MolCrystalFlow: Molecular Crystal Structure Prediction via Flow Matching

Cheng Zeng[1,2*], Harry W. Sullivan[3], Thomas Egg[5], Maya M. Martirossyan[5], Philipp Höllmer[5,6], Jirui Jin[1,2], Adrian Roitberg[1,2], Richard G. Hennig[2,4], Stefano Martiniani[5,6,7,8], Ellad B. Tadmor[9], and Mingjie Liu[1,2*]

[1]Department of Chemistry, University of Florida, Gainesville, FL 32611, USA
[2]Quantum Theory Project, University of Florida, Gainesville, FL 32611, USA
[3]Department of Chemical Engineering and Material Science, University of Minnesota - Twin Cities, Minneapolis, MN 55455, USA
[4]Department of Materials Science & Engineering, University of Florida, Gainesville, FL 32611, USA
[5]Center for Soft Matter Research, Department of Physics, New York University, New York 10003, USA
[6]Simons Center for Computational Physical Chemistry, Department of Chemistry, New York University, New York 10003, USA
[7]Courant Institute of Mathematical Sciences, New York University, New York 10003, USA
[8]Center for Neural Science, New York University, New York 10003, USA
[9]Department of Aerospace Engineering and Mechanics, University of Minnesota - Twin Cities, Minneapolis, MN 55455, USA

March 2, 2026

Abstract

Molecular crystal structure prediction represents a grand challenge in computational chemistry due to large sizes of constituent molecules and complex intra- and intermolecular interactions. While generative modeling has revolutionized structure discovery for molecules, inorganic solids, and metal–organic frameworks, extending such approaches to fully periodic molecular crystals is still elusive. Here, we present MolCrystalFlow, a flow-based generative model for molecular crystal structure prediction. The framework disentangles intramolecular complexity from intermolecular packing by embedding molecules as rigid bodies and jointly learning the lattice matrix, molecular orientations, and centroid positions. Centroids and orientations are represented on their native Riemannian manifolds, allowing geodesic flow construction and graph neural network operations that respects geometric symmetries. We benchmark our model against state-of-the-art generative models for large-size periodic crystals and rule-based structure generation methods on two open-source molecular crystal datasets. We demonstrate an integration of MolCrystalFlow model with universal machine learning potential to accelerate molecular crystal structure prediction, paving the way for data-driven generative discovery of molecular crystals.

Introduction

Molecular crystal structure prediction (CSP) represents a fundamental challenge in computational chemistry. Unlike inorganic crystals, molecular crystals are governed by an interplay of intramolecular and intermolecular interactions along with periodic lattice constraints, leading to complex energy landscapes with many competing low-energy minima. This complexity is also known as polymorphism, where a single molecule can adopt multiple crystal packings that can be thermodynamically accessible. Polymorphism has major implications for drug efficacy, stability, and manufacturability [1, 2]. A notable example is Ritonavir, an antiretroviral drug initially released in a crystal form known as Form A. Years after commercialization, a previously unknown Form B appeared during routine laboratory analysis [3]. Although chemically identical, the two forms differ subtly in molecular packing, leading to drastically

*c.zeng@ufl.edu, mingjieliu@ufl.edu

different solubility. Form B dissolves poorly in common solvents such as water and methanol, causing drug precipitation and a sharp drop in bioavailability. Its unexpected emergence forced product withdrawals and substantial reformulation costs, demonstrating how failing to anticipate a stable polymorph can have serious clinical and economic consequences.

Predicting stable polymorphs is essential, as applications in optoelectronics [4], pharmaceuticals [5], organic semiconductors [6], and energy storage materials [7] are highly sensitive to slight variations in molecular packing, where small structural differences can result in significant changes in stability and performance. Computational approaches have thus become indispensable for polymorph discovery, as experimental screening is slow and resource-intensive. However, predicting all possible crystal structures for a molecule remains a long-standing challenge in computational chemistry and materials science [8]. Current CSP workflows rely on stochastic or evolutionary searches to generate candidate structures, followed by large-scale lattice-energy evaluation [9]. These methods often demand millions of CPU hours for a single compound and are typically system-specific, limiting generalizability across chemical families [10]. Machine-learning-based models such as AIMNet2 [10] and FastCSP [11] accelerate energy prediction but primarily help with ranking; the underlying search over possible packings remains exhaustive and computationally intensive. Moreover, this prevailing "generate-and-rank" paradigm—relying on brute-force enumeration through stochastic or evolutionary sampling—produces large pools of candidate structures that must subsequently be deduplicated, clustered, and filtered [12, 13] before stability rankings are computed. These ranking, typically obtained by DFT-based screening [11, 8] or machine learning models, continue to face a fundamental bottleneck arising from the sheer size of the search space. More efficient and generalizable strategies are needed to guide exploration of the polymorph landscape without exhaustive sampling.

Generative modeling is a promising alternative approach to the brute-force approach using random structure search. It learns a mapping between an easy-to-sample base distribution and a target distribution [14, 15, 16, 17, 18, 19], often using neural networks that respect geometric symmetry constraints. A generative model trained on known stable polymorphs might directly propose plausible crystal packings, overcoming kinetic barriers that are normally inaccessible in conventional atomistic simulations. Recent advances in generative AI have demonstrated potential in accelerating the prediction and discovery of small molecules [20, 21, 22], protein sequences [23], and inorganic crystals [24, 25] using diffusion models and flow-based generative processes. However, extending these methods to fully atomistic, periodic molecular crystals is non-trivial. Single molecules lack the periodic constraints intrinsic to molecular crystals, whereas all-atom inorganic crystal models suffer from scaling up. For instance, given the compositions and numbers of atoms in the crystal, our recent model OMatG accurately matches $\sim$70% of the ground-truth crystal structures with up to 20 atoms but drops to only 27.4% at sizes near 50 atoms [24]. This decline reflects the difficulty of learning long-range, many-body interactions, enforcing periodic symmetry, and navigating the vast configuration spaces inherent to larger crystalline systems.

To address complexity in larger systems, recent cross-scale approaches, such as MOFFlow for metal-organic frameworks (MOFs) [26] and AssembleFlow for molecular clusters [27], represent materials as assemblies of rigid-body building blocks and learn their spatial arrangements. These models are promising for framework materials and supramolecular architectures, yet they are not directly applicable to molecular crystals because molecular clusters modeled in AssembleFlow do not naturally exhibit periodic lattices, and MOFFlow does not enforce periodic translational invariance. Generative models for molecular crystal structure prediction are just emerging—Jin et al. [28] convert the periodic molecular crystal structure prediction challenge to learning the packing for non-periodic molecular clusters with explicit long-range inter-molecular interactions. All-atom coordinates are then generated using non-equivariant Alphafold3-style transformer backbones. However, energies cannot be directly evaluated on the generated clusters without post hoc lattice inference, limiting the utility for downstream applications. Generative modeling for molecular crystals remains an open frontier. In addition, challenges remain for large structure sizes and complex energy landscapes, and no existing generative model can produce stable and diverse molecular crystal polymorphs with periodic lattice constraints. This work aims to close this gap by developing a generative framework designed for molecular crystal structure prediction. Such an approach will enable more efficient exploration of polymorph landscapes and accelerate the discovery of molecular crystals.

We introduce MolCrystalFlow, a periodic E(3)-invariant flow-based generative model that predicts the crystal packing with explicit lattice periodicity. MolCrystalFlow disentangles the intramolecular and intermolecular complexity by conditioning the generative process on given molecular conformers that are treated under rigid-body approximation. It then generates all modalities of molecular crystals needed to recover the full structures. Flow matching of molecular crystals operates on the corresponding Riemannian manifolds. We benchmark MolCrystalFlow on two publicly available datasets, including a small-dataset of 11.5K structures derived from CSD experimental data [29] and a subset of the largest open-source OMC25 dataset [30]. We demonstrate superior performance of MolCrystalFlow over MOFFlow [26], the state-of-the-art hierarchical flow-based generative models for large-size periodic crystalline materials. We integrate MolCrystalFlow with universal machine learning interatomic potentials (u-MLIP) and density functional theory (DFT) calculations to perform molecular crystal structure predictions on three open CSP competition

targets. The combined pipeline identifies polymorphs for two target compounds with energies and crystal packing geometries close to experimental structures. We anticipate that the model and pipeline developed in this work, together with the open data, will enable new research avenues toward end-to-end generative discovery of molecular crystals.

Results

Overview of MolCrystalFlow

The overarching goal of molecular crystal structure prediction is to identify low-energy polymorphs on the free-energy landscape (Figure 1a). Given a molecular graph as input, the task is to predict its crystal packing. This problem is typically decomposed into two stages. First, a stable molecular conformer* needs to be identified based on the corresponding 2D topological graph, a task for which tools such as OpenEye OMEGA [31] and RDKit ETKDGv2 [32] provide robust solutions. The subsequent and substantially more challenging step is to predict the crystal packing of these molecular conformers. Finally, candidate polymorphs are ranked according to their thermodynamic stability, commonly evaluated using DFT-based calculations.

MolCrystalFlow is designed to reduce the search cost and improve the quality of crystal packing by directly sampling physically plausible polymorphs for a given molecular conformer. To accelerate this search, each molecule within the crystal is represented as a rigid-body conformer, which is a reasonable approximation in many cases [33, 28]. Under this assumption, a crystal packing configuration is fully specified by the lattice matrix, along with the centroids and rotational orientations of all molecular rigid bodies (Figure 1b). The centroid of each molecule is defined as its center of position, while the rotational orientation is parameterized by unit principal axes obtained via principal component analysis (PCA), followed by normalization to ensure consistent and unique local coordinate frames across molecules.

Building on this representation, we adopt a hierarchical modeling framework for molecular crystal structure prediction (Figure 1c). Similar approaches have been demonstrated in previous generative models for large crystalline MOFs [26] and molecular clusters [27]. In the first stage, each molecular building block is embedded using an equivariant graph neural network (EGNN) based on the E(n)-equivariant architecture introduced by Hoogeboom et al. [20]. The final node features are aggregated to produce an invariant embedding that characterizes the molecular building block. Details of this building-block embedder are provided in Figure 2a. Some molecular information—such as size or functional-group complexity—may be attenuated when projected into a neural network latent space. As shown in Figure S2, the latent embedding derived from a trained EGNN model alone exhibits weak correlation with the number of atoms in the building block. To enrich the molecular representation, we concatenate 18 auxiliary molecular descriptors to the invariant embedding before passing it to the second stage. The full list of these auxiliary features is described in the 'Hierarchical Representation of Molecular Crystals' section.

The second stage consists of a joint flow-matching generative process parameterized by an E(3) periodic-invariant neural network, which generates the lattice matrix, molecular centroid positions, and molecular orientations (Figure 2b). Flow matching learns conditional velocity fields that transport samples from a simple base distribution to the target data distribution [14, 16]. By employing a joint flow formulation, distinct interpolants are defined for each modality. The generative process gradually refines each modality until a final structure is generated. To construct flows on non-trivial geometries, each modality is defined on its intrinsic Riemannian manifold [34]. Specifically, molecular centroids are represented using fractional coordinates to naturally respect periodic translational invariance, corresponding to a three-dimensional torus. Molecular orientations evolve on the $\mathrm{SO}(3)$ manifold. We adopt a data-informed base distribution for the lattice matrix and use a simple linear interpolation, which we find to be both stable and effective in practice. During inference, the generative model simultaneously evolves the lattice parameters, centroid positions, and molecular orientations by integrating the learned velocity fields (Figure 1d). By combining the generated lattice matrix, centroid positions and rotational orientations with the internal fixed coordinates of molecular building blocks, MolCrystalFlow reconstructs all-atom molecular crystal structures suitable for direct optimization and energy evaluation. Details of choices of base distributions and joint flow matching are included in the 'Base Distributions and Optimal Transport' and 'Joint Flow Matching on Riemannian manifolds' sections.

Neural Network Architecture

As a rigid body, the representation of each individual molecular building block must be invariant to global translations and rotations. Accordingly, the embedding is designed to capture the intrinsic chemical and geometric structure of the molecule, independent of its absolute position or orientation in space. The global placement of molecules—namely, their

*A *molecular conformer* is a specific, 3D spatial arrangement of a molecule's atoms corresponding to a local minimum on the molecule's potential energy surface.

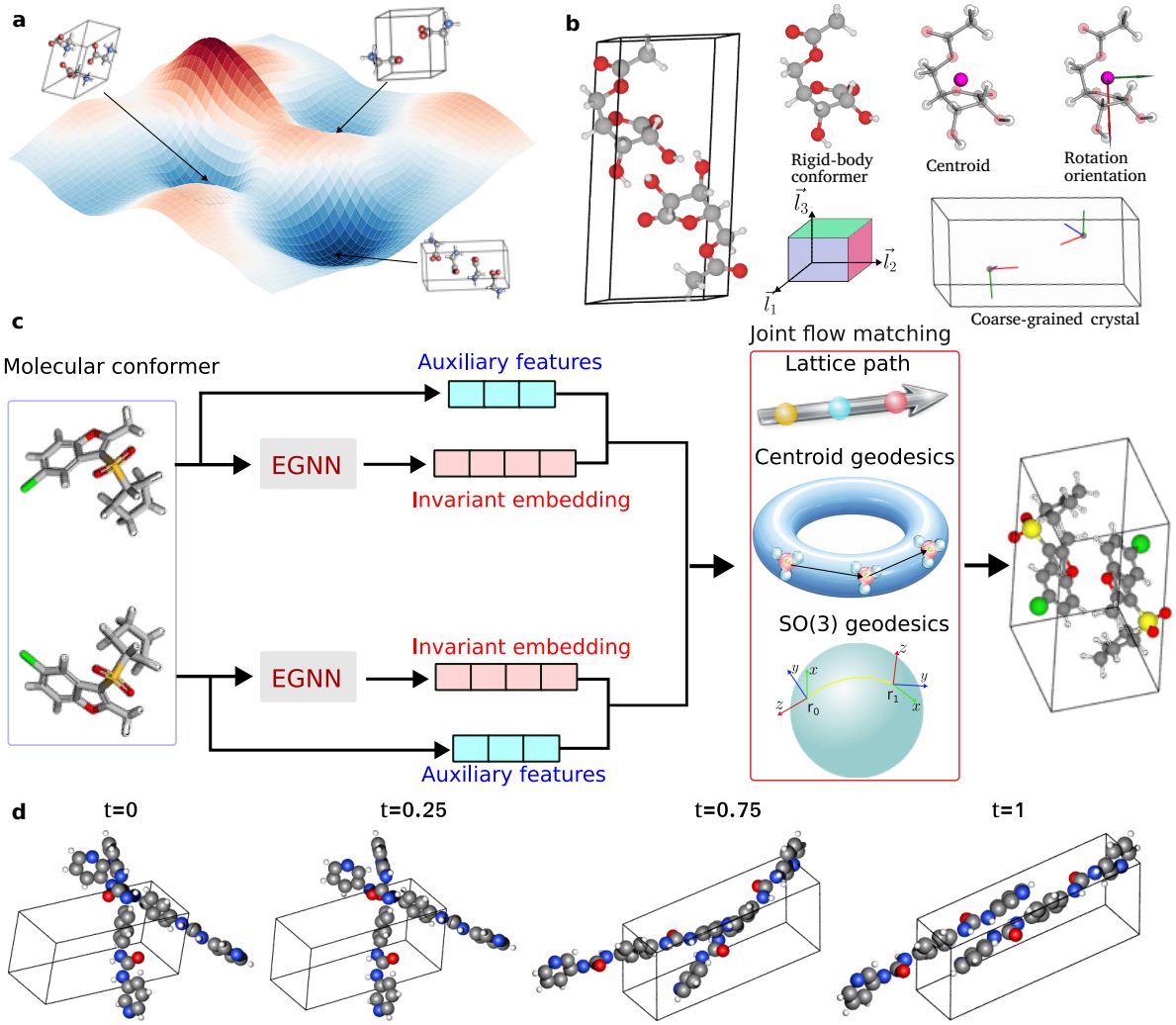

Figure 1: **Overview of the MolCrystalFlow framework.** a, Schematic illustration of a polymorph energy landscape in molecular crystals. The topology, barrier heights, and relative basin depths are not quantitative and are shown for conceptual purposes only. b, Representation of a molecular crystal structure. A molecular crystal is decomposed to a lattice matrix spanning by three vectors $\overrightarrow{l_1}$, $\overrightarrow{l_2}$ and $\overrightarrow{l_3}$, and a constituent molecular conformer building block which can be characterized by a rigid-body centroid and rotational orientation given by three eigenvectors derived from a principal component analysis over the atomic positions of the building block. Hence a coarse-grained crystal includes all the modalities to be generated. c, MolCrystalFlow works by encoding each building block in the crystal (for example two molecules in the schematic) with an equivariant graph neural network (EGNN) that produces an invariant embedding. The invariant embedding is augmented with auxiliary features to form the input for the joint flow-matching generative process that simultaneously produce the lattice matrix using a linear interpolant for each lattice matrix entry, centroid fractional coordinates on a torus geodesic, and rotational orientation on an $SO(3)$ geodesic. Together, the three learned modalities will reconstruct the full atomic structure of the molecular crystal. d. An example generation trajectory with four time steps, showing the co-evolution of centroid positions, molecular orientations and lattice.

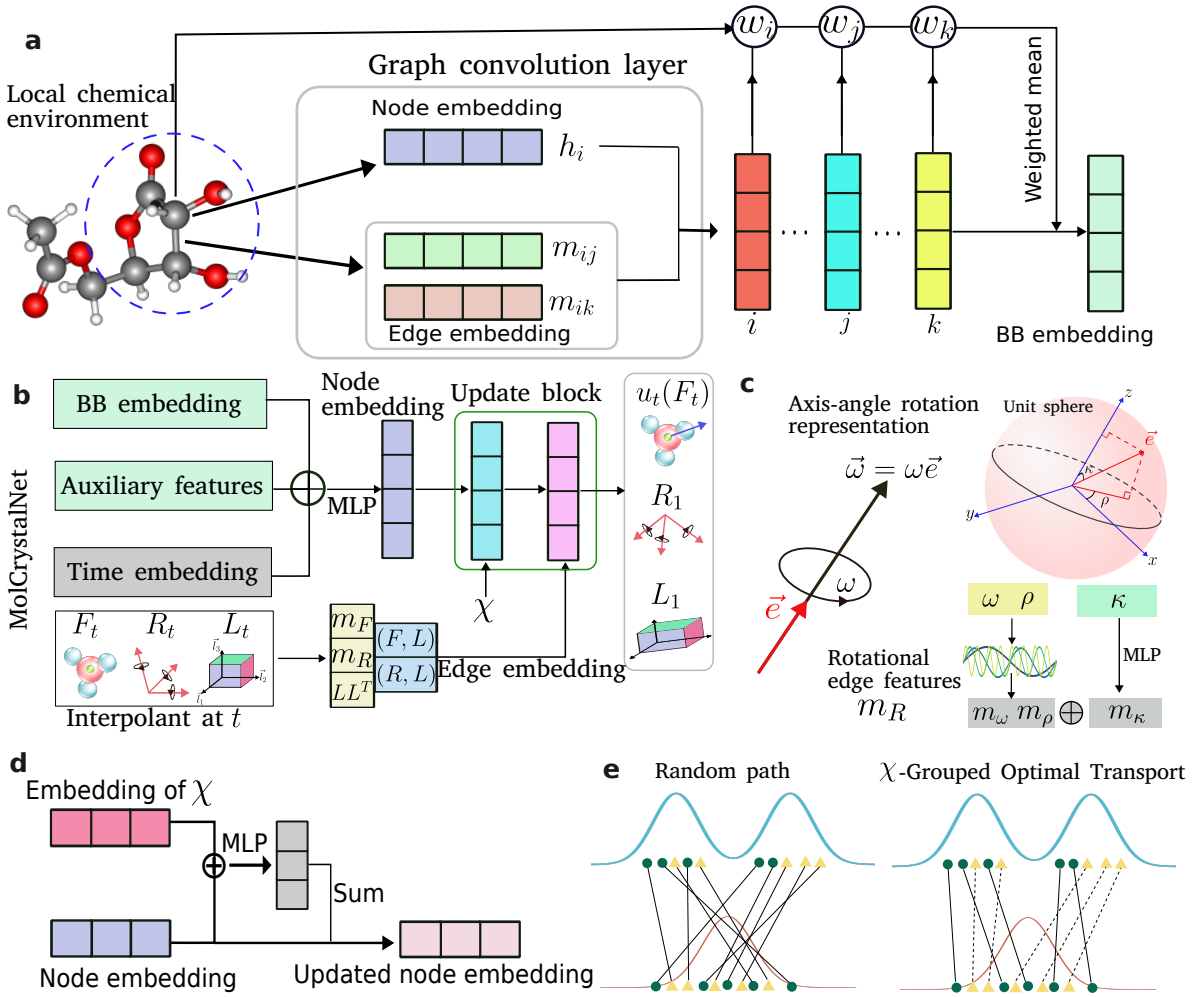

Figure 2: **Neural network architecture of MolCrystalFlow.** a, Molecular building block (BB) embedding network with EGNN. An atomic node embedding is constructed by the atom interacting with neighbors in the local chemical environment through multiple graph convolution layers. The final building block embedding is a weighted atomic embedding where each weight is learned and comes from the node embedding interacting with invariant atomic distance to the centroid. b, MolCrystalNet parameterizes the joint flow matching for lattice matrix L, centroid fractional coordinates (F), and rotational orientation (R). Given interpolants at time t, a concatenation of BB embedding, auxiliary features and time embedding, followed by a multilayer perceptron (MLP), gives rise to the initial node embedding for each molecular building block. In the update block, the initial embedding interacts with a χ embedding for each building block. Message passing relies on the relative positions and orientations between building blocks, and current lattice matrix at time t. The final output is the velocity field for centroid fractional coordinates ($u_t(F_t)$), the denoised rotational orientation (R_1) and lattice matrix (L_1). Interactions of lattice with fractional coordinates (F, L) and orientations (R, L) are also included in the message passing. c, For rotation-dedicated message passing (m_R), we use the axis-angle representation of relative rotational orientation between two building blocks. We embed the rotational angle ω and azimuthal angle ρ using Fourier-series like trigonometric functions while using an MLP to create the embedding for the inclination angle κ. d, Concatenation-summation operation to strengthen the signal of each building block's χ. e, χ-grouped optimal transport to reduce cross-links between paths of different building blocks and facilitate inference.

centroid positions and rigid-body orientations within the crystal lattice—is learned in a subsequent stage of the model. To obtain an E(3)-invariant molecular representation, we employ an EGNN (Figure 2a). Each atom i is initialized with an atom-type embedding and its Cartesian coordinates. Node embeddings h_i are updated through message passing m_{ij} between atoms within a cutoff radius, which depend on interatomic distances. Therefore, the resulting node embeddings are invariant to global translations and rotations. The molecular building-block (BB) embedding is then computed as a learnable weighted mean of the final node embeddings, where the weights w_i are predicted by a shallow MLP using the final node features and each atom's distance to the molecular centroid. Details of the BB embedding network are provided in the 'Hierarchical Representation of Molecular Crystals' section.

Given the molecular building-block embeddings, the remaining task is to generate the crystal lattice matrix, fractional centroid positions, and rigid-body orientations. To this end, we employ a periodic E(3)-invariant graph neural network adapted from DiffCSP [35] that explicitly respects the physical symmetries governing molecular crystal packing, including periodic translational invariance, equivariance of lattice vectors, and equivariance of molecular orientations under global rotations (Figure 2b). Geometric symmetries of molecular crystals are discussed in the 'Geometric Symmetry of Molecular Crystal Structures' section. By construction, the network also satisfies permutation equivariance with respect to molecules in the unit cell. The initial node embedding for each building block is formed by concatenating its invariant BB embedding, auxiliary features, and a time embedding [36, 37], followed by projection through a shallow MLP. In the flow-matching framework, the objective is to learn the velocity field $u_\theta(x_t, t)$ associated with an interpolant state x_t at time t. The interpolants for all structural modalities are described in the 'Joint Flow Matching on Riemannian manifolds' section. Rather than directly predicting velocity fields for all variables, we adopt a denoising parameterization for selected modalities, as intermediate states are less critical than the final configuration [38]. Specifically, the loss functions are reparameterized to predict denoised targets for the lattice matrix (L_1) and rotational orientations (R_1). For fractional coordinates, we retain a velocity-field objective ($u_t(F_t)$) to avoid ambiguities arising from periodic wrapping when comparing denoised targets.

Message passing operates on the current states of molecular building blocks. For fractional coordinates, the message m_F is constructed using a Fourier embedding of relative fractional coordinate differences, ensuring periodic translational invariance of centroid positions [35]. To model rigid-body orientations—an essential modality absent in inorganic crystals—we introduce a dedicated rotational message passing term m_R. The rotational message depends on relative orientations between building blocks. Each relative rotation matrix is converted to an axis–angle representation $\vec{\omega}$, where the direction vector $\vec{e}$ specifies the rotation axis and the magnitude ω denotes the rotation angle (Figure 2c). Since $\vec{e}$ has two degrees of freedom, it is parameterized using polar coordinates with an inclination angle κ and an azimuthal angle ρ. Notably, ω and ρ are periodic with modulo 2π, and are therefore embedded using Fourier features, while κ is embedded via an MLP. The resulting embeddings are combined to form the rotational edge feature m_R. In addition, we incorporate interactions of the lattice matrix with fractional coordinate (F, L) and orientation differences (R, L), which represent the product of lattice Gram matrix with the relative fractional coordinates and rotation matrices, following prior work in MatterGen [25] and FlowMM [39]. Despite the two-stage encoding of molecular and crystal-level information, the building-block embedder network is trained jointly with MolCrystalNet. Details of the periodic E(3)-invariant flow matching and message passing can be found in 'Joint Flow Matching on Riemannian manifolds' section.

Molecular crystals may adopt packings in which molecules share the same or opposite axis-flip states (χ). This feature χ arises naturally to account for the degeneracy of local reference frames obtained from principal component analyses. The formal definition of χ within the equivariant local frame construction is provided in 'Extracting the axis-flip state χ' section. To account for this additional geometric degree of freedom for each building block, we introduce an element embedding of χ at each update block, which is fused into the node embedding via a concatenate_sum operation (Figure 2c,d). To mitigate cross-links between flow paths of different building blocks within a crystal, we employ χ-grouped SE(3) optimal transport, which is detailed in the 'Base Distributions and Optimal Transport' section. This design reflects the fact that this axis-flip state is an intrinsic property of each building block and cannot be altered through rotational evolution during the generative process (Figure 2e).

Accurate and fast generation of densely packed molecular crystals

We trained MolCrystalFlow on an open-source molecular crystal dataset curated by Thürlemann from CSD [29]. The dataset comprises 11,488 structures, split to 10,000 training, 738 validation, and 750 test structures. Further details on data curation and preprocessing are provided in the 'Open-Source Molecular Crystal Dataset' section. We benchmark MolCrystalFlow against the state-of-the-art generative model MOFFlow designed for large-size periodic crystalline materials under rigid-body approximation such as metal-organic frameworks [26], as well as the rule-based Genarris-3 method [40]. We do not include fully all-atom generative models designed for periodic inorganic crystals, as these methods struggle to generate physically reasonable structure large crystalline materials [26]. We also do not compare to Oxtal [28], which reformulates crystal generation as finite cluster prediction without explicit lattices and is trained on

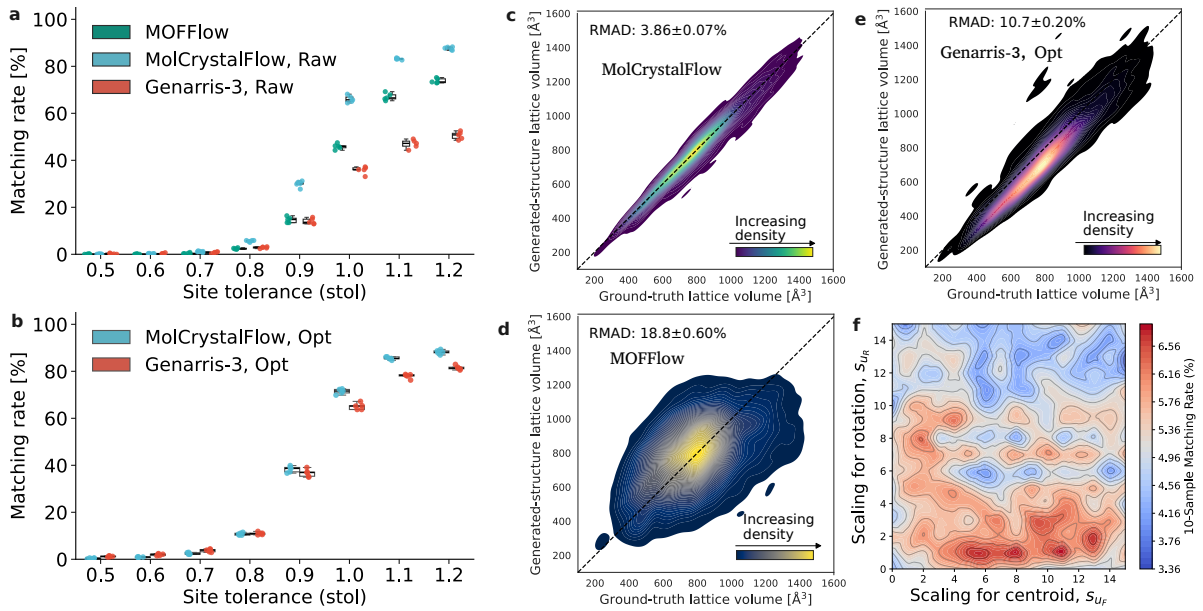

Figure 3: **Performance of MolCrystalFlow against MOFFlow and Genarris-3 baselines.** a, Comparison of MolCrystalFlow with MOFFlow and Genarris-3 over 10-sample matching rates evaluated for directly generated structures without any optimization under different levels of site tolerances. b, Comparison of MolCrystalFlow with Genarris-3 over 10-sample matching rates evaluated for generated structures after a rigid-body optimization under different levels of site tolerances. Regarding the box plots, the median is shown as a solid line. The edges of the box correspond to the first and third quartiles, and the whiskers extend to values within 1.5 × interquartile ranges. c–e, Lattice volume deviation of generated structures for MolCrystalFlow (c), MOFFlow (d), and Genarris-3 with rigid-press optimization (e). A kernel density estimation is used to show data distributions for 10 samples. Relative mean absolute deviation (RMAD) is reported as mean and standard deviation across 10 samples. d, Grid search for the optimal velocity annealing scalings for centroid (s_{u_F}) and rotation (s_{u_R}) using 10 samples and stol=0.8. For figure a and b, five independent runs for 10 samples are generated the statistics and the results for each run are overlaid on the box plots.

the proprietary full CSD molecular crystal data. Cluster-based and lattice-explicit models represent complementary formulations of molecular crystal generation. The former emphasizes local packing motifs, while the latter enforces global periodicity. In this work, we focus on the periodic formulation with an equivariant graph neural network. This choice ensures that both the model and dataset remain fully open for community development, enables the direct application of standard crystallographic metrics for periodic structures, and allows structure optimization and energy ranking to be performed without requiring any post hoc lattice inference. We first assess prediction quality using the structure matching rate, a metric widely adopted in the inorganic crystal community [41]. In addition, we analyze lattice volume deviations between generated and reference structures, which serve as an important indicator of whether densely packed crystals have been identified.

Unlike MolCrystalFlow, MOFFlow operates directly on Cartesian coordinates and generates lattice parameters without enforcing periodic translational invariance. We retrained MOFFlow on our molecular crystal datasets. Both MolCrystalFlow and MOFFlow use 50 integration time steps, which is the default setting in MOFFlow [26] and which we find to be robust for MolCrystalFlow (Figure S3). MolCrystalFlow and MOFFlow generate structures in a one-shot manner, whereas Genarris-3 adopts a multi-sample strategy, producing a large candidate pool followed by selection and deduplication to enhance structural diversity. Additionally, MolCrystalFlow and MOFFlow do not impose explicit space-group constraints, while Genarris-3 conditions structure generation on informative space-group symmetries derived from the molecular point group and the number of molecules per unit cell. For Genarris-3, we report both the directly generated structures ('Raw') and those after the "Rigid Press" optimization ('Opt'), where a hard-sphere potential is applied to refine molecular positions and lattice parameters under space-group constraints. To enable a fair comparison with these optimized structures, we additionally perform rigid-body optimization with fixed lattices on MolCrystalFlow generated structures using the UMA-OMC model [42, 11] . Details of Genarris-3 structure generation are provided in 'Computational settings' section.

We first compare structure matching rates across site tolerances ranging from 0.5 to 1.2. Site tolerance is the maximum allowed positional difference between atoms when deciding whether two crystal structures are considered equivalent [41].

For each test compound, 10 samples are generated, and a prediction is deemed successful if at least one sample matches the corresponding ground-truth structure. This multi-sample evaluation is well motivated, as molecular crystals frequently exhibit polymorphism. MolCrystalFlow substantially outperforms MOFFlow and Genarris-3 for directly generated structures across all site tolerances (Figure 3a). For optimized structures, Genarris-3 achieves higher matching rates under stricter site tolerances, whereas MolCrystalFlow performs better at looser tolerances (Figure 3b). This trend persists when increasing the number of samples up to 200 (Figure S4b,c). Beyond structure matching, MolCrystalFlow also demonstrates accurate lattice volume prediction. Across lattice volumes in the range between 200 and 1600 Å^3, the relative mean absolute deviation (RMAD) for all structures of ten samples is 3.86$\pm$0.07% (Figure 3b). As a comparison, MOFFlow exhibits substantially larger lattice volume deviations of 18.8$\pm$0.6%, and Genarris-3 shows deviations of 59.0$\pm$0.35% for directly generated and 10.7$\pm$0.20% after rigid-press optimization. Per-structure lattice volume deviation for MolCrystalFlow is detailed in Figure S5.

In Riemannian flow matching, introducing velocity scaling factors on the manifold for fractional coordinates (s_{u_F}) and rotational degrees of freedom (s_{u_R}) has empirically been shown to improve performance [24, 23]. To determine suitable values, we performed a grid search optimizing the 10-sample matching rate at stol $= 0.8$. As shown in Figure 3d, performance varies substantially (3.36–6.8%), with optimal results obtained for $s_{u_F} \in [5, 13]$ and $s_{u_R} \in [1, 3]$. Unless otherwise stated, Figures 3a–c use $s_{u_F} = 9$ and $s_{u_R} = 3$. In terms of sampling efficiency, MolCrystalFlow generates structures in an average of 22 ms per structure, approximately twice as fast as Genarris-3 (43 ms per structure). One should however note that MolCrystalFlow inference runs on a single GPU, whereas Genarris-3 operates entirely on multiple CPUs. While fast in absolute terms, MolCrystalFlow is slower than the non-equivariant transformer-based MOFFlow baseline (6 ms per structure under identical inference settings). However, the performance gap between MolCrystalFlow and MOFFlow cannot be offset by increasing the number of samples, as shown in Figure S4a. We further evaluate models trained on a subset of the open OMC25 molecular crystal dataset [30]. On this larger and more diverse dataset, MolCrystalFlow exhibits an even greater advantage over MOFFlow in both structure matching rates (Figure S11) and lattice volume deviations (Figure S12). Details of the selected OMC25 subset are provided in the 'Open-Source Molecular Crystal Dataset' section.

Integrating MolCrystalFlow with u-MLIP for Accelerated Crystal Structure Prediction

To assess MolCrystalFlow under a realistic CSP scenario, we construct a pipeline that integrates the MolCrystalFlow generator, a universal machine-learning interatomic potential (u-MLIP), and density functional theory (DFT) calculations (Figure 4a). We evaluate this workflow on three publicly released targets from the 3rd CCDC CSP Blind Test [33] (Figure 4b). All three experimental crystal packings contain two different axis-flip states with equal numbers of molecular building blocks. The iodine-containing target IX is excluded, as iodine is not present in the training data. The CSP workflow consists of two stages: structure generation and structure ranking. Structure generation itself consists of two steps, namely conformer generation followed by crystal packing generation. MolCrystalFlow targets the crystal packing generation stage by efficiently sampling putative polymorphs. However, these generated structures are not fully optimized. Structural optimization is therefore critical for reliable ranking, as competing polymorphs, once relaxed, often differ in lattice energies by only a few kJ/mol, on the order of chemical accuracy (4.18 kJ/mol). To avoid prohibitively expensive large-scale DFT evaluations, we employ the Universal Model for Atoms (UMA), specifically the 6M uma-s-1.1 model with the OMC task, as an intermediate screening tool [42]. An initial energy landscape is constructed using u-MLIP through a two-stage relaxation procedure (Figure S7). From this landscape, we select the top-10 structures with relative lattice energies closest to the minimum for subsequent DFT evaluation. This two-stage relaxation procedure is detailed in 'Computational settings' section.

Using the top-10 candidates selected by the UMA-OMC model, we performed more rigorous ranking using 0K DFT energies at the levels of PBE-D3 and PBE-MBD. While these dispersion-corrected functionals are widely used, it is important to note that quantitatively accurate lattice energy prediction for molecular crystals remains challenging even for state-of-the-art DFT methods [43, 44]. For each target, we identify three representative polymorphs, denoted by α, β and γ ordered by their PBE-MBD energies. Their relative energy rankings, together with the experimental structure, are shown at the levels of UMA-OMC u-MLIP, PBE-D3 and PBE-MBD (Figure 4c–e for targets VIII, X and XI, respectively). Overall, despite discrepancies in the exact ordering, u-MLIP captures the general ranking trend when compared to DFT methods. For target VIII and XI, both PBE-D3 and PBE-MBD identify a four-molecule ($Z = 4$) crystal to be close in energies to the experimental structures ($Z = 8$). For target X, UMA-OMC suggests a $Z = 2$ polymorph close to the experimental $Z = 4$ structure, but DFT places it approximately 3 kJ/mol higher in relative energy. These 0 K energy rankings indicate that MolCrystalFlow effectively generates molecular crystal structures that reside within low-energy basins of energy landscapes.

Beyond energy ranking, we evaluate packing similarity for the PBE-MBD lowest-energy polymorphs versus experimental structures using the COMPACK program implemented in CSD Python API [45]. This program aims to maximize the

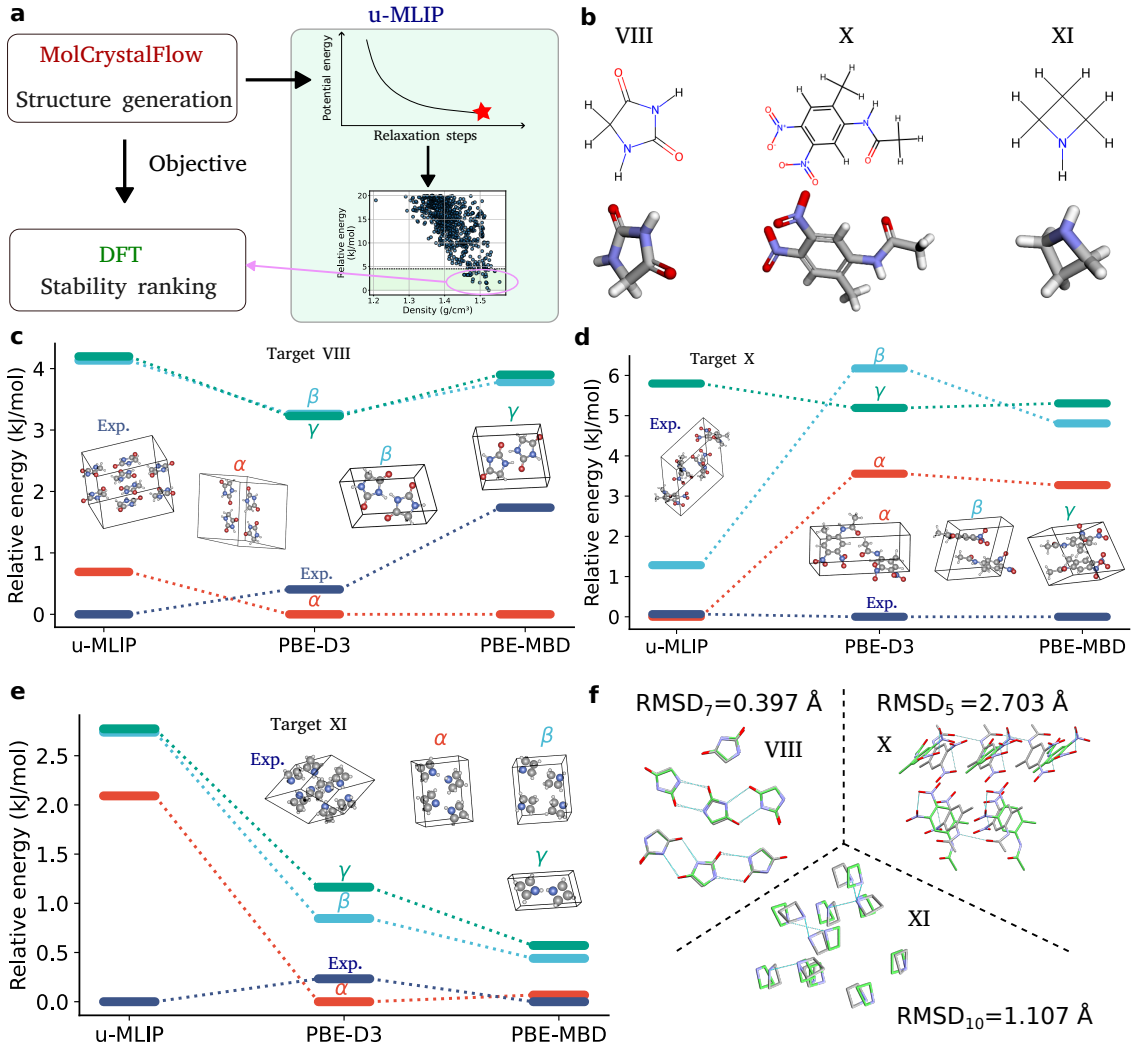

Figure 4: **Molecular crystal structure prediction by integrating MolCrystalFlow generator, u-MLIP and DFT stability ranking.** a, Four-step pipeline to integrate MolCrystalFlow for structure generation, u-MLIP for structural optimization and obtaining coarse energy landscape, and DFT for final stability ranking. b, Molecular graphs and conformer structures of three open targets from the 3rd CCDC CSP competition. c–e, Energy ranking results evaluated with u-MLIP, PBE-D3 and PBE-MBD for three generated and relaxed polymorphs whose PBE-MBD energies are the lowest, combined with the experimental structure evaluated at the same level of theory. Figure c, d and e correspond to target VIII, X and XI, respectively. f, Packing similarity analysis of the PBE-MBD lowest-energy polymorph relative to the experimental structure was performed using the COMPACK algorithm as implemented in the CSD Python API.

number of matched molecules and reports the root-mean-square deviation (RMSD) over the matched cluster. Consistent with the energy results, target VIII exhibits seven matched molecules with $\mathrm{RMSD}_7 = 0.397$ Å, and target XI shows $\mathrm{RMSD}_{10} = 1.107$ Å. For both targets, the predicted structures produce similar hydrogen-bonding networks compared to experiment, although the lowest-energy polymorph for target XI show more distinct packing differences. For target X, the model fails to recover a comparable hydrogen-bonding network ($\mathrm{RMSD}_5 = 2.703$ Å), highlighting the difficulty of identifying the exact experimental polymorph within a complex energy landscape. Details of DFT calculations and the COMPACK program are provided in the Computational settings section.

We further applied Genarris-3 to the same three targets. Its energy landscapes (Figure S8) show a smaller fraction of structures within 20 kJ/mol of the minimum compared to MolCrystalFlow, although for target XI it identifies more structures within 2.5 kJ/mol of the minimum. With space-group constraints, Genarris-3 successfully recovers the experimental structure for target VIII, but struggles for targets X and XI (Figure S9), resulting in performance comparable to MolCrystalFlow. Overall, this CSP study demonstrates the potential of integrating a flow-based generative model with universal machine-learning interatomic potentials to accelerate the discovery of stable molecular crystal polymorphs. At the same time, the difficulty in consistently reproducing experimental crystal structures underscores the remaining challenges and opportunities for further methodological advances.

Discussion

MolCrystalFlow leverages a hierarchical representation together with Riemannian flow matching to accelerate molecular crystal structure prediction. By efficiently sampling densely packed molecular crystals, it achieves state-of-the-art performance among flow-based models for large periodic systems and delivers competitive performance relative to multi-shot, rule-based crystal structure prediction methods.

Despite these advances, several limitations remain. First, the current model is fundamentally data-driven and trained solely on structural information, without explicit access to energetic information. As a result, generating intrinsically low-energy crystal packings still remains challenging. This limitation is common to many data-driven generative models and may be alleviated by incorporating energy-based formulations [46] or coupling generative sampling with inference-time scaling and reweighting strategies [47].

In addition, although open molecular crystal datasets used in this study satisfy rigid-body assumptions, conformational polymorphism plays a decisive role in many experimentally observed crystals [45, 3]. Extending MolCrystalFlow to include torsional degrees of freedom would allow the model to capture intramolecular flexibility without resorting to fully all-atom representations [48]. Such an extension would strike a balance between physical fidelity and computational tractability. Finally, experimental molecular crystals often crystallize in a limited subset of space groups, and explicitly leveraging space-group symmetry constraints could substantially reduce the generative search space [45]. Operating directly on space-group–constrained manifolds [49] or reduced asymmetric unit representations [50] represents a promising direction for improving both efficiency and accuracy.

Although molecular crystal structure prediction remains an open and challenging problem, MolCrystalFlow provides an important step toward end-to-end generative design of molecular crystals. By integrating hierarchical representations and Riemannian flow matching, it lays the groundwork for future advances in molecular crystal structure prediction based on generative models.

Methods

Hierarchical Representation of Molecular Crystals

Invariant embedding of individual molecules. We represent a molecule as a set of atoms with coordinates $x = (x_1, \ldots, x_N) \in \mathbb{R}^{N\times 3}$ and invariant atomic embedding $h = (h_1, \ldots, h_N) \in \mathbb{R}^{N\times n_f}$, where N is the number of atoms in the molecular building block. The atomic embeddings remain invariant under the E(3) coordinate transformation including translations, rotations, and reflections.

To embed the rigid molecular building block (BB), we use an equivariant graph neural network (EGNN). Node embeddings are updated through message passing between atoms within a cutoff radius:

$$m_{ij} = \phi_m(h_i^l, h_j^l, d_{ij}^2, \phi_d(d_{ij})), \quad h_i^{l+1} = \phi_h(h_i^l, \sum_{j\neq i} m_{ij}), \tag{1}$$

where l indexes the layers, $d_{ij} = \|x_i - x_j\|$ is the euclidean distance between atom i and j, and $\phi_d(d_{ij})$ is a Gaussian-expanded distance encoding detailed in Section S-14. ϕ_e and ϕ_h denote edge and node multilayer perceptrons (MLPs),

respectively. Because the message function depends only on distances, the embedding is E(3)-invariant. After message passing, the BB embedding is computed as a learnable weighted mean of the final node embeddings:

$$\hat{h}_{BB} = \sum_i w_i \cdot h_{i,\text{final}}, \qquad w_i = \phi_w([h_{i,final} \oplus \|x_i\|]), \tag{2}$$

where the weights w_i are predicted from a shallow MLP taking input as the concatenated final node features and each atom's distance to the molecular centroid.

To enrich the representation, we additionally incorporate auxiliary molecular-level descriptors ψ, inspired by Kilgour [51]. These descriptors include basic, chemical and geometric features. Three basic features consist of total atom count, number of heavy atoms and molecular weight. Eight chemical features are obtained through RDkit [52], comprising an indicator of chirality, counts of hydrogen bond donors acceptors, number of rotatable bonds, number of aromatic rings, octanol-water partition coefficient ($\log P$), and topological polar surface area (TPSA). Seven geometric features include the radius of gyration, asphericity, eccentricity, planarity and three lengths of each principal components. The combined 18 auxiliary descriptors are concatenated with the invariant EGNN-derived BB embedding to form the final building-block representation:

$$h_{\text{BB}} = [\hat{h}_{\text{BB}} \oplus \psi], \tag{3}$$

which is the input representation in the next flow stage that learns molecular packings in the crystal.

Molecules in the crystal. For flow matching, each molecule is represented as a rigid body characterized by a centroid fractional coordinate ($F_{\text{BB}} \in \mathbb{R}^3$) and a rotational orientation ($R_{\text{BB}} \in \mathbb{R}^{3\times3}$). We represent molecule centroids in fractional coordinates to enforce lattice periodicity and avoid discontinuities when centroid positions cross lattice boundaries during generation. F_{BB} is defined as the mean of the atomic coordinates transformed by by an inverse of lattice matrix, and the orientation is obtained by performing principal component analysis (PCA) on the atomic positions subtracted by a centroid position. To ensure a consistent choice of axes and preserve a right-handed coordinate system, we first align the PCA eigenvectors using an equivariant reference vector, followed by a normalization process to ensure that each molecular building block shares the same intramolecular atomic coordinates, which removes the ambiguity in SO(3) geodesic paths for isomorphic right-handed principal axes.

Combined with a lattice matrix $L \in \mathbb{R}^{3\times3}$, where $\det(L) > 0$, F_{BB} and R_{BB} specify the global arrangement of molecules in the crystal. The intra-molecular atomic coordinates $x_{i,\text{PCA}}$ are defined in the molecule's PCA frame and remain fixed throughout the flow matching stage. They are used only when reconstructing the final all-atom crystal structure from $(F_{\text{BB}}, R_{\text{BB}}, L)$. Under the rigid-body approximation, a molecular crystal is therefore specified by

$$\{(F_{i,\text{BB}}, R_{i,\text{BB}})\}_{i=1}^n, \quad L, \tag{4}$$

where n is the number of molecules (building blocks) in the crystal. For each building block $\mathcal{M}_{\text{BB}}$, we additionally retain its internal coordinates and atom types $\mathcal{M}_{\text{BB}} = (X, A)$, where $X = \{X_{j,\text{PCA}}\}_{j=1}^m$ are PCA-aligned atomic positions, $A = \{A_j\}_{j=1}^m$ are the corresponding element types, and m is the number of atoms in the molecule.

Thus, the crystal is parameterized by

$$\mathcal{C} = (F, R, L), \tag{5}$$

where $F \in [0, 1)^{3\times n}$ contains the fractional coordinates of all building block centroids, $R \in \mathbb{R}^{3\times3\times n}$ stores their rotation matrices, and L is the lattice matrix. This formulation separates local molecular geometry from global crystal packing and is therefore well-suited for scalable generative modeling.

Extracting the axis-flip state χ

For each molecular building block, we define an equivariant local reference from its atomic coordinates and species based on the previous works of Gao [53] and Kim et al. [26].

PCA axes. Consider a molecular building block with atomic coordinates $\{X_i \in \mathbb{R}^3\}_{i=1}^N$ where N is the number of atoms in the molecule and atomic masses as $\{m_i\}_{i=1}^N$. We first compute the geometric centroid $\bar{X} = \frac{1}{N}\sum_{i=1}^N X_i$. The centered coordinates are denoted as $Y_i = X_i - \bar{X}$. Principal axes for the rotational orientation of building blocks are obtained via an eigendecomposiiton of the covariance matrix $\text{Cov}(Y) = \frac{1}{N}\sum_{i=1}^N Y_i Y_i^T$. The eigenvectors are then identified and ordered by the eigenvalues to form a rotation matrix $U = [u_1, u_2, u_3]$, with the corresponding eigenvalues in the order of $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

Equivariant displacement vector. Consider the center of mass for the molecule $\bar{X}_{\text{CM}} = \sum_i m_i X_i / \sum_i m_i$. An equivariant displacement vector is then constructed as $D = \bar{X}_{\text{CM}} - \bar{X}$. If $D = 0$ for symmetric molecules, it is replaced by the displacement vector connecting the geometric centroid to its nearest neighbor.

Binary axis-flip state χ. We project D onto the PCA axes, which gives us $\nu_k = u_k^T D$, $k = 1, 2, 3$. To remove the ambiguities of PCA we define the discrete axis-flip vector $\eta = (\text{sign}(\nu_1), \text{sign}(\nu_2), \text{sign}(\nu_3)) \in \{-1, +1\}^3$. To remove symmetry-induced degeneracies, we introduces a indicator for the number of axes flipped with respect to D: $n_- = \sum_{i=1}^{3} \mathbb{I}[\eta_k = -1]$. The unique axis-flip state χ is hence given by

$$\chi = \begin{cases} 0 & \text{for} \quad n_- = 0, 2 \\ 1 & \text{for} \quad n_- = 1, 3 \end{cases}, \tag{6}$$

To ensure consistency of local coordinates across symmetry-equivalent building blocks, we renormalize local coordinates and rotation matrices based on the axis sign states so that local coordinates within the same χ group axis-flip states are the same and differ by an opposite sign for one axis across different χ groups. This procedure removes spurious axis inversions while preserving SE(3) equivariance and consistent local frame coordinates across different building blocks.

Geometric Symmetry of Molecular Crystal Structures

The molecular crystal distribution $p(L, F, R \mid h_{\text{BB}}(X, A))$ should satisfy geometric symmetries. For brevity, we drop the conditioning on the building-block embeddings and write the target distribution as $p(L, F, R)$. Geometric symmetries of inorganic crystals and molecular systems have been widely studied in prior work [54, 35, 55]. In our setting, the distribution respects permutation invariance, periodic translational invariance, and $SO(3)$ invariance. Permutation invariance requires that reordering the same types of building blocks does not change the distribution, which is naturally handled when the generative model is parameterized using a graph neural network. $SO(3)$ invariance ensures that a global rotation of the entire crystal leaves the distribution unchanged. In other words, for any orthogonal transformation $Q \in \mathbb{R}^{3\times3}$ where $Q^T Q = I$, we require $p(LQ^T, F, QR) = p(L, F, R)$. A right matrix multiplication is used for the lattice matrix since we adopt a row-vector convention for the lattice matrix throughout this work. Because fractional coordinates are defined relative to the lattice basis, they remain unaffected by such operations. Periodic translational invariance reflects the fundamental periodicity of crystals: translating all building-block centroids by any lattice vector should correspond to the same configuration. This is expressed as $p(L, \tau(F + T \cdot \mathbf{1}^\top), R) = p(L, F, R)$, where T is a translation vector, $\mathbf{1}$ is a vector of ones, and $\tau(\cdot)$ wraps coordinates back into the unit cell. These symmetries guide the design of the generative model so that the learned distribution is physically meaningful and consistent with the underlying geometry of molecular crystals. One should however note there is an ongoing debate whether the geometric symmetries can be comprised by simple extensive data augmentation on non-equivariant neural network architectures [28, 56, 57].

Joint Flow Matching on Riemannian manifolds

We employ a Riemannian flow matching generative process to learn the joint distribution $p(L, F, R)$. A conditional velocity field and an interpolant is defined between pairs of samples drawn from the base and data distribution. This conditional velocity field is constructed directly on the Riemannian manifolds for key molecular crystal modalities, which ensures that numerical integration of the learned velocity fields produce valid states that remain on the manifold. The flow matching process is parameterized by a period E(3)-invariant graph neural network detailed in the subsequent section.

The molecular centroids are represented by fractional coordinates which are wrapped around the range of [0, 1), hence living on three-dimensional tori for 3D centroid positions. The orientation can be intuitively represented by a 9-element rotation matrix on an $SO(3)$ manifold. Yet, entries of the 9-element matrix are not independent; to directly learn the flows of orientations, the rotation matrix is converted to a 3-element axis-angle representation, which is a reparameterization of the $SO(3)$ manifold through the Rodrigues' rotation formula, and any arbitrary vector in this representation corresponds to a valid rotational orientation. Despite enforcing the constraint $\det(L) > 0$, the lattice matrix is further normalized to a lower-triangular form with lattice lengths ordered as $a \leq b \leq c$. This standardization removes redundant lattice representations and simplifies the learning of lattice matrices. Lattice matrices are represented in its Euclidean space with informed base distributions fit from the training data lattice parameter distribution, as used in FlowMM [39] and OMatG [24].

MolCrystalFlow then performs flow matching jointly over three structural modalities, including lattice matrices, fractional coordinates, and molecular orientations on their corresponding manifolds [34]. A simple linear interpolant connecting each modality between $t = 0$ and $t = 1$ is used on its geodesic path, giving rise to the following forms:

$$L_t = (1-t)L_0 + tL_1, \qquad F_t = \tau\left((1-t)F_0 + tF_1'\right), \qquad R_t = \exp_{R_0}\left(t \log_{R_0}(R_1)\right), \tag{7}$$

where F_1' indicates a periodic image by unwrapping the end state F_1 in order to find the shortest geodesic path between F_1 and F_0, and $\tau(\cdot)$ wraps fractional coordinates back to the torus. Differentiating the interpolants yields the conditional

velocity fields, which define the training objective:

$$u_t(L_t|L_1, L_0) = \frac{L_1 - L_t}{1-t}, \qquad u_t(F_t|F_1, F_0) = \frac{F_1' - F_t}{1-t}, \qquad u_t(R_t|R_1, R_0) = \frac{\log_{R_t}(R_1)}{1-t}. \tag{8}$$

The loss combines denoising end-point predictions for lattices and rotations, along with a velocity-field objective for fractional coordinates,

$$\mathcal{L}(\theta) = \mathbb{E}_{t,(\mathcal{C}_0,\mathcal{C}_1)}\left[\frac{1}{(1-t)^2}\left(\lambda_L\|\hat{L}_1 - L_1\|^2 + \lambda_R\|\log_{R_t}(\hat{R}_1) - \log_{R_t}(R_1)\|^2\right) + \lambda_F\|\hat{u}_t(F_t) - u_t(F_t)\|^2\right].$$

We set $\lambda_L = 0.1$, $\lambda_R = 1$, and $\lambda_F = 2$ in this work. To prevent numerical instability near $t = 1$ during training, we clip time at $t = 0.9$; that is, we transform the denominator $(1 - t)$ in the loss function to $(1 - \min(t, 0.9))$ to avoid a division by zero.

Neural Network Architecture for Periodic E(3)-Invariant Flow Matching

We parameterize the joint flow matching over lattice matrices, rotation matrices, and centroid fractional coordinates using a periodic E(3)-invariant graph neural network adapted from the CSPNet [35]. The model learns the velocity fields of fractional coordinates ($u_t(F_t)$), rotational orientation ($u_t(R_t)$) and lattice matrices ($u_t(L_t)$). The initial node embeddings are given by the building block (BB) representations, concatenated with a time embedding for t and the projected into the initial dimension of the BB embedding. During message passing, interactions between two building blocks depend on their relative fractional coordinates, relative orientations, and the current lattice matrix. The lattice matrix L_t is shared across the entire crystal, while the fractional coordinates F_t and the rotation matrices R_t are defined for each building block. At layer s, message passing between building blocks i and j is defined as:

$$m_{ij}^{(s)} = \phi_m\left(h_i^{(s-1)}, h_j^{(s-1)}, L^\top L, \psi_{\mathrm{FT}}(F_j - F_i), \Psi_R(R_i^\top R_j), m_{R,L}, m_{F,L}\right), \tag{9}$$

$$m_i^{(s)} = \sum_{j=1}^{N} m_{ij}^{(s)}, \quad h_i^{(s)} = h_i^{(s-1)} + \phi_h\left(h_i^{(s-1)}, m_i^{(s)}\right), \tag{10}$$

where ψ_{FT} is a Fourier-based embedding using trigonometric functions enforcing periodicity on the torus [36], and $\Psi_R(R_i^\top R_j)$ encodes relative orientations. The matrix $R_i^\top R_j$ lies in $SO(3)$ because we have $\det(R_i^\top R_j) = 1$. Therefore, it can be expressed in an axis–angle representation. We compute the rotation angle ω and the azimuthal and inclination angles ρ and κ. Since ω and θ are periodic (mod 2π), we embed $\omega/2\pi$ and $\theta/2\pi$ using trigonometric functions, while κ is encoded via a shallow MLP $\phi(\cdot)$:

$$\Psi_R(R_i^T R_j) = [\psi_{\mathrm{FT}}(\omega/2\pi) \oplus \psi_{\mathrm{FT}}(\rho/2\pi) \oplus \phi(\kappa)]. \tag{11}$$

Regarding the interaction terms of L with R and F, we have unit vectors $m_{R,L} = L^T L R_i^T R_j / \|L^T L R_i^T R_j + \epsilon\|$ and $m_{F,L} = L^T L(F_j - F_i)/\|L^T L(F_j - F_i) + \epsilon\|$ where $\epsilon = 10^{-8}$ is to avoid divide-by-zero. All inputs to message passing are thus invariant to global rotations and translations, making the network periodic E(3)-invariant. After f message-passing layers, the final node embeddings $h_i^{(f)}$ are used to parameterize the outputs:

$$\Delta R_i = \phi_R\left(h_i^{(f)}\right), \quad R_{i,t=1} = R_i \times \Delta R_i$$

$$\Delta L_j = \phi_{L_j}\left(\frac{1}{m}\sum_{k\in\mathcal{C}_j} h_k^{(f)}\right) \times L_j^T, \quad L_{j,t=1} = L_j + \Delta L_j$$

$$u_t(F_t) = \phi_F(h_i^{(f)}),$$

The update of the lattice is invariant to global lattice rotation, and multiplication by L_t ensures equivariance, demonstrated by Jiao et al. [35]. Similar compositional operation is applied to the rotational update to ensure the global rotational equivariance of building block orientations.

During inference, the centroid velocity field $u_t(F_t)$ is integrated directly and the predicted fractional coordinate is mapped back onto the torus. In contrast, the velocity fields at time t for lattice matrix L and molecular orientations R are computed based on the denoised states and interpolant states at time t. The derived velocity fields are then integrated numerically in time to move to the next time step. The ordinary differential equations (ODEs) governing all velocity fields are solved using Euler integration with 50 discrete time steps. Velocity annealing is applied during integration for both the fractional coordinates and rotational orientations, with scaling factors s_{u_F} and s_{u_R}, respectively. The state (x_t) update at each step is given by $x_{t+\Delta t} = x_t + (1 + st) \cdot u(x_t) \cdot \Delta t$, where the annealing factors exert the strongest influence as $t \to 1$.

Base Distributions and Optimal Transport

Base distributions of fractional coordinates, lattice matrices and rotation matrices. Fractional coordinates are drawn from a uniform prior with support on [0,1). Lattice matrices are parameterized by six lattice parameters fitted from the training data and constructed in a normalized lower-triangular form. Lattice lengths are sampled from Gaussian distributions, while lattice angles are sampled from a uniform distribution with support [60°, 120°]. The sampled lattice lengths and angles are then used to construct the lattice matrix base distributions in the lower-triangular form. For molecular orientations, we adopt a symmetric-rotation prior. Specifically, the first rotation matrix is sampled uniformly from the SO(3) manifold, and the remaining rotation matrices are deterministically generated according to symmetry relations with respect to this reference orientation. As shown in the Supplementary Information, this symmetric-rotation prior preserves full SE(3) invariance, which is essential for ensuring that the generated molecular crystal structures, $p(L, F, R)$, are also SE(3)-invariant when propagated through an equivariant flow-matching process.

Optimal transport for fractional coordinates and rotation matrices. To mitigate trajectory crossings during flow matching, we employ a data-dependent optimal transport procedure that aligns source and target samples, thereby improving both inference quality and sampling efficiency. We permute building block indices within the same χ group to minimize geodesic distances in the spaces of fractional coordinates (wrapped distances) and molecular orientations (rotational angles). For example, given a batch of rotation matrix pairs (R_0, R_1), we seek the optimal permutation within each χ group that minimizes the geodesic distance on $SO(3)$, which can be formulated as $\arg\min_{P_\chi} d_{RM}\left(P_\chi(R_0), R_1\right)$, where d_{RM} denotes the geodesic distance between rotation matrices and P_χ represents a permutation operator acting only on indices within the same χ group.

Open-Source Molecular Crystal Dataset

Data preprocessing. Each molecular crystal is decomposed to constituent molecular building blocks based on atomic connectivity. Each building block is identified by using a graph-based traversal constructed from interatomic neighbors using element-dependent cutoff radii. The neighboring environment is determined through atomic simulation environment [58], and the cutoffs for each element is manually chosen based on covalent radius of each atom type. Two atoms are considered bonded if their separation distance falls within the add-up of a pair of atoms. A minimum-image convention is considered to account for cross-cell distances across periodic boundaries. Each grouped molecule will share a building block index that is used for determining the centroids and rotational orientation. In addition, each molecular building block is unwrapped so that all atoms are fully connected in their Cartesian coordinates. To facilitate learning of lattice matrices, we first standardize each lattice by sorting the lattice vectors by their lengths, then transform the reordered matrix into a lower-triangular form via a 'QR' decomposition applied to its transpose.

Thurlemann molecular crystal dataset. We evaluate MolCrystalFlow on the molecular crystal dataset curated by Thurlemann et al. [29], which is derived from the Cambridge Structural Database [45]. The dataset comprises molecular crystals containing seven elements (H, C, N, O, F, Cl, S) and originally includes 11,489×5 structures, where the five entries correspond to distinct lattice strain levels. All structures were relaxed with a fixed experimental cell at the DFT level using the PBE-XDM functional as implemented in Quantum ESPRESSO [59]. We selected the 11,489 unstrained structures as the initial candidate dataset. After preprocessing and filtering, the final dataset consists of 11,488 structures. The structure filtered out is one structure classified to have four building blocks but actually labeled with only two molecules in the original data source ($Z = 2$). This structure is shown in Figure S6. The dataset was split such that identical chemical formulas do not appear across different splits, with 10,000, 738, and 750 structures used for training, validation, and testing, respectively. A detailed explanatory data analysis is provided in Figure S1.

Selected OMC25 data (OMC25-MCF). OMC25 is the largest publicly available molecular crystal dataset generated at the PBE-D3 level by FAIR Chemistry using VASP [30]. It comprises more than 26 million structures, sampled from geometry-optimized conformers in the OE62 dataset [60]. The dataset contains 12 elements (C, H, N, O, P, S, Cl, F, I, B, Br, Si), and it is restricted to single-component molecular crystals and adopts a rigid-body approximation. For training MolCrystalFlow, we retain a single crystal packing per conformer, selected as the polymorph of the lowest lattice energy with a maximum atomic force below 0.05 eV/Å. After filtering, a total of 46,802 structures remain, and this selected subset of OMC25 is termed 'OMC25-MCF'. We split the dataset into 42,121 training structures (90%), 3,681 validation structures, and 1,000 test structures. Exploratory data analysis of the full OMC25-MCF data is also provided in Figure S10.

Evaluation Metrics

Evaluation on the raw generated structures. Crystal structure generation performance is first evaluated using *StructureMatcher* from the pymatgen library [41], with lattice tolerance *ltol=0.3*, angle tolerance *atol=10*, and site tolerances in the range of 0.5 to 1.2. Volume scaling is disabled because uniform rescaling would alter intramolecular bond lengths, which is undesirable for molecular crystal structures. Owing to the presence of numerous polymorphs for a given molecular conformer, multiple candidate structures are generated for each test case. A prediction is considered correct if at least one generated structure matches the corresponding ground-truth structure under the above criteria.

Evaluation using CCDC CSP compounds. We further benchmark MolCrystalFlow on three targets from the 3rd CCDC CSP Blind Test [33]. Among these targets, the experimental crystal packings of targets VIII and XI consist of eight molecules per unit cell ($Z = 8$), whereas target X has $Z = 4$. For sampling with MolCrystalFlow, we consider unit cells containing $Z \in \{2, 4\}$ molecules, which are among the most frequently observed values in the Cambridge Structural Database (CSD) [45]. Although $Z = 1$ and $Z = 8$ also account for a substantial fraction of molecular crystals in CSD, we do not sample these cases here. Preliminary tests indicate that $Z = 1$ structures are less stable than those with $Z = 2$ and $Z = 4$, while the Thurlemann dataset contains only a limited number of $Z = 8$ examples, making reliable generation of eight-molecule unit cells challenging. For each value of Z, we generate 200 candidate structures under a maximum intermolecular ellipsoidal volume overlap of 5%. We consider two χ-state scenarios: one in which all molecules share the same χ state, and another in which the crystal contains equal numbers of molecules in two distinct χ states. This results in a total of 800 sampled structures per molecular conformer. The entire sampling procedure is repeated twice using different random seeds. Structural optimization and energy ranking are performed using the UMA-OMC (6M UMA s-1.p.1 model) universal machine–learning interatomic potentials (u-MLIP), followed by DFT calculations at the PBE-D3(BJ) and PBE-MBD levels of theory. We further examine structural fidelity by computing the root mean squared distances (RMSDs) for 15-molecule assemblies, and RMSDs are evaluated on matched molecules using the *COMPACK* program as implemented in the CSD Python API [45]. Molecular matching is performed using distance and angular tolerances of 50%, excluding hydrogen atoms and without allowing molecular differences.

Computational settings

Model Training and Inference. Both MolCrystalFlow and MOFFlow models were built using PyTorch. Both training and inference use one Nvidia B200 graphic card with 192 GB memory. The baseline MOFFlow model was trained using the same hyperparameter settings of the original work [26]. Hyperparameters of MolCrystalFlow are included in Table [**?**].

u-MLIP model evaluation and structural optimization. UMA-OMC u-MLIP was used to optimize the directly generated structures, and it is a two-stage process implemented with atomic simulation environment (ASE) [58]. First, each generated structure is relaxed for up to 100 steps or until the maximum atomic force is below 0.05 eV/Å using a rigid-body constraint and a fixed lattice based on the BFGS algorithm. Second, the full atomic and lattice degrees of freedom are relaxed using the UMA-OMC model for up to 1000 steps or until the maximum atomic force falls below 0.01 eV/Å, implemented via the *FrechetCellFilter* of ASE. The 10 lowest-energy structures across all sampled Z values are selected for final DFT validation for each round. We have in total 20 structures for two rounds with different structure generation random seeds.

DFT calculations. We used VASP 6.4.3 for DFT evaluations [61]. The exchange-correlation interactions were described using Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation [62]. Core-valence interactions were treated using the projected augmented-wave (PAW) method, with a plane-wave energy cutoff of 520 eV. The structural relaxation was performed with the conjugate-gradient algorithm with a force convergence criterion of 0.01 eV/Å and an energy convergence threshold of 10^{-7} eV. A maximum of 1500 ionic steps was used. Brillium-zone integrations employ the F-centered Monkhorst-Pack scheme, with a Gaussian smearing of 0.05 eV. K-point meshes were generated automatically using the pymatgen based on a k-point density criterion of 1000 k-points per reciprocal atom [41]. Two types of dispersions were considered, including D3 (BJ) [63] and many-body dispersion (MBD) [64].

Structure generation with Genarris-3. We applied Genarris-3.0 [40] as a baseline comparison using a two-step workflow that first generates trial molecular crystal structures across compatible space groups and then compacts each candidate with Symmetric Rigid Press while preserving space-group symmetry. For the X, XI, and VIII targets, we generated 800 accepted structures per space group, while for the Thurlemann validation set we generated 200 accepted structures per space group. In the generation step we allowed up to 10,000,000 attempts per space group and per sampled volume. Unit-cell volumes were set from the built-in volume prediction and initialized at 1.5$\times$ the predicted mean to facilitate placement before compaction. Short-range contact screening used sr = 0.85 with tol =

0.01, and neighbor identification used natural_cutoff_mult = 1.2. In the Symmetric Rigid Press step, structures were compressed toward close packing using a smooth overlap-penalty model optimized with BFGS to tol = 0.01, with sr = 0.85 and natural_cutoff_mult = 1.2 applied consistently. For the validation set, the resulting structures were uniformly subsampled to produce Figure 3.

Acknowledgment

Cheng Zeng thanks Michael Kilgour (NYU) for insightful discussion during the AI for Chemistry symposium held at New York University. The authors acknowledge funding from NSF Grant OAC-2311632 and from the AI and Complex Computational Research Award of the University of Florida. S.M also acknowledges support from the Simons Center for Computational Physical Chemistry (Simons Foundation grant 839534). The authors also gracefully acknowledge UFIT Research Computing for computational resources and consultation, as well as the NVIDIA AI Technology Center at UF.

References

[1] Sheetal S. Buddhadev and Kevin C. Garala. Pharmaceutical Cocrystals—A Review. *Proceedings*, 62(1):14, 2021.

[2] Markus von Raumer and Rolf Hilfiker. Solid State and Polymorphism of the Drug Substance in the Context of Quality by Design and ICH Guidelines Q8–Q12. In *Polymorphism in the Pharmaceutical Industry*, pages 1–30. John Wiley & Sons, Ltd, 2018.

[3] John Bauer, Stephen Spanton, Rodger Henry, John Quick, Walter Dziki, William Porter, and John Morris. Ritonavir: An Extraordinary Example of Conformational Polymorphism. *Pharmaceutical Research*, 18(6):859–866, June 2001.

[4] Chuanxin Wei, Liang Li, Yingying Zheng, Lizhi Wang, Jingyao Ma, Man Xu, Jinyi Lin, Linghai Xie, Panče Naumov, Xuehua Ding, Quanyou Feng, and Wei Huang. Flexible molecular crystals for optoelectronic applications. *Chemical Society Reviews*, 53(8):3687–3713, April 2024.

[5] Dario Braga, Lucia Casali, and Fabrizia Grepioni. The Relevance of Crystal Forms in the Pharmaceutical Field: Sword of Damocles or Innovation Tools? *International Journal of Molecular Sciences*, 23(16):9013, August 2022.

[6] Kazuo Takimiya, Kirill Bulgarevich, and Kohsuke Kawabata. Crystal-Structure Control of Molecular Semiconductors by Methylthiolation: Toward Ultrahigh Mobility. *Accounts of Chemical Research*, 57(6):884–894, March 2024.

[7] Atta Muhammad, Leandro Iorio, Fabrizio Quadrini, Eliodoro Chiavazzo, and Matteo Fasano. Molecular dynamics analysis of plastic crystals as solid-solid phase-change materials for thermal energy storage. *Journal of Energy Storage*, 125:116978, July 2025.

[8] Gregory J. O. Beran. Frontiers of molecular crystal structure prediction for pharmaceuticals and functional organic materials. *Chemical Science*, 14(46):13290–13312, 2023. Publisher: Royal Society of Chemistry.

[9] L. M. Hunnisett, J. Nyman, N. Francia, N. S. Abraham, C. S. Adjiman, S. Aitipamula, T. Alkhidir, M. Almehairbi, A. Anelli, D. M. Anstine, J. E. Anthony, J. E. Arnold, F. Bahrami, M. A. Bellucci, R. M. Bhardwaj, I. Bier, J. A. Bis, A. D. Boese, D. H. Bowskill, J. Bramley, J. G. Brandenburg, D. E. Braun, P. W. V. Butler, J. Cadden, S. Carino, E. J. Chan, C. Chang, B. Cheng, S. M. Clarke, S. J. Coles, R. I. Cooper, R. Couch, R. Cuadrado, T. Darden, G. M. Day, H. Dietrich, Y. Ding, A. DiPasquale, B. Dhokale, B. P. van Eijck, M. R. J. Elsegood, D. Firaha, W. Fu, K. Fukuzawa, J. Glover, H. Goto, C. Greenwell, R. Guo, J. Harter, J. Helfferich, D. W. M. Hofmann, J. Hoja, J. Hone, R. Hong, G. Hutchison, Y. Ikabata, O. Isayev, O. Ishaque, V. Jain, Y. Jin, A. Jing, E. R. Johnson, I. Jones, K. V. J. Jose, E. A. Kabova, A. Keates, P. F. Kelly, D. Khakimov, S. Konstantinopoulos, L. N. Kuleshova, H. Li, X. Lin, A. List, C. Liu, Y. M. Liu, Z. Liu, Z.-P. Liu, J. W. Lubach, N. Marom, A. A. Maryewski, H. Matsui, A. Mattei, R. A. Mayo, J. W. Melkumov, S. Mohamed, Z. Momenzadeh Abardeh, H. S. Muddana, N. Nakayama, K. S. Nayal, M. A. Neumann, R. Nikhar, S. Obata, D. O'Connor, A. R. Oganov, K. Okuwaki, A. Otero-de-la Roza, C. C. Pantelides, S. Parkin, C. J. Pickard, L. Pilia, T. Pivina, R. Podeszwa, A. J. A. Price, L. S. Price, S. L. Price, M. R. Probert, A. Pulido, G. R. Ramteke, A. U. Rehman, S. M. Reutzel-Edens, J. Rogal, M. J. Ross, A. F. Rumson, G. Sadiq, Z. M. Saeed, A. Salimi, M. Salvalaglio, L. Sanders de Almada, K. Sasikumar, S. Sekharan, C. Shang, K. Shankland, K. Shinohara, B. Shi, X. Shi, A. G. Skillman, H. Song, N. Strasser, J. van de Streek, I. J. Sugden, G. Sun, K. Szalewicz, B. I. Tan, L. Tan, F. Tarczynski, C. R. Taylor, A. Tkatchenko, R. Tom, M. E. Tuckerman, Y. Utsumi, L. Vogt-Maranto, J. Weatherston, L. J. Wilkinson, R. D. Willacy, L. Wojtas, G. R. Woollam, Z. Yang, E. Yonemochi, X. Yue, Q. Zeng, Y. Zhang, T. Zhou, Y. Zhou, R. Zubatyuk, and J. C. Cole. The

seventh blind test of crystal structure prediction: structure generation methods. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 80(6):517–547, December 2024.

[10] Kamal Singh Nayal, Dana O'Connor, Roman Zubatyuk, Dylan M. Anstine, Yi Yang, Rithwik Tom, Wenda Deng, Kehan Tang, Noa Marom, and Olexandr Isayev. Efficient Molecular Crystal Structure Prediction and Stability Assessment with AIMNet2 Neural Network Potentials, June 2025.

[11] Vahe Gharakhanyan, Yi Yang, Luis Barroso-Luque, Muhammed Shuaibi, Daniel S. Levine, Kyle Michel, Viachaslau Bernat, Misko Dzamba, Xiang Fu, Meng Gao, Xingyu Liu, Keian Noori, Lafe J. Purvis, Tingling Rao, Brandon M. Wood, Ammar Rizvi, Matt Uyttendaele, Andrew J. Ouderkirk, Chiara Daraio, C. Lawrence Zitnick, Arman Boromand, Noa Marom, Zachary W. Ulissi, and Anuroop Sriram. FastCSP: Accelerated Molecular Crystal Structure Prediction with Universal Model for Atoms, August 2025.

[12] Yi Yang, Rithwik Tom, Jose A. G. L. Wui, Jonathan E. Moussa, and Noa Marom. Genarris 3.0: Generating Close-Packed Molecular Crystal Structures with Rigid Press, June 2025.

[13] Farren Curtis, Xiayue Li, Timothy Rose, Álvaro Vázquez-Mayagoitia, Saswata Bhattacharya, Luca M. Ghiringhelli, and Noa Marom. GAtor: A First-Principles Genetic Algorithm for Molecular Crystal Structure Prediction. *Journal of Chemical Theory and Computation*, 14(4):2246–2264, April 2018. Publisher: American Chemical Society.

[14] Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow Matching for Generative Modeling, February 2023.

[15] Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic Interpolants: A Unifying Framework for Flows and Diffusions, November 2023.

[16] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow, September 2022.

[17] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-Based Generative Modeling through Stochastic Differential Equations, February 2021.

[18] Yuanqi Du, Arian R Jamasb, Jeff Guo, Tianfan Fu, Charles Harris, Yingheng Wang, Chenru Duan, Pietro Liò, Philippe Schwaller, and Tom L Blundell. Machine learning-aided generative molecular design. *Nature Machine Intelligence*, 6(6):589–604, 2024.

[19] Cheng Zeng, Zulqarnain Khan, and Nathan Post. Data-efficient and interpretable inverse materials design using a disentangled variational autoencoder. *AI & Materials*, 1(1):1–16, December 2024.

[20] Emiel Hoogeboom, Victor Garcia Satorras, Clement Vignac, and Max Welling. Equivariant Diffusion for Molecule Generation in 3D. In *Proceedings of the 39th International Conference on Machine Learning*, pages 8867–8887. PMLR, June 2022.

[21] Cheng Zeng, Jirui Jin, Connor Ambrose, George Karypis, Mark Transtrum, Ellad B. Tadmor, Richard G. Hennig, Adrian Roitberg, Stefano Martiniani, and Mingjie Liu. PropMolFlow: property-guided molecule generation with geometry-complete flow matching. *Nature Computational Science*, pages 1–10, January 2026.

[22] Jirui Jin, Cheng Zeng, Pawan Prakash, Ellad B. Tadmor, Adrian Roitberg, Richard G. Hennig, Stefano Martiniani, and Mingjie Liu. MolGuidance: Advanced Guidance Strategies for Conditional Molecular Generation with Flow Matching, December 2025. arXiv:2512.12198 [cs].

[23] Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative Flows on Discrete State-Spaces: Enabling Multimodal Flows with Applications to Protein Co-Design, June 2024.

[24] Philipp Höllmer, Thomas Egg, Maya Martirossyan, Eric Fuemmeler, Zeren Shui, Amit Gupta, Pawan Prakash, Adrian Roitberg, Mingjie Liu, George Karypis, Mark Transtrum, Richard Hennig, Ellad B. Tadmor, and Stefano Martiniani. Open materials generation with stochastic interpolants. In *Forty-second International Conference on Machine Learning*, 2025.

[25] Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Zilong Wang, Aliaksandra Shysheya, Jonathan Crabbé, Shoko Ueda, Roberto Sordillo, Lixin Sun, Jake Smith, Bichlien Nguyen, Hannes Schulz, Sarah Lewis, Chin-Wei Huang, Ziheng Lu, Yichi Zhou, Han Yang, Hongxia Hao, Jielan Li, Chunlei Yang, Wenjie Li, Ryota Tomioka, and Tian Xie. A generative model for inorganic materials design. *Nature*, pages 1–3, January 2025.

[26] Nayoung Kim, Seongsu Kim, Minsu Kim, Jinkyoo Park, and Sungsoo Ahn. MOFFlow: Flow Matching for Structure Prediction of Metal-Organic Frameworks, March 2025.

[27] Hongyu Guo, Yoshua Bengio, and Shengchao Liu. AssembleFlow: Rigid Flow Matching with Inertial Frames for Molecular Assembly. October 2024.

[28] Emily Jin, Andrei Cristian Nica, Mikhail Galkin, Jarrid Rector-Brooks, Kin Long Kelvin Lee, Santiago Miret, Frances H. Arnold, Michael Bronstein, Avishek Joey Bose, Alexander Tong, and Cheng-Hao Liu. Oxtal: An all-atom diffusion model for organic crystal structure prediction, 2025.

[29] Moritz Thürlemann, Lennard Böselt, and Sereina Riniker. Regularized by Physics: Graph Neural Network Parametrized Potentials for the Description of Intermolecular Interactions. *Journal of Chemical Theory and Computation*, 19(2):562–579, January 2023.

[30] Vahe Gharakhanyan, Luis Barroso-Luque, Yi Yang, Muhammed Shuaibi, Kyle Michel, Daniel S. Levine, Misko Dzamba, Xiang Fu, Meng Gao, Xingyu Liu, Haoran Ni, Keian Noori, Brandon M. Wood, Matt Uyttendaele, Arman Boromand, C. Lawrence Zitnick, Noa Marom, Zachary W. Ulissi, and Anuroop Sriram. Open Molecular Crystals 2025 (OMC25) Dataset and Models, August 2025.

[31] Paul C. D. Hawkins, A. Geoffrey Skillman, Gregory L. Warren, Benjamin A. Ellingson, and Matthew T. Stahl. 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*, 50(4):572–584, April 2010. Publisher: American Chemical Society.

[32] Sereina Riniker and Gregory A. Landrum. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *Journal of Chemical Information and Modeling*, 55(12):2562–2574, December 2015. Publisher: American Chemical Society.

[33] G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, and P. Verwer. A third blind test of crystal structure prediction. *Acta Crystallographica Section B: Structural Science*, 61(5):511–527, October 2005.

[34] Ricky T. Q. Chen and Yaron Lipman. Flow Matching on General Geometries, February 2024.

[35] Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal Structure Prediction by Joint Equivariant Diffusion, March 2024.

[36] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is All you Need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[37] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020.

[38] Ian Dunn and David Ryan Koes. Mixed Continuous and Categorical Flow Matching for 3D De Novo Molecule Generation, April 2024.

[39] Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: Generating Materials with Riemannian Flow Matching, June 2024.

[40] Yi Yang, Rithwik Tom, Jose A. G. L. Wui, Jonathan E. Moussa, and Noa Marom. Genarris 3.0: Generating Close-Packed Molecular Crystal Structures with Rigid Press. *Journal of Chemical Theory and Computation*, 21(21):11318–11332, November 2025.

[41] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, February 2013.

[42] Brandon M. Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R. Kitchin, Daniel S. Levine, Kyle Michel, Anuroop Sriram, Taco Cohen, Abhishek Das, Ammar Rizvi, Sushree Jagriti Sahoo, Zachary W. Ulissi, and C. Lawrence Zitnick. UMA: A Family of Universal Models for Atoms, June 2025.

[43] Cody J. Perry, Sebastian A. Ramos, Maria C. Phelps, Leonard J. Mueller, and Gregory J. O. Beran. Taming Tautomerism in Organic Crystal Structure Prediction. *Journal of the American Chemical Society*, 147(30):26865–26876, July 2025.

[44] Amrita Chattopadhyay, Adam R. Hill, Sarah E. Wright, Gregory J. O. Beran, and Aurora J. Cruz-Cabeza. Lattice Energy Partitions in Crystals of Flexible Molecules and the 40% Limit. *Journal of the American Chemical Society*, 147(43):39192–39203, October 2025.

[45] Colin R. Groom, Ian J. Bruno, Matthew P. Lightfoot, and Suzanna C. Ward. The Cambridge Structural Database. *Acta Crystallographica Section B, Structural Science, Crystal Engineering and Materials*, 72(Pt 2):171–179, April 2016.

[46] Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky T. Q. Chen. Adjoint schrödinger bridge sampler. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

[47] Konstantin Mark, Leonard Galustian, Maximilian P.-P. Kovar, and Esther Heid. Feynman-Kac-Flow: Inference Steering of Conditional Flow Matching to an Energy-Tilted Posterior, September 2025. arXiv:2509.01543 [cs].

[48] Nayoung Kim, Seongsu Kim, and Sungsoo Ahn. Flexible MOF Generation with Torsion-Aware Flow Matching, September 2025.

[49] Omri Puny, Yaron Lipman, and Benjamin Kurt Miller. Space Group Conditional Flow Matching, September 2025.

[50] Daniel Levy, Siba Smarak Panigrahi, Sékou-Oumar Kaba, Qiang Zhu, Kin Long Kelvin Lee, Mikhail Galkin, Santiago Miret, and Siamak Ravanbakhsh. SymmCD: Symmetry-Preserving Crystal Generation with Diffusion Models. October 2024.

[51] Michael Kilgour, Jutta Rogal, and Mark Tuckerman. Geometric Deep Learning for Molecular Crystal Structure Prediction. *Journal of Chemical Theory and Computation*, 19(14):4743–4756, July 2023.

[52] RDKit. RDKit: Open-Source Cheminformatics, 2024.

[53] Nicholas Gao and Stephan Günnemann. Sampling-free inference for ab-initio potential energy surface networks. In *The Eleventh International Conference on Learning Representations*, 2023.

[54] Jonas Köhler, Leon Klein, and Frank Noe. Equivariant Flows: Exact Likelihood Generative Learning for Symmetric Densities. In *Proceedings of the 37th International Conference on Machine Learning*, pages 5361–5370. PMLR, November 2020.

[55] Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. GeoDiff: a Geometric Diffusion Model for Molecular Conformation Generation, March 2022.

[56] Filippo Bigi, Sanggyu Chong, Agustinus Kristiadi, and Michele Ceriotti. FlashMD: long-stride, universal prediction of molecular dynamics, January 2026.

[57] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, May 2022.

[58] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, June 2017.

[59] 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 Jr, 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. Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 29(46):465901, November 2017.

[60] Annika Stuke, Christian Kunkel, Dorothea Golze, Milica Todorović, Johannes T. Margraf, Karsten Reuter, Patrick Rinke, and Harald Oberhofer. Atomic structures and orbital energies of 61,489 crystal-forming organic molecules. *Scientific Data*, 7(1):58, February 2020.

[61] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, October 1996.

[62] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, October 1996.

[63] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15):154104, April 2010.

[64] Alexandre Tkatchenko, Robert A. DiStasio, Roberto Car, and Matthias Scheffler. Accurate and Efficient Method for Many-Body van der Waals Interactions. *Physical Review Letters*, 108(23):236402, June 2012.