand, there may be potential for discovering a new polymorph at higher pressure. The YT04 and PO13 polymorphs are predicted to become more stable than form Y in the ~ 1 –10 GPa range. Around 10 GPa, however, a new, unknown candidate structure is predicted to become enthalpically more stable than all experimentally-known polymorphs. Given the structural similarities of this candidate to form Y, perhaps this new structure might be realized experimentally by subjecting form Y to pressures around 10 GPa. Alternatively, liquid recrystallizations at high pressures could also be considered.

Finally, the work here demonstrates how conformational energy corrections can be applied to entire crystal energy landscapes to improve their accuracy considerably. Given the small computational cost of these corrections—only a few percent of the time spent on DFT structure optimizations—they should be considered whenever conformational polymorphs exhibit varying degrees of intramolecular π conjugation or any other features that are likely to manifest in significant delocalization error.

Author contributions

The Imperial College team performed the initial stages of the crystal structure prediction. G. J. O. B. performed the subsequent quantum mechanical refinement, with help from C. G., and analyzed the data. G. J. O. B. wrote the paper with input from all other co-authors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

G. J. O. B. gratefully acknowledges funding for this work from the National Science Foundation (CHE-1955554) and



- 76 C. C. Pantelides, C. S. Adjiman and A. V. Kazantsev, *Top. Curr. Chem.*, 2014, **345**, 25–58.
- 77 I. Sugden, C. S. Adjiman and C. C. Pantelides, *Acta Cryst. B*, 2016, **72**, 864–874.
- 78 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision E.01*, Gaussian Inc., Wallingford CT, 2009.
- 79 M. Habgood, I. J. Sugden, A. V. Kazantsev, C. S. Adjiman and C. C. Pantelides, *J. Chem. Theory Comput.*, 2015, **11**, 1957–1969.
- 80 I. J. Sugden, C. S. Adjiman and C. C. Pantelides, *Acta Cryst. B*, 2019, **75**, 423–433.
- 81 S. R. Cox, L.-Y. Hsu and D. E. Williams, *Acta Cryst. A*, 1981, **37**, 293–301.
- 82 D. E. Williams and S. R. Cox, *Acta Cryst. B*, 1984, **40**, 404–417.
- 83 D. S. Coombes, S. L. Price, D. J. Willock and M. Leslie, *J. Phys. Chem.*, 1996, **100**, 7352–7360.
- 84 T. Beyer and S. L. Price, *J. Phys. Chem. B*, 2000, **104**, 2647–2655.
- 85 A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman and C. C. Pantelides, *J. Chem. Theory Comput.*, 2011, **7**, 1998–2016.
- 86 S. L. Price, M. Leslie, G. W. A. Welch, L. S. P. M. Habgood, P. G. Karamertzanis and G. M. Day, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8478–8490.
- 87 A. D. Becke, *J. Chem. Phys.*, 1986, **38**, 7184–7187.
- 88 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
- 89 A. Otero-de-la Roza and E. R. Johnson, *J. Chem. Phys.*, 2012, **136**, 174109.
- 90 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys.: Condens. Matter*, 2009, **21**, 395502.
- 91 P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni, *J. Phys.: Condens. Matter*, 2017, **29**, 465901.
- 92 D. H. Bowskill, I. J. Sugden, C. C. Pantelides and C. S. Adjiman, in preparation, 2021.
- 93 M. A. Neumann, *J. Phys. Chem. B*, 2008, **112**, 9810–9829.
- 94 M. A. Neumann, F. J. J. Leusen and J. Kendrick, *Angew. Chem., Int. Ed.*, 2008, **47**, 2427–2430.
- 95 A. J. Misquitta, G. W. Welch, A. J. Stone and S. L. Price, *Chem. Phys. Lett.*, 2008, **456**, 105–109.
- 96 K. Szalewicz, *Acc. Chem. Res.*, 2014, **47**, 3266–3274.
- 97 M. P. Metz and K. Szalewicz, *J. Chem. Theory Comput.*, 2020, **16**, 2317–2339.
- 98 D. McDonagh, C.-K. Skylaris and G. M. Day, *J. Chem. Theory Comput.*, 2019, **15**, 2743–2758.
- 99 D. H. Bowskill, I. J. Sugden, N. George, A. Keates, J. Webb, C. C. Pantelides and C. S. Adjiman, *Comput.-Aided Chem. Eng.*, 2020, **48**, 493–498.
- 100 C. Greenwell, J. Řezáč and G. J. O. Beran, *ChemRxiv*, 2021, DOI: 10.33774/chemrxiv-2021-jdv3c.
- 101 *The MP2D software can be downloaded at*, <https://github.com/Chandemonium/MP2D/relax>.
- 102 L. Yu, *Acc. Chem. Res.*, 2010, **43**, 1257–1266.
- 103 J. A. Chisholm and W. D. S. Motherwell, *J. Appl. Crystallogr.*, 2005, **38**, 228–231.
- 104 E. L. Harty, A. R. Ha, M. R. Warren, A. L. Thompson, D. R. Allan, A. L. Goodwin and N. P. Funnell, *Chem. Commun.*, 2015, **51**, 10608–10611.
- 105 N. P. Funnell, C. L. Bull, C. J. Ridley and S. Capelli, *CrystEngComm*, 2019, **21**, 4473–4483.
- 106 L. R. Warren, E. McGowan, M. Renton, C. A. Morrison and N. P. Funnell, *Chem. Sci.*, 2021, **12**, 12711–12718.
- 107 B. A. Nogueira, C. Castiglioni and R. Fausto, *Nat. Commun.*, 2020, **3**, 34.
- 108 X. Feng, A. D. Becke and E. R. Johnson, *CrystEngComm*, 2021, **23**, 4264–4271.
- 109 A. J. Cruz-Cabeza and J. Bernstein, *Chem. Rev.*, 2014, **114**, 2170–2191.

