Efficient Molecular Crystal Structure Prediction and Stability Assessment with AIMNet2 Neural Network Potentials

Kamal Singh Nayal,[†] Dana O'Connor,[‡] Roman Zubatyuk,[†] Dylan M. Anstine,[†] Yi Yang,[‡] Rithwik Tom,[¶] Wenda Deng,[‡] Kehan Tang,[‡] Noa Marom,[*,†,‡,¶] and Olexandr Isayev[*,†,‡]

†*Department of Chemistry, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, United States*

‡*Department of Materials Science and Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, United States*

¶*Department of Physics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, United States*

E-mail: nmarom@andrew.cmu.edu; olexandr@olexandrisayev.com

Abstract

Identifying thermodynamically stable crystal structures remains a key challenge in materials chemistry. Computational crystal structure prediction (CSP) workflows typically rank candidate structures by lattice energy to assess relative stability. Approaches using self-consistent first-principles calculations become prohibitively expensive, especially when millions of energy evaluations are required for complex molecular systems with many atoms per unit cell. Here, we provide a detailed analysis of our methodology and results from the seventh blind test of crystal structure prediction organized

1

by the Cambridge Crystallographic Data Centre (CCDC). We present an approach that significantly accelerates CSP by training target-specific machine learned interatomic potentials (MLIPs). AIMNet2 MLIPs are trained on density functional theory (DFT) calculations of molecular clusters, herein referred to as n-mers. We demonstrate that potentials trained on gas phase dispersion-corrected DFT reference data of n-mers successfully extend to crystalline environments, accurately characterizing the CSP landscape and correctly ranking structures by relative stability. Our methodology effectively captures the underlying physics of thermodynamic crystal stability using only molecular cluster data, avoiding the need for expensive periodic calculations. The performance of target-specific AIMNet2 interatomic potentials is illustrated across diverse chemical systems relevant to pharmaceutical, optoelectronic, and agrochemical applications, demonstrating their promise as efficient alternatives to full DFT calculations for routine CSP tasks.

Introduction

Molecular crystals form through favorable packing of constituent species, stabilized by complex interactions including intramolecular forces, intermolecular dispersion, electrostatics, and polarization. The arrangement of molecules in crystals is crucial for their functionality, as it directly influences the physical and chemical properties of the solid form.[1] The crystal energy landscape often contains numerous energetically competitive minima within a narrow energy window (typically 2-7 kJ/mol), leading to the ability of a compound to crystallize in multiple solid forms, known as polymorphism. A Cambridge Structural Database (CSD) study has found that nearly 46% of crystal structures containing organic molecules exhibit polymorphism.[2]

The occurrence of polymorphism is influenced by several factors, most commonly by variations in experimental crystallization conditions.[3–5] At the molecular level, polymorphs may

differ in orientation, directional bonding, conformational degrees of freedom, and subtle distortions in bond distances and angles, shaped by the surrounding crystalline environment.[6] These polymorphic variations significantly impact physicochemical properties despite identical molecular composition, affecting melting points, solubility, density, hardness, optical characteristics, and electrical properties.[7–12] Understanding polymorphism is thus crucial for applications in pharmaceuticals, organic electronics, and energetic materials, where specific polymorphs can deliver targeted performance improvements.[12–14] It is often desirable to discover previously unknown, and possibly more stable, polymorphs with improved properties. In fact, commercial CSP studies of 41 pharmaceutical compounds have revealed that 15-45% of currently available small-molecule drugs exist in polymorphic forms other than their most thermodynamically stable counterpart.[15] This presents significant risks within the pharmaceutical industry, as the use of less stable polymorphs could result in inconsistent drug performance, reduced efficacy, and potential safety concerns, impacting both patient health and the reliability of treatments.

Computational CSP offers an alternative to the time-consuming, expensive and potentially hazardous (in the case of energetic materials) process of experimental polymorph discovery by developing methods that accurately identify molecular crystal polymorphism and determine structure(s) that exhibit desirable properties. CSP involves two key challenges: structure generation across a vast configuration space and accurate stability ranking of candidate structures. Ranking methods must distinguish between polymorphs separated by mere kilojoules per mole (kJ/mol). This demands computationally intensive electronic structure approaches[16] for modeling molecular crystals, most commonly dispersion-corrected DFT.[17–29]

A good metric of progress in CSP is provided by the results and methods presented in the CCDC blind tests.[30–35] The CCDC CSP blind test is an internationally recognized scientific challenge that serves as the gold standard for evaluating CSP methods. Participants

3

attempt to predict the 3D crystal structures of "target" molecules solely from their 2D molecular structures, without prior knowledge of their experimentally determined solid forms. The CCDC blind test holds particular importance for the pharmaceutical industry, as predicting polymorphism impacts drug development and formulation. With each iteration, the CSP blind test has featured increasingly complex target molecules, reflecting the evolving capabilities of prediction methods. The seventh CCDC CSP blind test represented a significant step up in complexity, featuring a diverse array of challenging systems including metal-organic complexes, a co-crystal with varying stoichiometry, a salt with a disappearing polymorph, and a large flexible pharmaceutical compound with multiple rotatable bonds. The breadth of these targets illustrates the widespread impact that improving computational CSP workflows can have, with the growing challenge continuing to propel methodological innovation in the field of crystal structure prediction. The seventh CCDC CSP blind test employed a two-phase approach for the first time. In Phase I[36] (structure generation), researchers submitted a ranked list of predicted structures that they determined to be most likely to match the experimental forms of the given targets. In Phase II[37] (structure ranking), participants were provided with a common set of structures and were required to relax and rank them based on their stability.

In the first few CSP blind tests, force fields were the only option due to the computing resources available at the time. It was only in the fourth blind test that dispersion-corrected DFT was used on a large scale for final energy minimization and energy ranking of the putative crystal structures.[33] Although dispersion-corrected DFT has become the community-accepted standard for final structure ranking, force fields are still typically employed for pre-screening in the early stages of CSP workflows, and in some cases also in the final ranking process. Recent examples of this include the force fields employed in the sixth and seventh CCDC CSP blind tests,[36–38] such as COMPASS (2.8),[39] the DREIDING,[40] atomic multipole force field, SAPT (DFT) fitted force fields,[41] XtalPi force fields (XFF)[42] and tailor-made

4

force field.[43]

The CSP blind tests have shown repeatedly that general-purpose force fields failed to reach the desired accuracy. This has led to the emergence of tailor-made force fields (TMFFs), fit specifically to a particular target molecule using first principles data. TMFFs can greatly reduce computational expense while still achieving high accuracy.[44,45] One of the earliest TMFF utilized for CSP, developed by Neumann et al., was effectively applied starting from the fourth CCDC CSP blind test, consistently achieving high success rates.[33,34,38,46–50] Another example is the CSP procedure developed by Szalewicz et al.,[25] which begins with a two-dimensional monomer graph and employs a two-body, rigid-monomer ab initio-based force field (aiFF) derived from quantum mechanical calculations for molecular dimers. This approach was shown to work well for a set of selected test systems, but not in a blind test. Polarizable force fields have also been shown to be effective for CSP.[51–53] However, fitting them requires extensive chemical expertise and laborious parameterization, leading to a high computational cost and significant effort, as is often the case with TMFFs and other system-specific force field approaches. Inconsistent parameterization protocols and variable fidelity have limited the adoption of polarizable force fields.

A promising alternative to both the empirical TMFFs and full DFT calculations are machine-learned interatomic potentials (MLIPs) which can capture complex atomic interactions efficiently while maintaining accuracy comparable to high-level computational methods. MLIPs can achieve near *ab initio* accuracy while scaling linearly ($O(N)$) by using a flexible functional form and training on reference data of energies, forces, and other multimodal atom state information to precisely capture the potential energy surface (PES) of the target chemical system.[54–60] Recent neural network potentials like AIMNet2[61,62] and MACE-OFF23[63] have demonstrated improved accuracy and transferability for neutral and, in the case of AIMNet2, charged molecules spanning diverse areas of organic and biochemical space. This

5

progress is especially relevant to the CSP problem, where the use of MLIPs and other machine learning techniques has steadily grown in recent years.[64–76] Another aspect of CSP that can benefit from MLIPs is free energy calculations. Polymorph ranking traditionally relies on static lattice energies at 0 K, neglecting thermal effects.[77] Hoja et al.[78] demonstrated that accounting for vibrational free energy is essential for accurately describing polymorph stabilities at finite temperatures, as it accounts for entropic effects in the Helmholtz free energy. Previous studies have shown that vibrational and thermal effects contributed an average of 6 kJ/mol and 1.6 kJ/mol in the revised X23 benchmark,[79] and 5.5 kJ/mol each in a set of 31 energetic materials.[80] Evaluating these effects with DFT is computationally demanding and can only be performed for a handful of crystal structures. With MLIPs, these calculations become more feasible, even for a large number of crystal structures with many atoms.

MLIPs were used in a blind test for the first time in the seventh CCDC CSP blind test. Our team (Group 16) used system-specific AIMNet2 MLIPs for both the structure generation and ranking phases. In the structure generation phase, MLIPs were used for relaxation and ranking of structures generated by Genarris.[81] In the ranking phase, MLIPs were used for relaxation, energy evaluation, and free energy corrections. Here, we present a detailed description of our CSP workflow and report a post hoc analysis of our method's performance. AIMNet2 (second generation atoms-in-molecules network) is a chemically inspired, modular deep neural network molecular potential that combines ML-parameterized short-range and physics-based long-range terms to attain generalizability. This approach captures the necessary physics while drastically reducing computational costs, enabling the screening of millions of candidate crystal structures. Rather than using general-purpose AIMNet2 models in the CSP blind test, system-specific versions (analogous to TMFFs) were fine-tuned and applied for better accuracy. Instead of relying on DFT calculations for periodic crystal structures, we trained AIMNet2 exclusively on molecular cluster (n-mer) data, which significantly reduced the computational cost of training data acquisition. In addition to static

6

lattice energy, we incorporated vibrational and thermal effects through the harmonic and quasi-harmonic approximations. The effectiveness of our methodology is highlighted by the results of the seventh CCDC CSP blind test. In the structure generation phase our submission achieved the highest success rate (67%; 4 out of 6 possible structures generated for the targets we submitted) among academic teams, and third overall. In the structure ranking phase two other teams employed MLIPs, based on Gaussian process regression (Group 12) and a transfer learning-based neural network model (Group 15). Our system-specific AIMNet2 potentials were the only successful machine learning submission that consistently aligned with experimental results and DFT. This demonstrates the promise of system-specific AIMNet2 potentials as an efficient alternative to DFT-based methods.

Methods

Training and Validation Dataset

To construct a reliable and extensible AIMNet2 interatomic potential for accurate modeling of periodic crystal structures, we assembled an extensive dataset of n-mer systems. This dataset includes a broad distribution of molecular and intermolecular geometries formed mainly by monomers, dimers, trimers and tetramers, with a smaller number of pentamers, hexamers, heptamers and octamers. Our preliminary studies indicated that the higher order set of n-mers is necessary for accurately capturing the intermolecular interactions specific to the CSP blind test targets. This section provides an overview of the procedure for generating the training datasets using a large pool of candidate crystal structures and various sampling techniques.

Conformer Generation

Targets XXVII, XXXI, XXXII and XXXIII were selected from the seventh CCDC CSP blind test. For each CSP blind test target, a dense conformer pool was generated and used as the

7

input for the random crystal structure generator, Genarris.[81] The approach involved exhaustive enumeration of torsional degrees of freedom based on pre-defined dihedral angles from a knowledge-based list and the rotatable bond count, utilizing either OpenEye OMEGA[82] or the RDKit ETKDGv2 method.[83,84] Subsequently, this set was sampled using geometric and energy considerations. The criteria included generating all flexible ring conformations in a molecule, controlling the enumeration of non-terminal nitrogen atoms, sampling hydrogen locations for amines, and establishing an energy window. Conformers were accepted if the similarity threshold exceeded 0.3 Å and the calculated strain energy was less than the sum of the energy window and the energy of the global minimum conformer, set at 15 kcal/mol (62.76 kJ/mol). Other codes and techniques, such as the BIOVIA Materials Studio (MS) Conformers,[85] and the Confab algorithm,[86] were also applied to aid in the systematic generation of diverse low-energy conformers for this task. BIOVIA MS enables the traversal of the potential energy surface (PES) by sampling torsion angles using methods such as systematic grid searching, random sampling, and Boltzmann's jump, which vary torsion angles to generate multiple conformers and identify low-energy structures. Confab adopts a methodology similar to OpenEye OMEGA, focusing on diversity selection through heavy-atom root-mean-square deviation (RMSD) relative to the already stored conformers. Our conformer generation strategy integrated all the aforementioned methods to enhance coverage of the conformational space and ensure a diverse, low-energy set of initial conformers for each of the target molecule. For the co-crystal salt Target XXXIII, conformers were only generated for the anion fragment. The cation (morpholine) fragment is often locked in a chair conformation that best satisfies its hydrogen bond and packing interactions within the crystal lattice, so no additional morpholine conformations were explored.

Crystal Structure Generation with Genarris

In the structure generation phase of the seventh CCDC CSP blind test, Genarris[81] was used to generate crystal structures. These structures were also used for the purpose of extracting

8

n-mers to train system-specific AIMNet2 potentials (for the additional targets we submitted in the ranking stage, n-mers were extracted from the structures provided by the CCDC). Genarris is an open source Python package for generating random homomolecular crystal structures with physical constraints. After selecting a molecular configuration from the conformer pool prepared using the methods described above, Genarris initiates the generation process by estimating the unit cell volume using PyMoVE,[76] an open source Python package for molecular volume estimation. It then automatically identifies all compatible space groups based on the molecular point group symmetry and the specified number of molecules per unit cell (Z), including those where molecules occupy special Wyckoff positions. Following this, Genarris selects a compatible space group and generates a unit cell within a volume distribution window around the predicted unit cell volume while maintaining space group symmetry. Genarris then places the first molecule randomly in the unit cell. The remaining molecules are positioned using space group symmetry operations. To ensure physically plausible structures, Genarris conducts the following intermolecular distance analysis to ensure that atoms from different molecules are sufficiently spaced apart by assessing the comparison

$$d_{A,B} > s_r(r_A + r_B) \tag{1}$$

where $d_{A,B}$ is the distance between atoms A and B from different molecules, s_r is the specific radius ratio (chosen between 0.7 - 0.9), and r_A and r_B correspond to their respective van der Waals radii. In the seventh CCDC CSP blind test, a development version of Genarris was used. Structures were constructed by picking a molecular geometry uniformly at random from a set of diverse conformers for each generation attempt. In addition, this version of Genarris included a preliminary implementation of the "Rigid Press" algorithm (to be described elsewhere). When Rigid Press is used, the unit cell is initially expanded to facilitate molecule placement and then compressed using a rigid body potential to achieve close packing. During this process, the molecular positions and orientations can change under

9

the s_r constraint, while their internal geometries remain fixed (or rigid). The preliminary version of Rigid Press that was used for the seventh CCDC CSP blind test did not preserve space group symmetries. The number of generated crystal structures varied depending on the target.

Molecular Cluster Sampling

Following candidate crystal structure generation, the molecular shell functionality of the CCDC Python Crystal API[87] combined with a custom-made sampling function, was used to extract molecular clusters from the pool of generated crystal structures. This function yields a set of intermolecular complexes that include atoms within a specified distance cutoff from the central molecule, from which n-mers were extracted employing multiple combinations of varying n size (n = 1 - 8). n-mers were further sampled through short constrained gas phase molecular dynamics simulations in xTB package using the GFNFF-xTB force field[88] at 300 K for up to 5000 ps to capture out-of-equilibrium conformations. The SHAKE algorithm,[89] which is used to constrain bond lengths, was disabled. Tethering spring forces were applied to atomic positions using force constants of 0.005 eV/Å^2. It should be noted that GFNFF-xTB was only used for the purpose of efficient sampling, and was not used for training the AIMNet2 MLIPs. Consequently, the final distribution of n-mers ranging from 1 to 8, and the size of the n-mer dataset, varied for each CSP blind test target based on heuristics. After molecular cluster sampling, single-point DFT calculations were performed to obtain the reference energy, atomic forces, partial charges, and dipole data, but only for the monomer set, not the enitre n-mer dataset.

The final training set was derived from a pool of n-mer data (n = 1 - 8) through iterative cycles of active learning (AL), according to the following procedure. We began by calculating energy uncertainties for dimers using four AIMNet2 models, each trained on a small random subset of monomer reference DFT data. The uncertainty σ was defined as

$\frac{E_{std}}{\sqrt{N}}$, where E_{std} represents the standard deviation among the predictions from four AIMNet2 models and N is the number of atoms in the sample (n-mer). If the uncertainty of a dimer exceeded twice the median uncertainty of the entire dimer set ($\sigma > 2 \times median(\sigma)$), the model predictions for that structure were deemed highly uncertain. We then carried out single-point DFT calculations at the same level of theory for these high-uncertainty dimer structures and added them to the initial monomer training pool. The AIMNet2 models were subsequently retrained for the next AL cycle to process trimers. The procedure was then continued iteratively for the remaining n-mer sets, one by one. This iterative approach allowed for the identification of the most diverse and chemically informative n-mers, thereby improving the accuracy and robustness of the AIMNet2 potentials. The AL procedure was not applied to the co-crystal Target XXXIII due to differences in the types of n-mers generated for this system. The final training set for this target comprised charged cation and anion monomers; neutral and charged dimers of the cation and anion; cation-anion fragment pairs; charged cation trimers; combinations of one anion with two cations; and two anions with one cation.

Model Architecture and Training

In this work, a developmental version of the AIMNet2[62] model was trained to learn atomic contributions to the total interaction energy, atomic forces, and atom-centered partial charges calculated using Hirshfeld partitioning. The models were trained to reproduce DFT-calculated properties and included pairwise C6 and C8 coefficients for explicit long-range atomic-charge dependent London dispersion corrections. All AIMNet2 calculations were performed using the PyTorch framework.[90] For more information about the model architecture and the computational resources used for training, please refer to SI[†] and the original AIMNet2 publication.[62]

System-specific AIMNet2 models for the attempted CSP blind test targets were trained

11

on diverse molecular clusters extracted from crystal structures (either generated by Genarris or provided by the CCDC). For these molecular clusters, single-point DFT data (energies, forces, partial charges and dipoles) was calculated, circumventing costly periodic calculations. In the interest of computational efficiency, low accuracy models were first trained using energies, atomic forces, and partial charges derived from the semi-empirical tight-binding GFN2-xTB[91] molecular dynamics trajectories of these clusters. During the preliminary testing stage, it was observed that initializing the AIMNet2 model weights with GFN2-xTB data led to enhanced training stability and faster convergence, in contrast to training directly on data at the DFT-level of theory. Essentially, this pre-training allows the AIMNet2 models to establish rudimentary, albeit lower accuracy, coverage of target-specific potential energy surfaces with a computational cost that is negligible compared to exhaustive DFT. The final set of models underwent training on the DFT reference data for the n-mers of each target molecule, obtained through the AL procedure, while excluding all tight-binding initialization data.

These trained models were subsequently employed to relax crystal structures with a custom BFGS optimizer implemented in PyTorch, each adhering to a root-mean-squared (RMS) atomic force convergence criterion of 0.001 eV/Å. Only the two-body dispersion term was considered during the optimization process. Once the final optimized crystal geometry was obtained, the three-body dispersion correction was also included. D4-dispersion[92] corrections and Ewald summation[93] (with an accuracy of 10^{-8} and a real space cutoff of 15 Å) were incorporated throughout all geometry optimization and thermodynamic calculations to more precisely account for long-range interaction contributions, especially those extending beyond the atomic environment cutoff of the AIMNet2 models. For more information on the training set and the number of crystal structure optimizations conducted for each target, please refer to Table S1 in the SI†.

12

Free Energy Corrections

The relative stability evaluation in the ranking phase of the CSP blind test was performed based on the Gibbs free energy (G) using the quasi-harmonic approximation (QHA). The QHA is essentially a collection of harmonic approximations (HAs) executed across several fixed volumes, aiming to extract physical insights about thermally induced volume effects. The finite temperature and vibrational contributions to the the Helmholtz free energy (F) of a crystal were calculated using the HA and QHA. In the HA, which neglects thermal expansion,[20,22,94] the Gibbs free energy is defined as

$$G = E_{latt} + F_{vib} + pV, \tag{2}$$

where E_{latt} represents the equilibrium ground-state energy, F_{vib} accounts for the vibrational contribution to the Helmholtz free energy (F), and pV is the pressure-volume work, which is usually negligible at ambient pressure. In the HA, the vibrational free energy per unit cell is calculated from the harmonic phonon frequencies as follows

$$F_{vib} = \frac{1}{N_{\mathbf{q}}} \sum_{i=1}^{3n} \sum_{\mathbf{q}}^{N_{\mathbf{q}}} \frac{\hbar \omega_{i,\mathbf{q}}}{2} + k_B T \ln\left[1 - \exp\left(-\frac{\hbar \omega_{i,\mathbf{q}}}{k_B T}\right)\right], \tag{3}$$

where k_B is Boltzmann's constant, T is the temperature, $\hbar$ is the reduced Planck's constant, and $\omega_{i,\mathbf{q}}$ represents the phonon frequency of mode i at wave vector $\mathbf{q}$. The sums extend over the $3n$ phonon branches (n is the number of atoms in the unit cell) and the $N_{\mathbf{q}}$ sampled $\mathbf{q}$-points in the first Brillouin zone. The first term in Eq. (3) denotes the zero-point energy (ZPE), which corresponds to F_{vib} at $T = 0$ K, while the second term represents the temperature-dependent component of F_{vib}, commonly referred to as the Kirchhoff term.

The vibrational effects were computed via the HA as implemented in Phonopy,[95] an open-source Python package designed for phonon calculations at both harmonic and quasi-harmonic

levels. Within the Phonopy framework, the phonon density of states and the corresponding thermodynamic properties are computed by constructing the dynamical matrix from force constants. To obtain these force constants, the default finite displacement method was employed, where small atomic displacements (0.005 Å) were systematically applied to atoms in the optimized crystal structure. For each displacement, forces on all atoms were calculated via a single-point calculation using DFT. Supercells were constructed to ensure that the lattice vectors extended at least 10 Å in each direction, minimizing interaction between periodic images and providing accurate force calculations. The phonon frequencies were then obtained by diagonalizing the dynamical matrix at various $\mathbf{q}$-points in the Brillouin zone. Using a sufficiently dense $\mathbf{q}$-mesh (8×8×8), the vibrational free energy contribution was calculated with Eq. (3). The Helmholtz free energy (F) was then determined by combining the static lattice energy with these vibrational contributions at the fixed optimized volume, providing a framework for comparing relative stabilities among predicted crystal structures at finite temperatures.

The harmonic phonon theory cannot predict thermal expansion, which can be modeled within the QHA.[96] In this approach, the effect of volume on the phonon frequencies is taken into account, and the free energy at a given temperature is minimized with respect to volume as follows

$$G(T, P) = \min_{V} \left[U(V) + F_{vib}(T, V) + pV \right]. \tag{4}$$

The QHA was employed to account for thermal expansion effects on the free energy of crystal structures. Using Phonopy, we extended beyond the standard HA by explicitly calculating vibrational properties at multiple unit cell volumes. After the initial crystal optimization without any applied external pressure, a series of 16 volume-constrained geometry optimizations was conducted for each crystal structure, wherein external isotropic pressure was exerted to modulate the unit cell volume (V) within the range of −5% to +15%, spanning from 0.95 to 1.15 times the equilibrium volume at 300 K with uniform increments of 0.0125. For

14

each volume-constrained structure, we calculated the phonon spectrum and corresponding Helmholtz free energy (F) at both 0 K and 300 K, using a well-converged 8×8×8 reciprocal space sampling mesh. The free energy versus volume ($F - V$) data were then fitted to the Birch-Murnaghan equation of state[97] to determine the minimum Helmholtz free energy and corresponding volume at each temperature of interest. The Gibbs free energy (G) was obtained as the minimum value along this $F - V$ curve, effectively capturing anharmonic contributions arising from thermal expansion. Temperature-dependent properties, including the constant pressure heat capacity (C_p), were subsequently derived from the temperature dependence of the Gibbs free energy according to the following relation

$$C_p(T, p) = -T \frac{\partial^2 G(T, p)}{\partial T^2}. \tag{5}$$

During the second phase of the CCDC CSP blind test, free energy ranking was performed by evaluating the Helmholtz free energy (F), with vibrational contributions and thermal expansion effects included. It is worth emphasizing that the efficiency of the AIMNet2 models supports sufficiently dense volume sampling without a significant computational overhead, with the final choice of 16 samples being informed by an accuracy-volume scan density analysis performed during initial testing. Each expanded or contracted crystal structure underwent a fixed-volume variable-cell geometry optimization using a custom BFGS optimizer implemented in PyTorch. The optimization convergence criteria are outlined in the computational details section in SI[†]. For each optimized geometry along the volume scan, a HA calculation was performed using Phonopy. Unlike the typical and computationally expensive finite displacement approach with DFT, atomic force constants were obtained directly from the analytical Hessian, derived through automatic differentiation applied to the AIMNet2 model outputs. Hence, this approach offers a significant advantage in conducting QHA/HA calculations with MLIPs, as it is more efficient than its numerical counterpart and avoids potential errors associated with choosing improper displacement sizes.

Density Functional Theory Calculations

In the molecular cluster (n-mer) reference data generation process, we utilized the B97M[98] functional with the Karlsruhe triple-zeta basis set featuring two sets of polarization functions (def2-TZVPP)[99,100] in the gas phase to obtain data for training each system specific MLIP. All single-point energy calculations were carried out using the Orca 5.0.3 software.[101] Additionally, molecular dipole-dipole dispersion coefficients and polarizabilities were included using the DFT-D4 python package[92] based on the Axilrod–Teller–Muto (ATM) formula.

All post hoc DFT calculations were conducted using the FHI-aims electronic structure code[102](version 200112). A k_grid of size $n \times a \geq 25$, where n is the number of k-points and a represents the length of the lattice vector, was utilized for all calculations.[20,21,27] The generalized gradient approximation of Perdew, Burke, and Ernzerhof (PBE)[103] was paired with the many-body dispersion (MBD)[104] method. Geometry relaxations, lattice energy evaluations, and free energy corrections were performed using *lower-level* settings, which correspond to the *tier 1* basis sets and *light* species defaults of FHI-aims. The experimental structures were further re-optimized with PBE+MBD using *higher-level* settings, which correspond to the *tier 2* basis sets and *tight* species defaults of FHI-aims. These structures were re-ranked based on single-point energy evaluations using the PBE-based hybrid functional, PBE0,[105] paired with the MBD method.

Results and Discussion

The seventh CCDC CSP blind test was conducted in two phases to highlight the two primary aspects of any CSP workflow: structure generation and structure ranking. In Phase I (structure generation), a list of 1500 generated crystal structures ranked by the lattice energy was submitted for each of the attempted target compounds. This list was later examined by CCDC organizers for matches with the experimentally observed polymorphs. In Phase

II (structure ranking), CCDC provided participants with a set of either 100 or 500 selected crystal structures (depending on the target), which needed to be relaxed and ranked according to their relative stability. The lists included the experimentally observed polymorphs as well as putative structures drawn from a global dataset comprising all submissions from the structure generation phase.

Figure 1 shows a high-level overview of our MLIP-based CSP workflow. The approach involves generating a large number of random crystal structures, conducting geometry optimization with system-specific AIMNet2 potentials, and subsequently choosing structures with low lattice energies. The pipeline begins with a conformational search, as described above, proceeds to crystal structure generation with Genarris, and subsequently to n-mer sampling and AIMNet2 training by active learning. Once the system-specific AIMNet2 potentials are trained, relaxation can be performed efficiently for millions of structures, producing an energy landscape. A smaller subset of the structures predicted to be most stable by AIMNet2 can be re-ranked with more accurate methods and/or with free energy corrections.
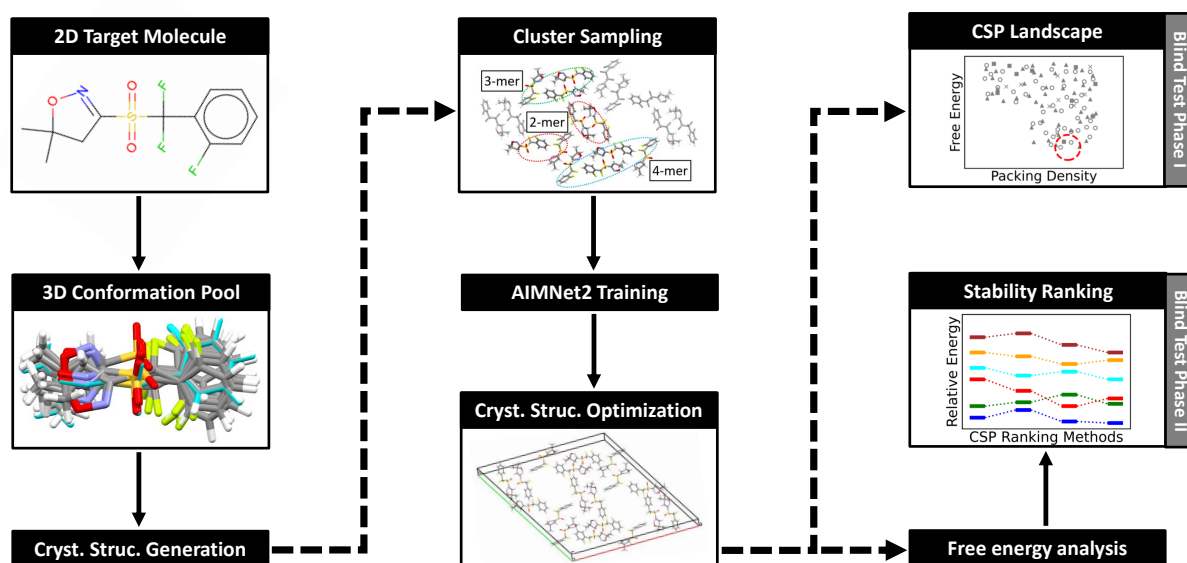

Figure 1: Schematic diagram of the CSP workflow, which was used in the structure generation phase of the seventh CCDC CSP blind test.

17

The selected CSP blind test Targets XXVII, XXXI, XXXII and XXXIII, shown in Figure 2, can be categorized into optoelectronic materials, pharmaceutical or agrochemical-like compounds, and molecules with diverse chemistry and industrial significance. Molecule XXVII, [(2,3-diiodopentacene-6,13-diyl)bis-(ethyne-2,1-diyl)]bis(triisopropylsilane) is a silicon- and iodine-substituted triisopropylsilyl (TIPS) pentacene derivative intended for use in organic electronics.[106,107] Molecule XXXI, 3-((difluoro-(2-fluorophenyl)methyl)sulfonyl)-5,5-dimethyl-2l2-isoxazolidine, is a simple agrochemical with three rotatable bonds. Molecule XXXII, or (N-(3-[2-(difluoromethoxy)-5-(methylthio)phenyl]-1-[2-(4-morpholinopiperidin-1-yl)-2-oxoethyl]-1H-pyrazol-4-yl)pyrazolo-[1,5-a]pyrimidine-3-carboxamide) is a large pharmaceutical compound with eleven rotatable bonds. Lastly, Molecule XXXIII is a 1:1 morpholine salt of 4-amino-N-(5-methylisoxazol-3-yl)-benzenesulfonamide (commercially known as Sulfamethoxazole). CSP for Target XXVII was challenging because of the conformationally flexible substituents, leading to a very large conformational space. Target XXXI was challenging, owing to the rotational disorder of the *ortho*-fluorophenyl ring and difluoromethyl group, and fluorine atoms susceptible to positional disorder. This required screening millions of trial structures for each of these targets. This was enabled thanks to the computational efficiency of the AIMNet2 models, which are faster than DFT by a factor of over a thousand. In the structure ranking phase, the efficiency of AIMNet2 facilitated free energy evaluations for hundreds of structures, including structures with large unit cells (Target XXXII). It is also noteworthy that AIMNet2 can handle charged species, such as the salt Target XXXIII, and elements such as Si, I (Target XXVII), S and F (Targets XXXI, XXXII), which other MLIPs may not support or struggle with.
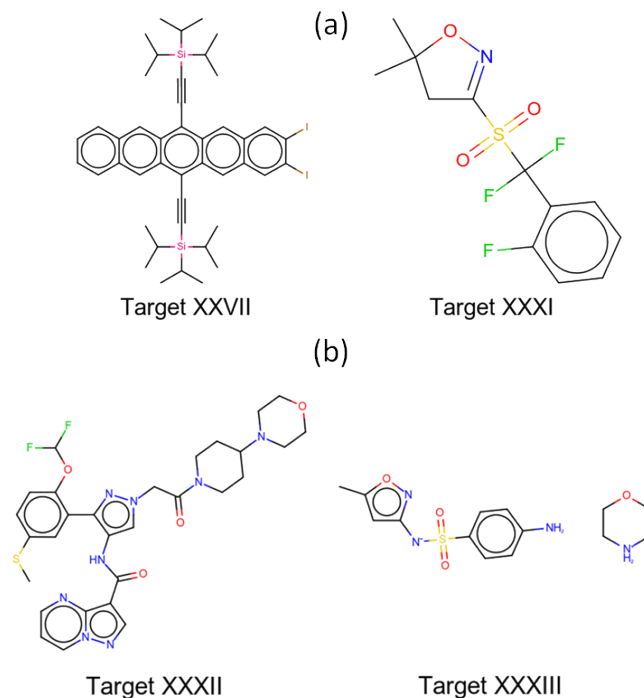
18

Figure 2: (a) Selected targets attempted in both Phase I and II, and (b) additional targets attempted only in Phase II of the seventh CCDC CSP blind test.

CSP Blind Test Phase I: Structure Generation

For target XXVII, AIMNet2 models were trained within 8 AL iterations on a total of 9.5×10^4 n-mers, comprising 13×10^3 monomers, 75×10^3 dimers, and 7×10^3 trimers. An ensemble of 1.8×10^4 conformers was used to generate 5.0×10^6 crystal structures with Z = 2, and 4 with Genarris. After geometry optimization with AIMNet2, the structures were symmetrized using Spglib,[108] with a distance tolerance in cartesian coordinates (symprec) of 0.2 Å. Duplicate structures were then removed using pymatgen's[109] StructureMatcher, applying a site tolerance of 0.2 Å along with default fractional length (lattice) and angle tolerances (stol=0.2 Å, ltol=0.2 Å, angle_tol=5°).

For Target XXXI, AIMNet2 models were trained within 4 AL iterations on a total of 22.13×10^4 n-mers, comprising 6.3×10^3 monomers, 125×10^3 dimers, 76×10^3 trimers, and 14×10^3 tetramers. An ensemble of 658 conformers was used to generate 3.5×10^6 crystal struc-

19

tures with Z = 2, 4, and 8 with Genarris. Following geometry optimization with AIMNet2, the structures were symmetrized, and duplicates were removed using the same settings as those applied to Target XXVII. A small subset of the top-ranked AIMNet2 structures (from a total of 1.7×10^6 structures after removing duplicates) was selected for DFT calculations and re-optimized with FHI-aims using PBE+MBD method and *lower-level* settings. The final submission included a combination of 1006 structures optimized by DFT and 494 structures optimized by AIMNet2, labeled by the prefixes 'aimnet' or 'dft', respectively.
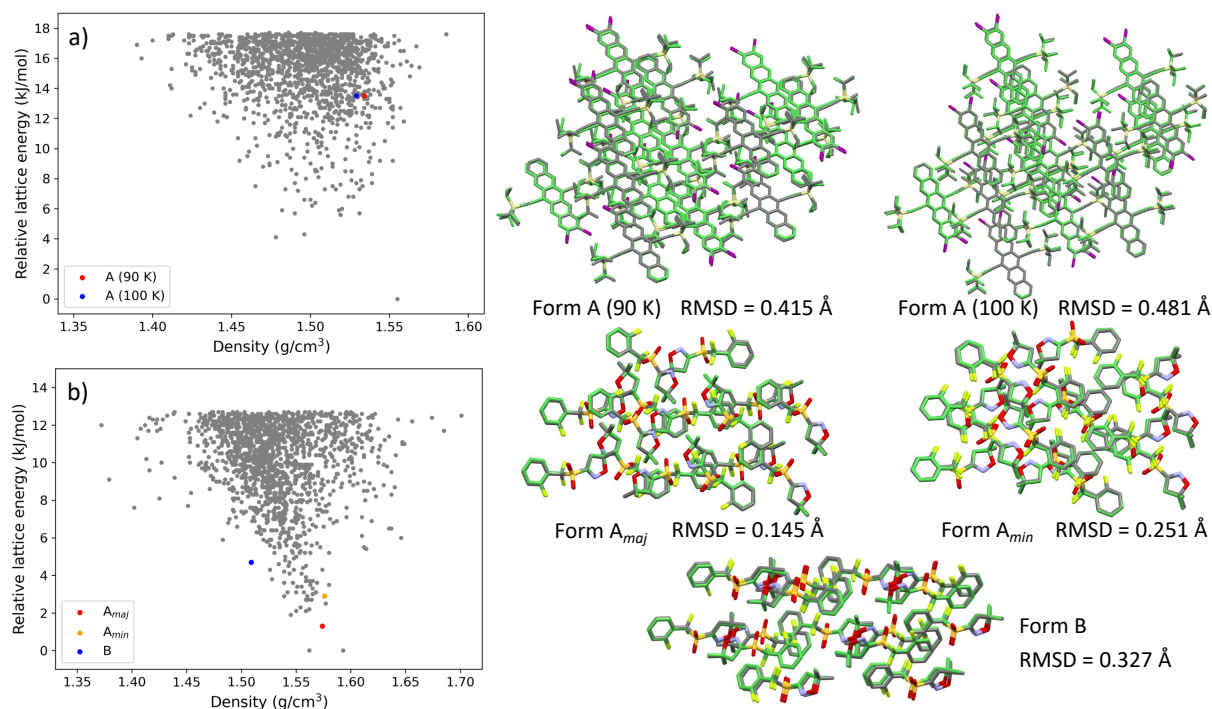

Figure 3: Results of the structure generation phase: Relative lattice energy vs density landscapes obtained with AIMNet2 for (a) Target XXVII and (b) Target XXXI. Overlays of the structures relaxed with AIMNet2 on top of the respective experimental structures (colored by element) are also shown. Hydrogen atoms were omitted for clarity.

The known crystal structure of Target XXVII, Form A, crystallizes in the $P\bar{1}$ space group with a single molecule in the asymmetric unit. An initial structure was obtained at 90 K before the blind test began and later re-refined as having I/Br disorder as a result of Br contamination during the synthesis (CSD refcode: *XIGYUL*). Eventually, the structure was re-determined from pure material with diffraction data collected at 100 K (CSD refcode:

XIFZOF01). It was found that the bromine impurity did not significantly impact the crystal structure.

The experimental form of Target XXVII was successfully generated during the structure generation phase. In Figure 3a, the AIMNet2-based relative lattice energy is shown as a function of crystal density for the 1500 submitted structures. One structure is found to be significantly lower in energy than the others, although it is not the closest to the experimental form. The experimental form (90 K) is ranked much higher in energy (ranked 297/1500), nearly 13 kJ/mol above the minimum energy structure.

For Target XXXI, three crystal structures are known, labeled as Forms A, B, and C. This target was the most polymorphic among those included in the seventh CCDC CSP blind test, which offered significant challenges in both the structure generation and structure ranking phases. Form A (CSD refcode: *ZEHFUR02*) crystallizes in space group $P2_1/c$ with one molecule in the asymmetric unit. The *ortho*-fluorophenyl ring is disordered over two sites, designated Form A_{maj} and A_{min}, in a 60:40 ratio. Form B (CSD refcode: *ZEHFUR*) also crystallizes in space group $P2_1/c$ with one molecule in the asymmetric unit. Form C (CSD refcode: *ZEHFUR01*) crystallize in space group $R\bar{3}$. It is a porous structure with void channels, whose formation is likely promoted and stabilized by the presence of solvent molecules. It has been experimentally determined that Form B is more stable than Form A at 346 K and below.[36]

Figure 3b shows the AIMNet2-based relative lattice energy vs crystal density plot for the 1500 submitted structures of Target XXXI. Three of the four known experimental structures were generated, A_{maj}, A_{min}, and B. The solvent stabilized Form C, which is expected to be relatively high in energy due to its porous nature, was not generated by any of the participants. It is generally challenging to predict the occurrence of porous molecular crystal

21

structures that are relatively high in energy.[110] Here, the experimental forms are positioned in the relatively low-energy region of the landscape compared to Target XXVII. The A_{maj} form (ranked 3/1500) is correctly ranked as more stable than the A_{min} form (ranked 15/1500). However, Form B (ranked 46/1500) is incorrectly ranked as less stable than Form A (see further discussion of stability ranking below). Only two roughly degenerate structures were found to have a lower energy than Form A_{maj} by less than 2 kJ/mol. Upon further examination, we found that these two structures share the same space group, $P2_1/c$, as Form A_{maj}, but differ in their unit cell orientations and molecular packing arrangements compared to Form A_{maj}. Additionally, a separation of ~ 1 kJ/mol exists between the Forms A_{min} and A_{maj}, with a further gap of about 1-2 kJ/mol between the Forms A_{min} and B.

For both targets the energy landscapes obtained from our CSP workflow manifest the so-called over-prediction problem.[111] CSP workflows tend to significantly overestimate the number of possible crystal structures because structures that appear to be distinct energy minima at 0 K may correspond to the same structure at finite temperature. In particular, for Target XXVII, whose crystal structure exhibits disorder in the TIPS groups, multiple local minima with different TIPS group conformations coalesce into one disordered structure at finite temperature. In such cases, our workflow could benefit from further clustering and coarse-graining of the energy landscape in the future.

CSP Blind Test Phase II: Structure Ranking

In the structure ranking phase, we received 100 structures of Target XXVII. The experimental structure of Form A determined at 90 K was absent from this set. Instead, four structures that exhibited the same core packing but varied in their isopropyl conformations were included. These were numbered 28, 38, 59, and 61 and had $RMSD_{30}$ of 0.53, 0.80, 0.83, and 0.57 Å, respectively. For Target XXXI, three experimental forms were included in the list of 100 structures: A_{maj} (structure 98), A_{min} (structure 1), and B (structure 25).

Form C, identified as a channel-type solvate with unresolved solvent, was excluded from the ranking exercise, but is included in our analysis. For Target XXXII, a list of 500 structures was provided, which included the two polymorphs, whose structures were determined at 90 K. Form A (CSD refcode: *JEKVII*) crystallizes in the $P\bar{1}$ space group, with one molecule in the asymmetric unit and rotational disorder of the difluoromethyl group (structure 317). Form B (CSD refcode: *JEKVII01*) also crystallizes in the $P\bar{1}$ space group, but with two molecules in the asymmetric unit (structure 232). For Target XXXIII, we received a list of 500 structures, including its two known forms. Both Form A (CSD refcode: *ZEGWAN*) and Form B (CSD refcode: *ZEGWAN01*) contain one formula unit in the asymmetric unit. In both forms, proton transfer from the sulfonamide nitrogen to morpholine results in the formation of a salt. Form A (structure 233), determined at 296 K, crystallizes in the monoclinic space group $C2/c$ with a tetrameric packing motif. Form B (structure 452), determined at 297 K, crystallizes in the orthorhombic space group $Pna2_1$. Its crystal structure consists of zigzag chains of sulfamethoxazole linked via morpholine molecules.

For all four targets, the structures we received from the CCDC were relaxed and ranked with system-specific AIMNet2 potentials. Further free energy calculations were also performed using AIMNet2. Our submission consisted of two lists, one ranked by lattice energy and the other by free energy. For Target XXVII and Target XXXI, we used the AIMNet2 potentials already trained during the first phase of the blind test. For Target XXXII and Target XXXIII, new AIMNet2 potentials were trained for the ranking phase. For Target XXXII, AIMNet2 potentials were trained within 4 AL iterations on a total of 0.62×10^4 n-mers, comprising 1.03×10^3 monomers, 2.68×10^3 dimers, and 2.52×10^3 trimers. For Target XXXIII, AIMNet2 potentials were trained on a total of 2.25×10^4 n-mers, comprising 1.59×10^3 monomers, 7.37×10^3 dimers, and 13.58×10^3 trimers without using AL. Fewer n-mers were used during the ranking phase than during structure generation, as the latter required an interatomic potential capable of broadly exploring the crystal energy

23

landscape. In contrast, the n-mers for the ranking phase were extracted from a fixed set of structures provided by CCDC, resulting in a smaller and more focused training dataset. Free energy calculations within the QHA were performed using AIMNet2 on a set of 16 distinct volume-constrained structures at 300 K, which was used to determine the Helmholtz free energy (F) as a function of volume (V). This amounted to a total of 1600 or 8000 calculations for each target, depending on the number of structures provided by the CCDC (100 or 500). Notably, AIMNet2 enabled these QHA calculations at a speed exceeding that of DFT by orders of magnitude. In addition, for Target XXVII, Target XXXI, and Target XXXIII, we relaxed and ranked the structures provided by the CCDC using PBE+MBD with *lower-level* settings. These DFT calculations were performed for validation purposes and were not included in our blind test submission.
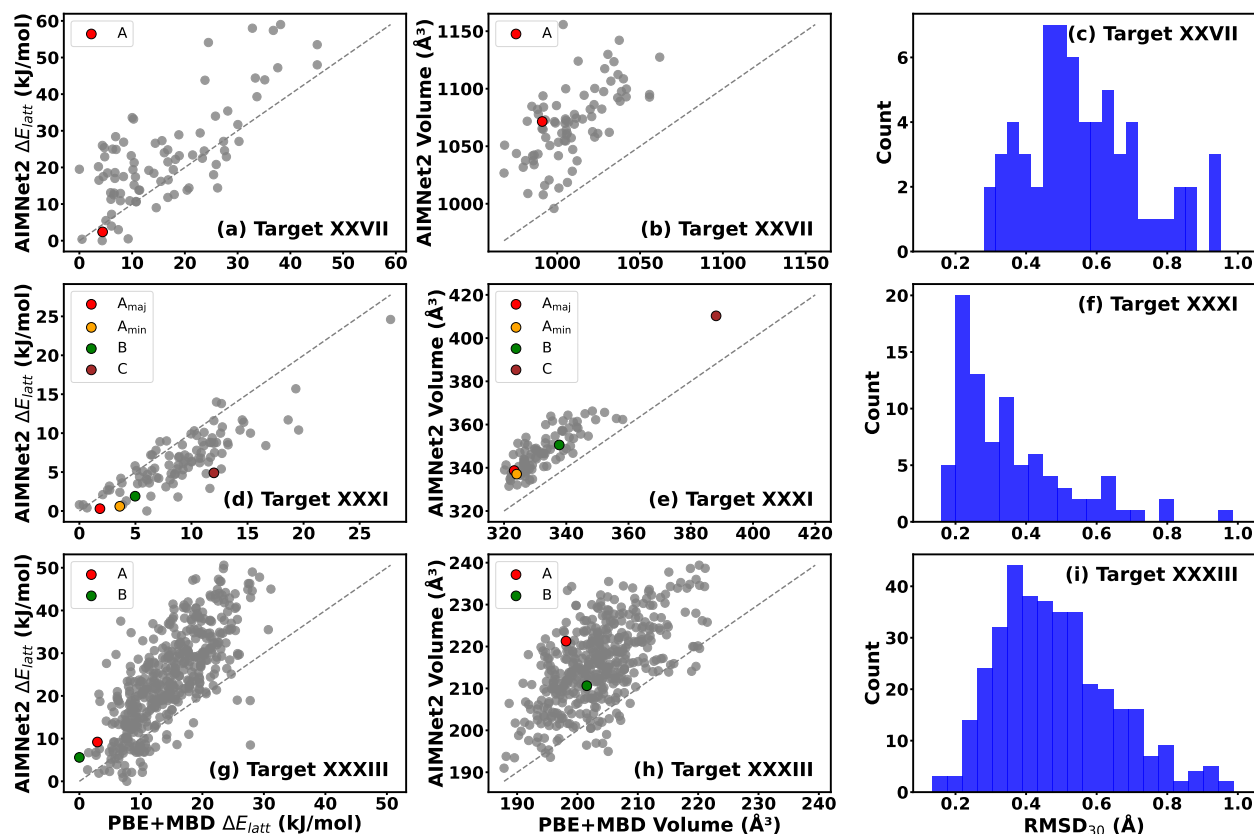

Figure 4: Assessment of the performance of AIMNet2 for the crystal structures provided in the ranking phase of the seventh CCDC CSP blind test. Comparison of the relative energy ranking, unit cell volume, and RMSD$_{30}$ between AIMNet2- and DFT-relaxed structures for Targets XXVII (top row, a–c), XXXI (middle row, d–f), and XXXIII (bottom row, g–i).

Figure 4 presents the results of the DFT validation for Target XXVII, Target XXXI, and Target XXXIII. The performance of AIMNet2 is assessed based on two metrics: the accuracy of the relaxed geometry, evaluated by comparing the unit cell volume and $RMSD_{30}$ to the DFT reference, and the relative energy ranking. Overall, AIMNet2 tends to overestimate the volume compared to PBE+MBD. This could be attributed to differences in the DFT functionals (B97M vs PBE) and dispersion correction schemes (D4 vs MBD) used in the AIMNet2 training and the reference assessment method. The best relaxation performance is obtained for Target XXXI, for which the $RMSD_{30}$ is below 0.4 Å for the vast majority of structures. The performance of AIMNet2 for Target XXXIII is somewhat worse with the $RMSD_{30}$ histogram peaked around 0.4 Å. For Target XXVII most structures have $RMSD_{30}$ values above 0.4 Å. This is also reflected in the more significant unit cell volume overestimation for this target.

For all three targets, the relative energy ranking obtained with AIMNet2 is overall reasonably consistent with DFT and the experimentally observed structures are ranked relatively low in energy (with the exception of Form C of Target XXXI). The relative energies obtained for Target XXXI match DFT more closely than for Target XXXIII and Target XXVII. The lower performance observed for Target XXVII is likely due to its larger molecular and unit cell size, which increases the system's complexity. Moreover, the presence of highly flexible TIPS side groups introduces additional interaction variability, requiring more extensive training data for the interatomic potential to achieve higher accuracy. The AIMNet2 model achieved a mean absolute error (MAE) of 2.88 kJ/mol and a root mean square error (RMSE) of 3.53 kJ/mol for Target XXXI when compared to the resulting DFT values for relative lattice energies. For Target XXXIII, the model yielded an MAE of 10.08 kJ/mol and an RMSE of 12.08 kJ/mol, while for Target XXVII, the MAE and RMSE were 8.66 kJ/mol and 11.16 kJ/mol, respectively.

Table 1 summarizes the performance of target-specific AIMNet2 models and the medium variant of the MACE-OFF23[63] pre-trained transferable organic force field, denoted as MACE-OFF23(M), with respect to DFT reference values. The MACE-OFF23 medium model was selected for comparison because it matches the AIMNet2 models from both the phases in terms of size, specifically the total number of model parameters involved. As MACE-OFF23(M) does not support the chemical element Si, results for Target XXVII are unavailable. Additionally, only relative energy rankings were computed using MACE-OFF23(M) without retaining optimized geometries; hence, unit cell volume comparisons for all three targets have not been reported at this time. Overall, AIMNet2 consistently outperforms MACE-OFF23(M), exhibiting lower MAE and RMSE values, along with higher Kendall rank correlation coefficients. Specifically, for Target XXXI, AIMNet2 achieves lower MAE by 42.3% and lower RMSE by 40.1%; for Target XXXIII, MAE and RMSE reduce by 57.9% and 55.1%, respectively, relative to MACE-OFF23(M)—demonstrating much better agreement with DFT. This improved performance can be attributed to the target-specific training of the AIMNet2 models.

Table 1: Comparison of the performance of AIMNet2 and MACE-OFF23(M) against DFT reference for the crystal structures provided in the ranking phase of the seventh CCDC CSP blind test. The best results are highlighted in **bold**. Lower MAE and RMSE, and higher Kendall rank correlation coefficient, indicate better agreement with the reference.

Target	Model	$RMSD_{30}$ (Å)	Matched (%)	MAE (kJ/mol)	RMSE (kJ/mol)	Kendall rank
XXVII	AIMNet2	0.56 ± 0.17	76	8.66	11.16	0.44
	MACE-OFF23(M)	N/A	N/A	N/A	N/A	N/A
XXXI	AIMNet2	0.36 ± 0.16	88	**2.88**	**3.53**	**0.63**
	MACE-OFF23(M)	-	-	4.99	5.89	0.52
XXXIII	AIMNet2	0.48 ± 0.16	78	**10.08**	**12.08**	**0.57**
	MACE-OFF23(M)	-	-	23.94	26.84	0.48

Table 2 presents further analysis comparing the results of geometry relaxation with AIMNet2 to PBE+MBD with *lower-level* and *higher-level* settings for the experimental forms of all four targets. We find that switching from *lower-level* to *higher-level* settings in the PBE+MBD calculations only leads to a slight reduction in the RMSD. In most cases, the

26

Table 2: Comparison of the structures optimized with AIMNet2 and DFT (PBE+MBD with *lower-level* and *higher-level* settings) for the polymorphs of Targets XXVII, XXXI, XXXII and XXXIII. The $RMSD_{30}$ values are with respect to the experimentally determined structures.

Target	Experimental Form	$RMSD_{30}$ (Å)		
		AIMNet2	PBE+MBD (*lower-level*)	PBE+MBD (*higher-level*)
XXVII	A	0.574	0.629	0.582
XXXI	A_{maj}	0.214	0.145	0.107
	A_{min}	0.240	0.207	0.226
	B	0.479	0.221	0.203
	C	0.178	0.133	0.123
XXXII	A	0.231	0.201	0.175
	B	0.353	0.300	0.296
XXXIII	A	0.359	0.182	0.151
	B	0.196	0.190	0.192

system-specific AIMNet2 potentials deliver an excellent performance with RMSD values on par with DFT.

Upon heating, molecular crystals undergo considerable expansion from their so-called electronic volume (V_{el}) at 0 K. This expansion can be split into the contribution of the zero-point vibrational energy at 0 K (V_{ZPVE}) and the contribution of thermal vibrational motion at room temperature (V_{RT}).[20,27,112] Both effects can be evaluated within the QHA, as described above. Figure 5 presents the percent error with respect to the experimental unit cell volume, obtained using PBE+MBD with *lower-level* settings when vibrational and thermal effects are taken into account, and with AIMNet2 at 0 K ($V_{AIMNet2}$), for the polymorphs of Target XXXI, Target XXXII, and Target XXXIII. The PBE+MBD electronic volume at 0 K is underestimated compared to the experimental volume in all cases except the solvent-stabilized Form C of Target XXXI. Adding the ZPVE contribution at 0 K results in some volume expansion. For Forms A_{maj}, A_{min}, and C of Target XXXI, this leads to a slightly

overestimated volume compared to experiment, whereas for Form B, and all forms of Targets XXXII and XXXIII, V_{ZPVE} is still underestimated. Considering vibrational effects and thermal expansion at 300 K leads to further volume expansion, which is particularly significant for Forms A_{maj}, A_{min}, and C of Target XXXI, resulting in considerable overestimation of the experimental volume. For Form B of Target XXXI and the two forms of Target XXXII, V_{RT} is slightly overestimated, whereas for the two forms of Target XXXIII, V_{RT} is slightly underestimated. The unit cell volume obtained with AIMNet2 is somewhat overestimated for all forms of Target XXXI and Form A of Target XXXIII, but slightly underestimated for the two forms of Target XXXII and Form B of Target XXXIII.
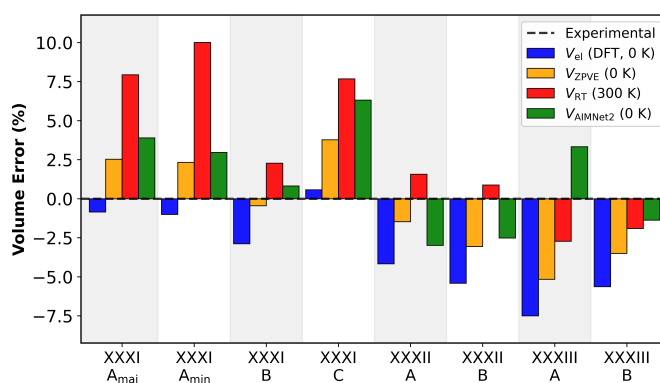

Figure 5: Comparison of the unit cell volume obtained with AIMNet2 at 0 K ($V_{AIMNet2}$) to that obtained with DFT (PBE+MBD with *lower-level* settings) at 0 K (V_{el}), with a zero point vibrational energy correction (V_{ZPVE}), and at room temperature with free energy corrections including thermal expansion within the QHA (V_{RT}), for the polymorphs of Target XXXI, Target XXXII, and Target XXXIII.

In addition to the unit cell volume, accounting for vibrational and thermal contributions can affect the relative stability of molecular crystal structures. For Target XXVII, although the experimental structure was ranked relatively high in energy out of the 1500 putative structures we submitted in the structure generation phase (see Figure 3a), out of the list of 100 structures we received in the ranking phase, the experimental structure was ranked by AIMNet2 as the most stable both in terms of lattice energy and free energy.[37] In comparison, the two other teams that used MLIPs, Group 12 and Group 15 ranked the experimental structure as #25 (lattice energy) and #56 (free energy), respectively. Seven teams (Groups

28

3, 5, 9, 10, 11, 20, 22) submitted ranking entries for Target XXVII using various DFT methods. The DFT ranking of the experimental structure ranged from #1 to #7, depending on the method used and whether free energy corrections were included. Our DFT lattice energy ranking obtained using PBE+MBD and *lower-level* settings (see Figure 4a) is #7. Hence, the DFT results presented here are consistent with the DFT results reported in the blind test.

For Target XXXI, the experimentally determined order of stability is B > A > C at both 0 K and 300 K.[37] Of the list provided by the CCDC, our team ranked Form A_{maj} as #2 based on AIMNet2 lattice energy and #3 based on free energy. Form A_{min} was ranked #4 based on lattice energy and #18 based on free energy. Form B was ranked as #10 based on both lattice energy and free energy. All three forms were within an energy window of less than 5 kJ/mol of the global minimum. The two other teams that employed MLIPs ranked the Forms A_{maj}, A_{min}, and B higher above their respective global minima as #33, #38, and #12 (Group 12, lattice energy) and #18, #13, and #12 (Group 15, free energy). Furthermore, during the analysis for Table 1, we observed that MACE-OFF23(M) ranked the Forms A_{maj}, A_{min}, and B as #9, #45, and #6, respectively, based on lattice energy. Ten teams submitted ranking entries for Target XXXI using various DFT methods (Groups 2, 3, 4, 5, 9, 10, 11, 14, 20, 22). All DFT methods predicted Forms A_{maj}, A_{min}, and B to be within 5.7 kJ/mol of the global minimum, although the order of stability varied. In most cases, Form B was ranked higher than Forms A_{maj} and A_{min}. Group 20 ranked Form B as the most stable of the three known forms and #6 overall based on free energy. Only two of these groups ranked Form B as the global minimum, Group 10 based on lattice energy (and #3 based on free energy) and Group 3, based on free energy (and #6 based on lattice energy).

Figure 6a shows the relative energy ranking of the polymorphs of Target XXXI, obtained using AIMNet2, compared to different DFT methods, with and without vibrational and thermal corrections. Form C is always ranked as the least stable and Form A_{maj} is always

29

ranked as the most stable. Based on lattice energy, Form B is ranked as less stable than Forms A_{maj} and A_{min} using PBE+MBD, PBE0+MBD, and AIMNet2. As shown in Figure 4d, PBE+MBD with *lower-level* settings ranks Form A_{maj} as #4, Form A_{min} as #9, Form B as #16, and Form C as #79. When the zero point vibrational energy contribution is added to the PBE0+MBD lattice energy, Form B is still ranked as less stable than Form A, but closer in energy to Form A_{min}. When the free energy at room temperature is calculated within the QHA, Form B becomes more stable than Form A_{min} using both PBE+MBD and AIMNet2. Thus, the ranking obtained with AIMNet2 is consistent with the DFT results presented here, which are consistent with some of the DFT results submitted by other groups in the ranking phase of the blind test.

For Target XXXII, Form B was experimentally found to be more stable than Form A.[37] This target proved particularly challenging in the ranking stage. All blind test participants ranked the two experimentally observed forms relatively high (over 3 kJ/mol) above the global minimum out of the list of structures provided by the CCDC. Our team was one of only four groups that correctly ranked Form B as more stable than Form A. Based on AIM-Net2 lattice energy, Form B was ranked as #3, approximately 6.3 kJ/mol above the global minimum and Form A was ranked as #41, around 12.4 kJ/mol above the global minimum. In this case, applying vibrational and thermal free energy corrections did not improve the results with Form B ranked as #15, approximately 13 kJ/mol above the global minimum, and Form A ranked as #38, around 15.9 kJ/mol above the global minimum. In comparison, the MLIP of Group 12 ranked Form B as #129 and Form A as #490, based on lattice energy. The MLIP of Group 15 failed to relax the structure of Form B and their optimized geometry no longer matched the experimental structure. They ranked Form A as #18 based on free energy.

Eight teams submitted ranking entries for Target XXXII using various DFT methods (Groups

30

2, 3, 4, 5, 10, 11, 20, 22). All of them ranked Form A as more stable than Form B. Figure 6b shows the relative energy ranking of the polymorphs of Target XXXII, obtained using AIMNet2, compared to different DFT methods, with and without vibrational and thermal corrections. All the methods used here consistently rank Form B as more stable than Form A. With PBE+MBD and *lower-level* settings there is a very small energy difference of 0.355 kJ/mol between the lattice energies of the two forms. The energy difference increases when switching to PBE0+MBD and when applying vibrational and thermal free energy corrections. Of the teams that submitted DFT rankings in the blind test, Group 10 used PBE0+MBD to compute lattice energies, however they ranked Form B (#30) as less stable than Form A (#13). It is possible that they may have employed an alternative implementation or different convergence settings.
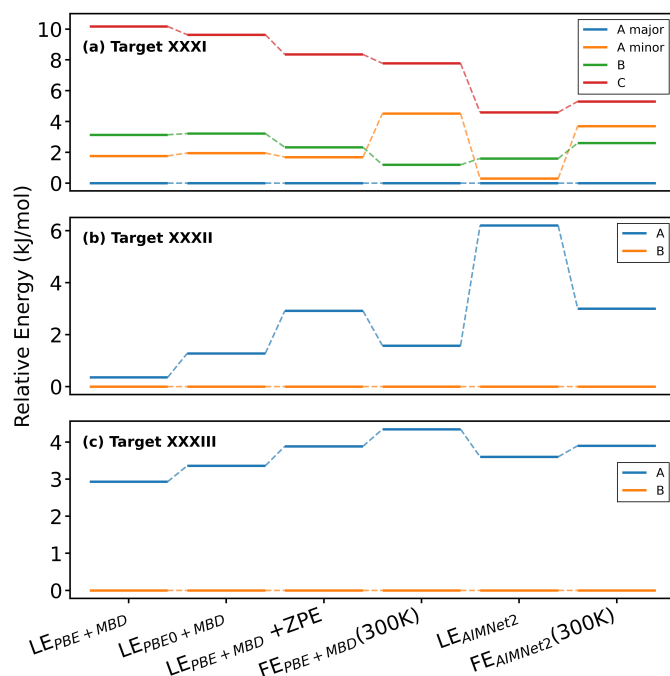

Figure 6: Relative energy ranking of the polymorphs of (a) Target XXXI, (b) Target XXXII, and (c) Target XXXIII obtained based on lattice energies (LE) calculated with PBE+MBD and PBE0+MBD, PBE+MBD with zero point energy (ZPE) corrections, and PBE+MBD free energies (FE) calculated using the QHA at 300 K, compared to AIMNet2 lattice energy and free energy.

For the co-crystal Target XXXIII, Form B was experimentally found to be more stable than

Form A, which is a meta-stable disappearing polymorph.[37] For this target, eight teams submitted ranking entries for using various DFT methods (Groups 3, 4, 5, 9, 10, 11, 20, 22). All correctly ranked Form B as more stable than Form A and all but one (Group 4) ranked Form B as the global minimum. Our PBE+MBD results, shown in Figure 4g, are consistent with the DFT results reported in the blind test, ranking Form B as the global minimum, and Form A as #6, 2.9 kJ/mol above the global minimum. AIMNet2 correctly ranked Form B as more stable than Form A, however both forms were ranked relatively high above the global minimum. Form B was ranked #20 based on both lattice energy and free energy, about 5 kJ/mol and 8.1 kJ/mol above the global minimum. Form A was ranked as #60 based on lattice energy, approximately 8.5 kJ/mol above the global minimum and #56 based on free energy, around 10.7 kJ/mol above the global minimum. The MLIPs used by other teams performed poorly for this target. Group 12 ranked Form A as #90 and Form B as #470. The MLIP used by Group 15 failed to relax Form A and its optimized structure no longer matched the experimental structure, whereas Form B was ranked as #288 based on free energy. Additionally, MACE-OFF23(M) ranked the Form A as #167 and Form B as #204 based on lattice energy. Figure 6c shows the relative energy ranking of the polymorphs of Target XXXIII, obtained using AIMNet2, compared to different DFT methods, with and without vibrational and thermal corrections. All the methods used here consistently rank Form B as more stable than Form A. Switching from PBE to PBE0 and adding vibrational and thermal corrections only increases the energy difference between the two forms.

Target XXXIII is the only blind test target for which the ranking performance of AIMNet2 was not on par with DFT, although it still performed considerably better than the MLIPs used by other participating groups. The poor performance of AIMNet2 for this target compared to Target XXVII, Target XXXI, and Target XXXII could be attributed to the different training procedure, which did not include crystal structure generation for n-mer extraction and active learning. Additionally, accurately capturing long-range Coulomb

32

interactions in salts remains a difficult challenge, representing a potential area for improvement in the AIMNet2 training. In the future, we intend to improve the training procedure for salts and co-crystals by incorporating a more sophisticated treatment of charged and long-range interactions. Overall, system-specific AIMNet2 potentials provided accuracy on par with DFT at a significantly lower computational cost. The results of the seventh CCDC CSP blind test position system-specific AIMNet2 potentials as an attractive alternative to DFT for CSP workflows.

Conclusion

In summary, in the seventh CCDC CSP blind test we have successfully demonstrated a CSP workflow based on target-specific second generation atoms-in-molecules neural network potentials, AIMNet2. System-specific AIMNet2 potentials were trained on datasets of gas phase dispersion-inclusive DFT data acquired for molecular clusters (n-mers). The training sets of n-mers were carefully curated using active learning. These system-specific potentials trained on n-mers demonstrated good transferability to crystals, thus avoiding the need for expensive periodic DFT calculations for training data acquisition.

In the structure generation phase of the blind test, the computational efficiency of AIMNet2 enabled conducting geometry optimization and energy ranking of millions of structures generated by Genarris. Our team achieved the highest success rate of academic groups and third overall with 4 out of 6 possible structures generated for the targets we submitted (67%). The structures we did not generate were the solvent-stabilized Form C of Target XXXI, which was not generated by any of the blind test participants, and the structure of Target XXIX (not discussed here), which has three molecules in the asymmetric unit (Z'=3), and was only generated by one group. In the future, our CSP workflow could be improved by automating and streamlining the training of system-specific AIMNet2 potentials, and by

addressing the over-prediction problem with clustering.

In the ranking phase of the blind test, system-specific AIMNet2 potentials demonstrated performance that was overall comparable with dispersion-inclusive DFT methods for both geometry optimization and energy ranking. It is noteworthy that the high computational efficiency of AIMNet2 made it possible to perform fast free energy evaluations accounting for vibrational contributions and thermal expansion within the quasi-harmonic approximation for thousands of crystal structures with large unit cells. The performance of AIMNet2 was superior to the MLIPs used by other groups, which showed lower ranking accuracy and occasionally encountered problems during geometry relaxations (we are unable to analyze the reasons for the poorer performance of these models, as they are not publicly available). In the future, the performance of AIMNet2 could be improved for salts and co-crystals by developing a tailored training procedure.

Finally, based on the results presented here and in the blind test papers,[36,37] we conclude that system-specific AIMNet2 machine-learned interatomic potentials have tremendous prospects as an efficient and accurate alternative to dispersion-inclusive DFT in CSP workflows.

Acknowledgement

Supporting Information Available

The supplemental information includes a description of the computational details for the DFT method, convergence information for crystal geometry optimizations, molecular volume parity plots prepared using AIMNet2 vs DFT optimized structures, and unit cell volume expansion plots for the experimental forms of different targets. ML models and submission structures for both phases (structure generation and structure ranking) of the seventh CCDC CSP blind test are available at https://github.com/isayevlab/csp7-blindtest-group16-MaromIsayev. The training data for the ML models in both phases can be found at https://figshare.com/s/bba111ee4289adc08201.

References

(1) Bernstein, J. *Polymorphism in molecular crystals 2e*; International Union of Crystal, 2020; Vol. 30.

(2) Cruz-Cabeza, A. J.; Bernstein, J. Conformational polymorphism. *Chemical reviews* **2014**, *114*, 2170–2191.

(3) Brittain, H. G.; others Polymorphism in pharmaceutical solids. *Drugs and the pharmaceutical sciences* **1999**, *95*, 183–226.

(4) Kitamura, M. Strategy for control of crystallization of polymorphs. *CrystEngComm* **2009**, *11*, 949–964.

(5) Wu, W.; Mehrman, S. J.; Zhou, Y.; Pu, S. X.; Huang, L.; Fermier, A.; Karki, S. The mechanism of the formation of enantiotropic polymorphs of carisbamate in solution crystallization. *Journal of crystal growth* **2009**, *311*, 3435–3444.

(6) Haleblian, J.; McCrone, W. Pharmaceutical Applications of Polymorphism. *Journal of Pharmaceutical Sciences* **1969**, *58*, 911–929.

(7) Cote, A.; Erdemir, D.; Girard, K.; Green, D.; Lovette, M.; Sirota, E.; Nere, N. *Cryst. Growth Des.* **2020**, *20*, 7568–7581.

(8) Pfund, L.; Matzger, A. *ACS Comb. Sci.* **2014**, *16*, 309–313.

(9) Tom, R.; Gao, S.; Yang, Y.; Zhao, K.; Bier, I.; Buchanan, E. A.; Zaykov, A.; Havlas, Z.; Michl, J.; Marom, N. Inverse design of tetracene polymorphs with enhanced singlet fission performance by property-based genetic algorithm optimization. *Chemistry of Materials* **2023**, *35*, 1373–1386.

(10) Bhatia, A.; Chopra, S.; Nagpal, K.; Deb, P. K.; Tekade, M.; Tekade, R. K. *Dosage Form Design Parameters*; Elsevier, 2018; pp 31–65.

(11) Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J. Facts and fictions about polymorphism. *Chemical Society Reviews* **2015**, *44*, 8619–8635.

(12) Shi, Q.; Chen, H.; Wang, Y.; Xu, J.; Liu, Z.; Zhang, C. Recent advances in drug polymorphs: Aspects of pharmaceutical properties and selective crystallization. *International journal of pharmaceutics* **2022**, *611*, 121320.

(13) Li, M.; Balawi, A. H.; Leenaers, P. J.; Ning, L.; Heintges, G. H.; Marszalek, T.; Pisula, W.; Wienk, M. M.; Meskers, S. C.; Yi, Y.; others Impact of polymorphism on the optoelectronic properties of a low-bandgap semiconducting polymer. *Nature communications* **2019**, *10*, 2867.

(14) Wang, H.; Xu, S.; Zhao, H.; Wu, W.; Wang, N.; Wang, T.; Huang, X.; Zhou, L.; Bao, Y.; Hao, H. Polymorphism of CL-20 and the Modification and Inhibition Strategies for Its Crystal Transformation. *Crystal Growth & Design* **2024**, *24*, 9266–9296.

(15) Neumann, M. A.; van de Streek, J. How many ritonavir cases are there still out there? *Faraday discussions* **2018**, *211*, 441–458.

(16) Beran, G. J. Modeling polymorphic molecular crystals with electronic structure theory. *Chemical reviews* **2016**, *116*, 5567–5613.

(17) Lehmann, C. Crystal Structure Prediction: Dawn of a New Era. *Angew. Chem. Int. Ed.* **2011**, *50*, 5616–5617.

(18) Day, G. Current approaches to predicting molecular organic crystal structures. *Crystallographr. Rev.* **2011**, *17*, 3–52.

(19) Nyman, J.; Day, G. M. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm* **2015**, *17*, 5154–5165.

(20) Dolgonos, G.; Hoja, J.; Boese, A. Revised values for the X23 benchmark set of molecular crystals. *Phys. Chem. Chem. Phys.* **2019**, *21*, 24333–24344.

(21) Hoja, J.; Tkatchenko, A. First-principles stability ranking of molecular crystal poly-morphs with the DFT+MBD approach. *Faraday Discuss.* **2018**, *211*, 253–274.

(22) Reilly, A.; Tkatchenko, A. Seamless and Accurate Modeling of Organic Molecular Materials. *J. Phys. Chem. Lett.* **2013**, *4*, 1028–1033.

(23) Brandenburg, J. G.; Grimme, S. Dispersion corrected Hartree–Fock and density func-tional theory for organic crystal structure prediction. *Prediction and Calculation of Crystal Structures: Methods and Applications* **2014**, 1–23.

(24) Wang, H.; Zhang, Y.; Zhang, L.; Wang, H. Crystal structure prediction of binary alloys via deep potential. *Frontiers in Chemistry* **2020**, *8*, 589795.

(25) Nikhar, R.; Szalewicz, K. Reliable crystal structure predictions from first principles. *Nature Communications* **2022**, *13*, 3095.

(26) Beran, G. J.; Sugden, I. J.; Greenwell, C.; Bowskill, D. H.; Pantelides, C. C.; Ad-jiman, C. S. How many more polymorphs of ROY remain undiscovered. *Chemical science* **2022**, *13*, 1288–1297.

(27) O'Connor, D.; Bier, I.; Hsieh, Y.-T.; Marom, N. Performance of Dispersion-Inclusive Density Functional Theory Methods for Energetic Materials. *Journal of Chemical Theory and Computation* **2022**, *18*, 4456–4471.

(28) Marom, N.; DiStasio Jr., R. A.; Atalla, V.; Levchenko, S.; Reilly, A. M.; Che-likowsky, J. R.; Leiserowitz, L.; Tkatchenko, A. Many-Body Dispersion Interactions in Molecular Crystal Polymorphism. *Angewandte Chemie International Edition* **2013**, *52*, 6629–6632.

(29) Mayo, R. A.; Price, A. J. A.; Otero-de-la Roza, A.; Johnson, E. R. Assessment of the exchange-hole dipole moment dispersion correction for the energy ranking stage

of the seventh crystal structure prediction blind test. *Acta Crystallographica Section B* **2024**, *80*, 595–605.

(30) Lommerse, J. P.; Motherwell, W. D.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W.; Leusen, F. J.; Mooij, W. T.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; Van Eijck, B. P.; Verwer, P.; Williams, D. E. A test of crystal structure prediction of small organic molecules. *Acta Cryst. B* **2000**, *56*, 697–714.

(31) Motherwell, W. D. et al. Crystal structure prediction of small organic molecules: A second blind test. *Acta Cryst. B* **2002**, *58*, 647–661.

(32) Day, G. M. et al. A third blind test of crystal structure prediction. *Acta Cryst. B* **2005**, *61*, 511–527.

(33) Day, G. M. et al. Significant progress in predicting the crystal structures of small organic molecules - A report on the fourth blind test. *Acta Cryst. B* **2009**, *65*, 107–125.

(34) Bardwell, D. A. et al. Towards crystal structure prediction of complex organic compounds - A report on the fifth blind test. *Acta Cryst. B* **2011**, *67*, 535–551.

(35) Reilly, A. M. et al. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Cryst. B* **2016**, *72*, 439–459.

(36) Hunnisett, L. M.; Nyman, J.; Francia, N.; Abraham, N. S.; Adjiman, C. S.; Aitipamula, S.; Alkhidir, T.; Almehairbi, M.; Anelli, A.; Anstine, D. M.; others The seventh blind test of crystal structure prediction: structure generation methods. *Structural Science* **2024**, *80*.

(37) Hunnisett, L. M.; Francia, N.; Nyman, J.; Abraham, N. S.; Aitipamula, S.; Alkhidir, T.; Almehairbi, M.; Anelli, A.; Anstine, D. M.; Anthony, J. E.; others

The seventh blind test of crystal structure prediction: structure ranking methods. *Structural Science* **2024**, *80*.

(38) Reilly, A. M.; Cooper, R. I.; Adjiman, C. S.; Bhattacharya, S.; Boese, A. D.; Brandenburg, J. G.; Bygrave, P. J.; Bylsma, R.; Campbell, J. E.; Car, R.; others Report on the sixth blind test of organic crystal structure prediction methods. *Structural Science* **2016**, *72*, 439–459.

(39) Sun, H.; Jin, Z.; Yang, C.; Akkermans, R. L.; Robertson, S. H.; Spenley, N. A.; Miller, S.; Todd, S. M. COMPASS II: extended coverage for polymer and drug-like molecule databases. *Journal of molecular modeling* **2016**, *22*, 1–10.

(40) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *Journal of Physical chemistry* **1990**, *94*, 8897–8909.

(41) Garcia, J.; Podeszwa, R.; Szalewicz, K. SAPT codes for calculations of intermolecular interaction energies. *The Journal of Chemical Physics* **2020**, *152*.

(42) Xue, B.; Yang, Q.; Zhang, Q.; Wan, X.; Fang, D.; Lin, X.; Sun, G.; Gobbo, G.; Cao, F.; Mathiowetz, A. M.; others Development and comprehensive benchmark of a high-quality amber-consistent small molecule force field with broad chemical space coverage for molecular modeling and free energy calculation. *Journal of Chemical Theory and Computation* **2023**, *20*, 799–818.

(43) Firaha, D.; Liu, Y. M.; van de Streek, J.; Sasikumar, K.; Dietrich, H.; Helfferich, J.; Aerts, L.; Braun, D. E.; Broo, A.; DiPasquale, A. G.; others Predicting crystal form stability under real-world conditions. *Nature* **2023**, *623*, 324–328.

(44) Wang, C.; Ni, Y.; Zhang, C.; Xue, X. Crystal structure prediction of 2, 4, 6, 8, 10, 12-Hexanitro-2, 4, 6, 8, 10, 12-hexaazaisowurtzitane (CL-20) by a tailor-made OPLS-AA force field. *Crystal Growth & Design* **2021**, *21*, 3037–3046.

(45) Bowskill, D. H.; Sugden, I. J.; Konstantinopoulos, S.; Adjiman, C. S.; Pantelides, C. C. Crystal structure prediction methods for organic molecules: State of the art. *Annual Review of Chemical and Biomolecular Engineering* **2021**, *12*, 593–623.

(46) Neumann, M. A. Tailor-made force fields for crystal-structure prediction. *The Journal of Physical Chemistry B* **2008**, *112*, 9810–9829.

(47) Kendrick, J.; Leusen, F. J.; Neumann, M. A.; Van De Streek, J. Progress in Crystal Structure Prediction. *Chemistry – A European Journal* **2011**, *17*, 10736–10744.

(48) Kendrick, J.; Stephenson, G. A.; Neumann, M. A.; Leusen, F. J. Crystal structure prediction of a flexible molecule of pharmaceutical interest with unusual polymorphic behavior. *Crystal growth & design* **2013**, *13*, 581–589.

(49) Neumann, M.; Van De Streek, J.; Fabbiani, F.; Hidber, P.; Grassmann, O. Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nature Communications* **2015**, *6*, 7793.

(50) Hoja, J.; Ko, H.-Y.; Neumann, M. A.; Car, R.; DiStasio Jr, R. A.; Tkatchenko, A. Reliable and practical computational description of molecular crystal polymorphs. *Science Advances* **2019**, *5*, eaau3338.

(51) Misquitta, A. J.; Welch, G. W.; Stone, A. J.; Price, S. L. A first principles prediction of the crystal structure of C6Br2ClFH2. *Chemical Physics Letters* **2008**, *456*, 105–109.

(52) Welch, G. W.; Karamertzanis, P. G.; Misquitta, A. J.; Stone, A. J.; Price, S. L. Is the induction energy important for modeling organic crystals? *Journal of chemical theory and computation* **2008**, *4*, 522–532.

(53) Beran, G. J.; Nanda, K. Predicting organic crystal lattice energies with chemical accuracy. *The Journal of Physical Chemistry Letters* **2010**, *1*, 3480–3487.

(54) Yoo, D.; Lee, K.; Jeong, W.; Lee, D.; Watanabe, S.; Han, S. Atomic energy mapping of neural network potential. *Physical Review Materials* **2019**, *3*, 093802.

(55) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine learning interatomic potentials as emerging tools for materials science. *Advanced Materials* **2019**, *31*, 1902765.

(56) Mishin, Y. Machine-learning interatomic potentials for materials science. *Acta Materialia* **2021**, *214*, 116980.

(57) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schutt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine learning force fields. *Chemical Reviews* **2021**, *121*, 10142–10186.

(58) Manzhos, S.; Carrington Jr, T. Neural network potential energy surfaces for small molecules and reactions. *Chemical Reviews* **2020**, *121*, 10187–10217.

(59) Unke, O. T.; Meuwly, M. Toolkit for the construction of reproducing kernel-based representations of data: Application to multidimensional potential energy surfaces. *Journal of chemical information and modeling* **2017**, *57*, 1923–1931.

(60) Jiang, B.; Li, J.; Guo, H. Potential energy surfaces from high fidelity fitting of ab initio points: the permutation invariant polynomial-neural network approach. *International Reviews in Physical Chemistry* **2016**, *35*, 479–506.

(61) Zubatyuk, R.; Smith, J. S.; Leszczynski, J.; Isayev, O. Accurate and transferable multitask prediction of chemical properties with an atoms-in-molecules neural network. *Science advances* **2019**, *5*, eaav6490.

(62) Anstine, D.; Zubatyuk, R.; Isayev, O. AIMNet2: A Neural Network Potential to Meet your Neutral, Charged, Organic, and Elemental-Organic Needs. **2023**,

(63) Kovács, D. P.; Moore, J. H.; Browning, N. J.; Batatia, I.; Horton, J. T.; Kapil, V.;

Magdău, I.-B.; Cole, D. J.; Csányi, G. MACE-OFF23: Transferable machine learning force fields for organic molecules. *arXiv preprint arXiv:2312.15211* **2023**,

(64) Podryabinkin, E. V.; Tikhonov, E. V.; Shapeev, A. V.; Oganov, A. R. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Physical Review B* **2019**, *99*, 064114.

(65) Butler, P. W.; Hafizi, R.; Day, G. M. Machine-learned potentials by active learning from organic crystal structure prediction landscapes. *The Journal of Physical Chemistry A* **2024**, *128*, 945–957.

(66) Žugec, I.; Geilhufe, R. M.; Lončarić, I. Global machine learning potentials for molecular crystals. *The Journal of chemical physics* **2024**, *160*.

(67) Taylor, C. R.; Butler, P. W.; Day, G. M. Predictive crystallography at scale: mapping, validating, and learning from 1000 crystal energy landscapes. *Faraday Discussions* **2025**, *256*, 434–458.

(68) Ryan, K.; Lengyel, J.; Shatruk, M. Crystal structure prediction via deep learning. *Journal of the American Chemical Society* **2018**, *140*, 10158–10168.

(69) Kadan, A.; Ryczko, K.; Wildman, A.; Wang, R.; Roitberg, A.; Yamazaki, T. Accelerated organic crystal structure prediction with genetic algorithms and machine learning. *Journal of Chemical Theory and Computation* **2023**, *19*, 9388–9402.

(70) Shapera, E. P.; Bučar, D.-K.; Prasankumar, R. P.; Heil, C. Machine learning assisted prediction of organic salt structure properties. *npj Computational Materials* **2024**, *10*, 176.

(71) Ye, Z.; Wang, N.; Zhou, J.; Ouyang, D. Organic crystal structure prediction via coupled generative adversarial networks and graph convolutional networks. *The Innovation* **2024**, *5*.

(72) Zhou, D.; Bier, I.; Santra, B.; Jacobson, L. D.; Wu, C.; Garaizar Suarez, A.; Almaguer, B. R.; Yu, H.; Abel, R.; Friesner, R. A.; others A robust crystal structure prediction method to support small molecule drug development with large scale validation and blind study. *Nature Communications* **2025**, *16*, 2210.

(73) Kilgour, M.; Rogal, J.; Tuckerman, M. Geometric deep learning for molecular crystal structure prediction. *Journal of chemical theory and computation* **2023**,

(74) Galanakis, N.; Tuckerman, M. E. Rapid prediction of molecular crystal structures using simple topological and physical descriptors. *Nature communications* **2024**, *15*, 9757.

(75) Kuroshima, D.; Kilgour, M.; Tuckerman, M. E.; Rogal, J. Machine Learning Classification of Local Environments in Molecular Crystals. *Journal of Chemical Theory and Computation* **2024**, *20*, 6197–6206, PMID: 38959410.

(76) Bier, I.; Marom, N. Machine learned model for solid form volume estimation based on packing-accessible surface and molecular topological fragments. *The Journal of Physical Chemistry A* **2020**, *124*, 10330–10345.

(77) Price, S. L. Why don't we find more polymorphs? *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2013**, *69*, 313–328.

(78) Hoja, J.; Tkatchenko, A. First-principles stability ranking of molecular crystal polymorphs with the DFT+ MBD approach. *Faraday discussions* **2018**, *211*, 253–274.

(79) Dolgonos, G. A.; Hoja, J.; Boese, A. D. Revised values for the X23 benchmark set of molecular crystals. *Physical Chemistry Chemical Physics* **2019**, *21*, 24333–24344.

(80) O'Connor, D.; Bier, I.; Tom, R.; Hiszpanski, A. M.; Steele, B. A.; Marom, N. Ab Initio Crystal Structure Prediction of the Energetic Materials LLM-105, RDX, and HMX. *Crystal Growth & Design* **2023**, *23*, 6275–6289.

44

(81) Tom, R.; Rose, T.; Bier, I.; O'Brien, H.; Vázquez-Mayagoitia, Á.; Marom, N. Genarris 2.0: A random structure generator for molecular crystals. *Computer Physics Communications* **2020**, *250*, 107170.

(82) Hawkins, P. C.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *Journal of chemical information and modeling* **2010**, *50*, 572–584.

(83) Riniker, S.; Landrum, G. A. Better informed distance geometry: using what we know to improve conformation generation. *Journal of chemical information and modeling* **2015**, *55*, 2562–2574.

(84) RDKit: Open-Source Cheminformatics Software, version 2021.09.4. https://www.rdkit.org.

(85) Meunier, M.; Robertson, S. Materials studio 20th anniversary. 2021.

(86) O'Boyle, N. M.; Vandermeersch, T.; Flynn, C. J.; Maguire, A. R.; Hutchison, G. R. Confab-Systematic generation of diverse low-energy conformers. *Journal of cheminformatics* **2011**, *3*, 1–9.

(87) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge structural database. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72*, 171–179.

(88) Spicher, S.; Grimme, S. Robust atomistic modeling of materials, organometallic, and biochemical systems. *Angewandte Chemie International Edition* **2020**, *59*, 15665–15673.

(89) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian

equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics* **1977**, *23*, 327–341.

(90) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; others Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* **2019**, *32*.

(91) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of chemical theory and computation* **2019**, *15*, 1652–1671.

(92) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *The Journal of chemical physics* **2017**, *147*.

(93) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der physik* **1921**, *369*, 253–287.

(94) Otero-De-La-Roza, A.; Johnson, E. R. A benchmark for non-covalent interactions in solids. *The Journal of chemical physics* **2012**, *137*.

(95) Togo, A.; Tanaka, I. First principles phonon calculations in materials science. *Scripta Materialia* **2015**, *108*, 1–5.

(96) Kantorovich, L. Thermoelastic properties of perfect crystals with nonprimitive lattices. I. General theory. *Physical Review B* **1995**, *51*, 3520.

(97) Birch, F. Finite elastic strain of cubic crystals. *Physical review* **1947**, *71*, 809.

(98) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *The Journal of chemical physics* **2015**, *142*.

(99) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.

(100) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Physical chemistry chemical physics* **2006**, *8*, 1057–1065.

(101) Neese, F. Software update: The ORCA program system—Version 5.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2022**, *12*, e1606.

(102) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications* **2009**, *180*, 2175–2196.

(103) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximations Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(104) Tkatchenko, A.; DiStasio, R.; Car, R.; Scheffler, M. Accurate and Efficient Method for Many-Body van der Waals Interactions. *Phys. Rev. Lett.* **2012**, *108*, 236402.

(105) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(106) Anthony, J. E. Functionalized acenes and heteroacenes for organic electronics. *Chemical reviews* **2006**, *106*, 5028–5048.

(107) Sorli, J. C.; Ai, Q.; Granger, D. B.; Gu, K.; Parkin, S.; Jarolimek, K.; Telesz, N.; Anthony, J. E.; Risko, C.; Loo, Y.-L. Impact of Atomistic Substitution on Thin-Film Structure and Charge Transport in a Germanyl-ethynyl Functionalized Pentacene. *Chemistry of Materials* **2019**, *31*, 6615–6623.

(108) Togo, A.; Shinohara, K.; Tanaka, I. Spglib: a software library for crystal symmetry search. *Science and Technology of Advanced Materials: Methods* **2024**, 2384822.

(109) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science* **2013**, *68*, 314–319.

(110) Zhao, C.; Chen, L.; Che, Y.; Pang, Z.; Wu, X.; Lu, Y.; Liu, H.; Day, G. M.; Cooper, A. I. Digital navigation of energy-structure-function maps for hydrogen-bonded porous molecular crystals. *Nature Communications* **2021**, *12*, 817.

(111) Francia, N. F.; Price, L. S.; Nyman, J.; Price, S. L.; Salvalaglio, M. Systematic Finite-Temperature Reduction of Crystal Energy Landscapes. *Crystal Growth & Design* **2020**, *20*, 6847–6862.

(112) Heit, Y. N.; Beran, G. J. O. How important is thermal expansion for predicting molecular crystal structures and thermochemistry at finite temperatures? *Acta Cryst. Sect. B* **2016**, *72*, 514–529.