FRAMEWORK-CONSTRAINED MATERIALS GENERATION

Maya M. Martirossyan[1, 2,*], Hillary Pan[3,*], Philipp Höllmer[1, 2], Stefano Martiniani[1, 2, 4, 5,†]
[1]Center for Soft Matter Research, Department of Physics, New York University
[2]Simons Center for Computational Physical Chemistry, Department of Chemistry, New York University
[3]Institute for Systems Genetics, NYU Langone Health
[4]Courant Institute of Mathematical Sciences, New York University
[5]Center for Neural Science, New York University

ABSTRACT

Theoretical predictions of novel materials rely on a variety of techniques, from data-mining-based ion substitution methods to fully *de novo* generative models. Each approach imposes specific constraints and encounters unique challenges when extrapolating beyond the distribution of known materials to discover new ones. The foremost goal of materials generation is to produce stable, unique, and novel (SUN) structures. Predicted structures that meet these criteria, however, typically result from novel chemical compositions mapped onto known structural prototypes or frameworks. We leverage this heuristic insight to present a new generative approach for the *de novo* prediction of inorganic crystalline materials constrained by their *structural framework—i.e.*, the unit cell shapes and respective particle positions. Our method, termed OMatG-FC (Open Materials Framework-Constrained Generation), learns *via* a combination of discrete flow matching on the atomic species and stochastic interpolants for the volume scaling of the unit cell lattice, while fixing in-place the fractional coordinates and the unit-volume lattice vectors, dramatically reducing the dimensionality of the problem compared to unconstrained (framework-free) generation. Our method is more flexible than existing structure-constrained methods, as it allows for symmetry-breaking or chemical disorder on sites of equivalent crystalline symmetry. We also investigate the number of unique frameworks in the MP-20 dataset, demonstrating the propensity of duplicate frameworks within the dataset. We benchmark the performance of our model using the LeMat-GenBench suite using three ML interatomic potentials and achieve state-of-the-art SUN rate, eschewing the need to learn atomic packings for materials discovery.

1 INTRODUCTION

A foundational goal of materials science is the design and discovery of novel materials with desired function. The advent of battery materials Kamaya et al. (2011); Jang et al. (2024), semiconductors Nomura et al. (2004), and alloys Gludovatz et al. (2016) have historically shaped the fabric of our society, with further technological leaps promised by "holy-grail" materials such as room-temperature superconductors Boeri et al. (2022); Prakash et al. (2025).

Theoretical predictions of novel materials can support the discovery process: Traditional approaches include high-throughput Jain et al. (2011); Curtarolo et al. (2013) or ab-initio random structure search Pickard & Needs (2011), which are based on first-principles methods like density functional theory (DFT), but suffer from poor sampling and high computational cost. Given the vast search space of candidate materials, machine-learning (ML) methods offer a promising avenue for dramatically accelerating materials discovery. Among these, generative models De Breuck et al. (2025); Metni et al. (2025); Bone et al. (2025) stand out for their ability to propose novel materials by learning from experimentally or computationally derived crystal structures represented as unit cells.

[*]These authors contributed equally.
[†]Corresponding author: sm7683@nyu.edu.

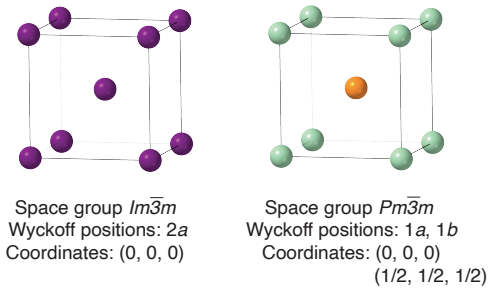

Figure 1: Two crystal structures—a BCC-type crystal (left) and a CsCl prototype (right)—with the same structural framework but different structural prototypes, labeled by their space groups and corresponding Wyckoff positions.

Structure is of chief importance for material properties, especially for inorganic crystals, whose behavior depends sensitively on atomic-level periodicity Vu et al. (2023). This has motivated extensive work on crystal structure prediction (CSP) from chemical composition, as well as *de novo* generation (DNG) of novel compositions and their structural packings. The primary challenge for such methods is learning to represent and sample the space of stable crystalline materials—spanning all possible structural packings and chemical species—in order to discover viable candidate materials that can be synthesized Cheetham & Seshadri (2024).

1.1 Related Works

Significant prior work in materials prediction leverage knowledge of existing materials combined with heuristic constraints. Such methods predate ML techniques: For example, the empirical Hume-Rothery rules Hume-Rothery (1939) identify plausible solid solutions via substitutions of chemical species within an existing crystal structure template. Modern probabilistic models have also been applied to predict new compounds from structural templates *via* ionic substitutions Hautier et al. (2011): A key feature of this method is the constraint for candidate structures to remain within their *structural prototype* (see Sec. 2.1). In contrast, machine-learned CSP models predict stable atomic packings given a fixed chemical composition. Several such CSP methods have been proposed with the goal of improving on *ab initio* random structure search Pickard (2016), such as CDVAE Xie et al. (2022), DiffCSP Jiao et al. (2023), FlowMM Miller et al. (2024), and OMatG Höllmer et al. (2025). These methods span architectures from variational autoencoders to stochastic interpolants Albergo et al. (2023) (encompassing diffusion models Song et al. (2021) and flow-matching Albergo & Vanden-Eijnden (2023); Lipman et al. (2023)). A separate line of work addresses unconstrained *de novo* generation, predicting both composition and structure jointly; this includes GNoME Merchant et al. (2023) and MatterGen Zeni et al. (2023), as well as variants of several methods mentioned above Höllmer et al. (2025); Miller et al. (2024).

A newer class of ML methods imposes physically-motivated inductive biases by exploiting the redundancy inherent in crystalline symmetries. Crystals possess translational and point-group symmetries that together describe a space group, which generates sets of symmetry-equivalent positions within the unit cell. When occupied by atoms, these positions are known as Wyckoff positions. Each Wyckoff position specifies both the multiplicity (how many symmetry-equivalent atoms it generates) and the degrees of freedom for atomic coordinates. For a given space group, operating directly on Wyckoff positions rather than on all atomic coordinates simplifies the learning objective by removing redundant degrees of freedom.

DiffCSP++ Jiao et al. (2024) takes this approach, requiring Wyckoff templates as input and learning only the continuous coordinates within those sites. Since ground-truth symmetry is unavailable, templates must be retrieved from the training set based on compositional similarity, restricting generation to known structural prototypes. SymmCD Levy et al. (2024), CrystalFormer Cao et al. (2025), and SGEquiDiff Chang et al. (2025) instead incorporate Wyckoff site selection as a discrete generative step, enabling the discovery of structural prototypes beyond those seen during training. Broadly, such symmetry-constrained models typically show improved space group distributions of

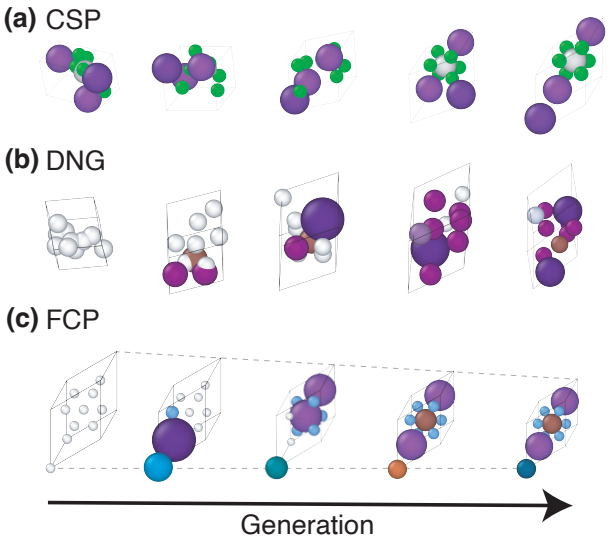

Figure 2: Illustration of the generation process with different types of generative models for the same test structure: **(a)** crystal structure prediction (CSP) samples the species at time $t = 0$ from the structure at time $t = 1$; **(b)** *de novo* generation (DNG) evolves the lattice, species, and fractional coordinates jointly; **(c)** framework-constrained prediction (FCP) evolves the lattice volume scaling and atomic species only.

predicted structures—towards more symmetric crystal structures—which in turn improves stable, unique, and novel (SUN) rates.

Finally, SCIGEN Okabe et al. (2025) takes a complementary approach, steering diffusion-based generation toward desired lattice motifs (such as kagome or honeycomb) by constraining the denoising process, and has led to the successful synthesis of two predicted materials.

1.2 Our Contribution

It is well-understood in materials science that two crystal structures sharing the same *structural framework* can exhibit different space group symmetries. A structural framework specifies the unit cell shape and the fractional coordinates of atomic sites, but not the chemical species occupying each site (see Sec. 2.1 and Fig. 1). In fact, partitioning sites among distinct species can reduce the symmetry to a subgroup of the framework's full geometric symmetry.

It is unknown whether symmetry-constrained generative models might benefit from operating at the structural framework level. Current models produce few novel structural prototypes Szymanski & Bartel (2025), and it remains unclear whether this limited novelty reflects shortcomings of the models or instead the fact that nature has already explored most viable frameworks. Indeed, the relatively small number of known structural prototypes—2050 entries in the AFLOW database Eckert et al. (2024); Hicks et al. (2021) at the time of writing—suggests that the space of geometrically stable arrangements is highly constrained. This motivates an alternative approach: rather than learning to generate novel atomic coordinates, one can leverage known frameworks and focus on predicting which species assignments yield stable materials.

We introduce OMatG-FC, the first 'framework-constrained' generative model, which conditions generation on structural frameworks drawn from known crystalline materials. Unlike symmetry-constrained methods, OMatG-FC fixes only the framework geometry; any additional realized symmetry-breaking emerges from species assignment rather than being imposed. This places our approach between symmetry-based methods, which allow geometric variation within space groups, and substitution-based methods, which start from a known compound and constrain candidate species by chemical similarity within the structural prototype. We demonstrate in Figure 2 how our model, which performs framework-constrained prediction (FCP), differs from standard CSP and DNG methods.

We show that our model learns the symmetries common in the dataset from species-to-site assignment by good performance across distributional benchmarks and SUN metrics. Moreover, our model demonstrates superior ability to predict stable, novel, and unique (SUN) structures. While not the focus of this paper, our approach is well-suited for the prediction of chemical disorder on crystallographic sites, an active area of research particularly for materials such as multi-component alloys Jakob et al. (2026); Antypov et al. (2025).

2 BACKGROUND

2.1 CRYSTAL STRUCTURE DEFINITIONS

A crystal structure is defined by its translational symmetry, meaning it can be represented compactly by a periodic repeating subunit—termed a unit cell. Because of the constraint of translational symmetry combined with point-group symmetry, crystals in three dimensions must obey the symmetries belonging to one of 230 space groups. An infinite number of crystal structures can exist within each space group, which defines the possible symmetry-inequivalent sites, known as Wyckoff positions, that can be occupied. To specify a particular crystalline material, one needs to define its unit cell: the particle positions, the chemical species occupying each position, and the lattice parameters. Defining all particle positions, however, is equivalent to defining the space group and all occupied Wyckoff positions, along with the choices of the free parameters (if extant) within them.

We define some terms that will be used throughout this manuscript.[1] The structural *prototype* is the *discrete* classification of a crystal structure by its space group, occupied Wyckoff positions, and stoichiometry. A prototype is a structural class: many materials with different chemical species and lattice parameters can belong to the same prototype. The AFLOW Prototype Encyclopedia Eckert et al. (2024); Hicks et al. (2021) follows this convention, encoding prototypes with labels such as `AB4C3_oI32_46_b_2bc_bc`, which specify the stoichiometry, Pearson symbol (combining the Bravais lattice and the number of atoms per conventional unit cell), space group number, and Wyckoff positions, respectively. AFLOW further indexes representative materials within each prototype using numerical suffixes (`-001`, `-002`), with each entry specifying the lattice parameters and Wyckoff coordinates for that material. A crystalline *material* is thus a specific instantiation of a prototype: it specifies which chemical species occupy each stoichiometric site (e.g., A = Ca, B = Ti, X = O for the ABX_3 perovskite $CaTiO_3$) along with the lattice parameters and atomic coordinates.

We define a structural *framework* as the geometric specification of a crystal independent of chemical identity: the unit cell shape (lattice parameters up to uniform scaling) together with the fractional coordinates of all atomic sites. A framework does not encode which species occupy these sites. Therefore, the same framework can correspond to multiple prototypes depending on species assignment. For example, BCC elemental metals (space group $Im\bar{3}m$, Wyckoff site $2a$, stoichiometry A) and CsCl (space group $Pm\bar{3}m$, Wyckoff sites $1a$ + $1b$, stoichiometry AB) share identical frameworks but constitute different prototypes, because the body-centered translation is a valid symmetry only when both sites are chemically equivalent (see Fig. 1).

2.2 STOCHASTIC INTERPOLANTS

Stochastic interpolants Albergo & Vanden-Eijnden (2023); Albergo et al. (2023) unify diffusion models Song et al. (2021) and conditional flow-matching Albergo & Vanden-Eijnden (2023); Lipman et al. (2023) under a single framework, bridging a base probability distribution ρ_0 with the true data distribution ρ_1 through samples ($x_0 \sim \rho_0, x_1 \sim \rho_1$). We consider interpolants of the form

$$x(t, x_0, x_1, z) = \alpha(t)x_0 + \beta(t)x_1 + \gamma(t)z, \tag{1}$$

where $t \in [0, 1]$ represents time such that $x(t = 0) = x_0$ and $x(t = 1) = x_1$. A learned time-dependent drift $b^\theta(x, t)$ and, optionally, a learned denoiser $z^\theta(x, t)$ give access to trajectories from the time-dependent density ρ_t through ODE-based or SDE-based sampling starting from $x_0 \sim \rho_0$. For training SDE-based sampling schemes, $\gamma(t) \neq 0$ for $t \in (0, 1)$ is required, while it is optional for ODE-based sampling.

[1] The language we define may not be universal.

For both ODE- and SDE-based sampling, the required velocity term $b^\theta(t,x)$ is learned by minimizing the loss function

$$\mathcal{L}_b(\theta) = \mathbb{E}_{t,z,x_0,x_1}\big[|b^\theta(t,x_t)|^2 - 2\,\partial_t x(t,x_0,x_1,z)\cdot b^\theta(t,x_t)\big], \tag{2}$$

with the expectation taken independently over $t \sim \mathrm{Uniform}\,[0,1]$, $z \sim \mathcal{N}(0,\boldsymbol{I})$, $x_0 \sim \rho_0$, and $x_1 \sim \rho_1$. For SDE-based sampling, an additional denoiser $z^\theta(t,x)$ must be learned by minimizing the additional loss

$$\mathcal{L}_z(\theta) = \mathbb{E}_{t,z,x_0,x_1}\left[|z^\theta(t,x_t)|^2 - 2\,z^\theta(t,x_t)\cdot z\right]. \tag{3}$$

We note that minimizing with respect to these loss functions is equivalent to minimizing a mean-squared error loss.

Prior work has explored stochastic interpolants in the context of crystal structure prediction and *de novo* materials generation in the OMatG model Höllmer et al. (2025). Following this and related work Miller et al. (2024); Bose et al. (2024), we apply inference-time velocity annealing which rescales the learned velocity field as $b^\theta(t,x) \to (1+st)\,b^\theta(t,x)$, where s is as an hyperparameter that increases the magnitude of the velocity with time.

2.3 Discrete Flow Matching

Discrete Flow Matching (DFM) Campbell et al. (2024); Gat et al. (2024) is a generative modeling framework for categorical variables which extends flow matching to discrete state spaces. Similarly to stochastic interpolants, DFM relies on connecting samples d_0 from the (discrete) base distribution to data samples d_1 from the data distribution with a conditional flow. Types of base distributions for sampling d_0 include uniform and completely masked. For the uniform conditional flow, the probability for the discrete variable d_t at time t given the target sample d_1 is given by $p^{\text{unif}}_{t|1}(d_t|d_1) = \mathrm{Cat}(t\,\delta\{d_1,d_t\} + (1-t)/S)$, where S is the number of possible states. This uniform conditional flow converges to d_1 at $t=1$ and is uniformly distributed over all possible states at $t=0$. Alternatively, the masked conditional flow is given by $p^{\text{mask}}_{t|1}(d_t|d_1) = \mathrm{Cat}(t\,\delta\{d_1,d_t\} + (1-t)\,\delta\{M,d_t\})$, which converges to an additional masking token M at $t=0$.

The conditional flows can be used to train a neural network to predict the denoising probability $p^\theta_{1|t}(d_1|d_t)$ of the final state d_1 given the current state d_t by minimizing a cross-entropy loss:

$$\mathcal{L}_{\mathrm{DFM}}(\theta) = \mathbb{E}_{t,d_1,d_0}\left[-\log p^\theta_{1|t}(d_1|d_t)\right]. \tag{4}$$

Once the denoising probability $p^\theta_{1|t}(d_1|d_t)$ is learned, one can construct the rate matrix R_t that determines the rate with which state d_t jumps to another state at time t during generation Campbell et al. (2024); Höllmer et al. (2025). In addition, it is possible to add stochasticity in the shape of a second rate matrix satisfying detailed balance, R^{DB}_t, with a tunable noise parameter η such that $R^\eta_t = R_t + \eta R^{DB}_t$.

3 Crystal Representation and Generation

A crystalline unit cell defines a distinct choice of periodic repeating unit which can be tiled to generate the infinite crystal structure. Generative models typically represent a unit cell with N atoms using three components: chemical species $\boldsymbol{A} \in \mathbb{Z}^N_{>0}$, lattice vectors $\boldsymbol{L} \in \mathbb{R}^{3\times3}$, and fractional coordinates $\boldsymbol{F} \in [0,1)^{3\times N}$. In this work, we further decompose the lattice vectors into unit-volume lattice vectors $\bar{\boldsymbol{L}} \in \mathbb{R}^{3\times3}$ where $\det(\bar{\boldsymbol{L}}) = 1$ and a log-volume scale factor $V = \log|\det(\boldsymbol{L})|$.

Our generative modeling framework seeks to generate samples from the joint conditional distribution $\rho_1\left(\boldsymbol{A}, V|\bar{\boldsymbol{L}},\boldsymbol{F}\right)$ where $\bar{\boldsymbol{L}}$ and $\boldsymbol{F}$ define a structural framework that is sampled from the dataset. We sample this joint distribution by co-evolving V by stochastic interpolants and $\mathbf{A}$ by discrete flow matching, with combined loss

$$\mathcal{L}(\theta) = \lambda^V_b\mathcal{L}_b(\theta) + \lambda^V_z\mathcal{L}_z(\theta) + \lambda^{\boldsymbol{A}}\mathcal{L}_{\mathrm{DFM}}(\theta) \tag{5}$$

where $\lambda^V_z = 0$ for ODE sampling schemes. Relative weights λ are optimized via hyperparameter tuning (see Sec. B).

3.1 CHEMICAL SPECIES

We represent chemical species by atomic number and model them using discrete flow matching, following OMatG's implementation. Each atom takes a value $\boldsymbol{a}_i \in \{1, 2, \ldots, 100\}$ and $M = 0$ is a mask token. The base distribution sets all atoms to the masked state $\rho_0(\boldsymbol{A}) = [M]^N$ and atomic identities evolve *via* a continuous-time Markov Chain and are progressively unmasked until all tokens are replaced at time $t = 1$. The learned denoising probability $p^{\theta}_{1|t}$, and hence the rate matrix, depend on all other state variables ($\boldsymbol{F}$, $\bar{\boldsymbol{L}}$, V) in addition to the atom types.

3.2 LATTICE VECTORS AND VOLUME SCALE

The lattice matrix $\boldsymbol{L}$ is defined by the three lattice vectors $[\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3]$ and can be decomposed into shape and scale: the unit-volume lattice $\bar{\boldsymbol{L}}$ and log-volume V, related by $\bar{\boldsymbol{L}} = \boldsymbol{L} * \exp(-V/3)$. We fix $\bar{\boldsymbol{L}}$ throughout generation using an identity interpolant with mirrored based distribution $\bar{\boldsymbol{L}}_t = \bar{\boldsymbol{L}}_0 = \bar{\boldsymbol{L}}_1$, meaning that samples x_0 are taken to have equivalent $\bar{\boldsymbol{L}}$ to the sampled structures $x_1 \sim \rho_1$.

For the volume scale V, we utilize a linear interpolant where $\alpha(t) = 1 - t$ and $\beta(t) = t$ with both an ODE ($\gamma(t) = 0$) and SDE sampling scheme ($\gamma(t) = \sqrt{at(1-t)}$ where a is a tunable parameter). As the stochastic interpolant framework permits flexible base distributions, we fit a log-normal distribution to unit cell volumes in the training set, yielding a Gaussian prior $\rho_0(V)$ on log-volume; previous work has shown that such priors improve generation performance Miller et al. (2024); Höllmer et al. (2025).

3.3 ATOMIC COORDINATES

Atomic positions live on a torus and are represented as fractional coordinates $\boldsymbol{F} \in [0, 1)^{3\times N}$, expressing each atom's position relative to the lattice vectors. As with the unit-volume lattice, we fix fractional coordinates throughout generation using an identity interpolant with mirrored base distribution, setting $\boldsymbol{F}_t = \boldsymbol{F}_1$ for all t, meaning that samples x_0 are taken to have equivalent $\boldsymbol{F}$ to the sampled structures $x_1 \sim \rho_1$.

4 METHODOLOGY

4.1 EQUIVARIANT MODEL FOR CRYSTAL STRUCTURES

In order to properly account for symmetries inherent to crystal structures, we utilize a CSPNet architecture based on that of Jiao et al. (2023) in the DiffCSP generative model. CSPNet is translation-invariant as well as rotation- and permutation-equivariant, meaning it correctly accounts for the symmetries which emerge as a result of treating the unit-cell as a periodic object. The full lattice matrix $\boldsymbol{L}$ is reconstructed and provided to the model as input.

The architecture uses learnable atomic embeddings and represents fractional coordinates through sinusoidal positional embeddings. It uses six message-passing layers with a fully-connected graph neural network with atomic positions as nodes. Node embeddings $\boldsymbol{h}^i_{(s)}$ are initialized as a function of the atom types $\boldsymbol{a}$ and updated through a message-passing scheme in which nearby node embeddings, the full lattice information $\boldsymbol{L}$, and sinusoidal distance embeddings are injected. This scheme is given

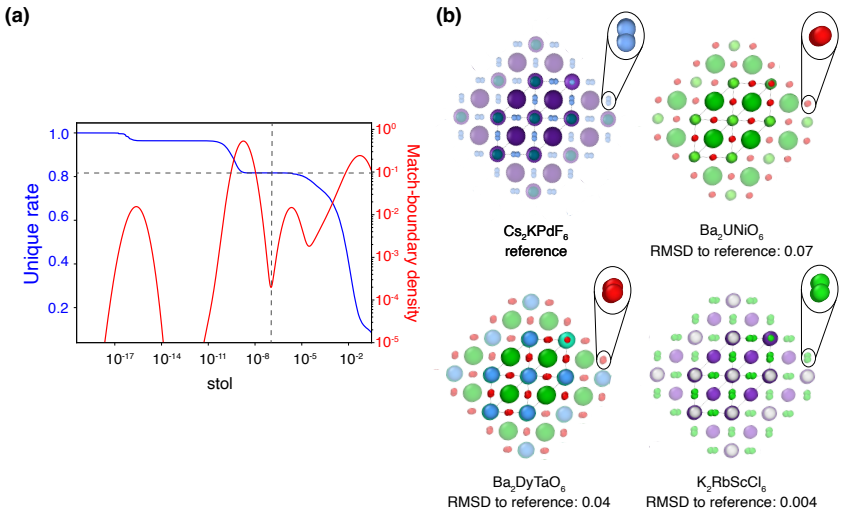

Figure 3: **(a)** Uniqueness rate and kernel density estimate (using a Gaussian kernel) of the match-boundary density as a function of `stol` of `StructureMatcher` plotted for the entire MP-20 dataset. **(b)** Examples of three structural frameworks with different compositions which 'match' by framework to the reference Cs_2KPdF_6 structure, with the RMSDs and structural differences on one of the symmetry sites magnified for each.

by

$$
\begin{aligned}
\boldsymbol{h}^i_{(0)} &= \phi_{\boldsymbol{h}_{(0)}}(\boldsymbol{a}_i) \\
\boldsymbol{m}^{ij}_{(s)} &= \varphi_m\left(\boldsymbol{h}^i_{s-1}, \boldsymbol{h}^j_{s-1}, \boldsymbol{l}, \text{SinEmb}\left(\boldsymbol{x}^j - \boldsymbol{x}^i\right)\right) \\
\boldsymbol{m}^i_{(s)} &= \sum_{j=1}^{N} \boldsymbol{m}^{ij}_{(s)} \\
\boldsymbol{h}^i_{(s)} &= \boldsymbol{h}^i_{(s-1)} + \varphi_h\left(\boldsymbol{h}_{(s-1)}, \boldsymbol{m}^i_{(s)}\right) \\
b_V &= \varphi_V\left(\boldsymbol{h}^i_{(\max s)}\right) \\
b_{\boldsymbol{A}} &= \varphi_{\boldsymbol{A}}\left(\boldsymbol{h}^i_{(\max s)}\right)
\end{aligned}
\tag{6}
$$

wherein φ_m is a parameterized neural network. All necessary drift $b^\theta(t,x)$ and denoiser $z^\theta(t,x)$ terms are computed from linear layers applied to the hidden final layer node embeddings (of dimension 1 for the volume scale and dimension 101 for the atomic species). The network also predicts $\log p^\theta_{1|t}(\boldsymbol{a}_1|x_t)$ for the atomic species.

5 Datasets

For this work, we utilize the ***MP*-20** Jain et al. (2013); Xie et al. (2022) dataset from the Materials Project that contains 45 229 structures with at maximum $N = 20$ atoms per unit cell. This dataset utilizes a 60–20–20 train-validation-test split.

5.1 Framework Duplicate Analysis

We investigate structural framework duplicates within MP-20, demonstrating that frameworks are commonly shared across different materials—motivating our framework-constrained approach. Duplicate analysis follows Martirossyan et al. (2025) using `StructureMatcher` (see Appendix D for details), though we perform analysis only using `stol`. Structure pair comparisons are made to

find the `stol` tolerance at which they transition from matching to non-matching, which is equivalent to the RMSD between atomic positions at this transition.

All pairs of structures are compared to one another using PyMatGen's `StructureMatcher` with the default tolerances: the site tolerance `stol`$= 0.3$, lattice length tolerance of `ltol`$= 0.2$ and angle tolerance of `angle_tol`$= 5.0$. Several non-standard arguments are added: the `primitive_cell` argument is set to False to avoid reduction to primitive cells for matching; `scale` is set to True, allowing for identical structural frameworks with different volume scales to match; and the comparator ignores the identity of the chemical species. Our setup thus compares the structural frameworks only.

Figure 3 reveals large peaks in match-boundary density at `stol` $\approx 1 \times 10^{-9}$ and $\approx 1 \times 10^{-16}$, identifying exact framework duplicates. These structures differ only by what we posit to be from dataset preparation—numerical noise (possibly from structure symmetrization) or stopping criteria for DFT-based structure relaxation. These duplicates account for $\sim 18\%$ of MP-20 materials. The match boundary density also reveals multiple valleys (one at 1×10^{-7} and another at 1×10^{-4}) which would both suffice as duplicate cutoff values of `stol`, with only a modest difference in the fraction of unique structures between them. At `stol` tolerances above 1×10^{-4}, structures are strikingly similar, but not duplicates. Visual inspection (Fig. 3b) illustrates how Ba_2DyTaO_6 and $K_2RbScCl_6$ share similar geometric motifs as the reference Cs_2KPdF_6 with completely different chemistries and slightly varying structure. It is worth noting, however, that at the default tolerance for `StructureMatcher` of `stol`$= 0.3$ only $\sim 10\%$ structural frameworks would be considered unique.

In summary, the analysis shows that the inductive bias underlying OMatG-FC is well-motivated by the strong similarities between structural frameworks present in the dataset samples.

6 Benchmarks

6.1 Distributional Properties

To evaluate whether our model has learned the data distribution, we compare distributional properties of generated structures against the test set using Wasserstein distance. The properties we compare include mass density ρ, N-arity (number of unique elements), averaged coordination number by element $\langle CN \rangle_S$, element counts, and space groups. We also report validity metrics for the generated structures, including structural validity (all bond lengths > 0.5Å), compositional validity as measured by the SMACT software (version 3.0), and the combined validity requiring both criteria to be satisfied.

6.2 Compositional Overlap

Finally, we examine the distribution of compositional (element) overlaps between structure pairs to assess the propensity of our model to explore new chemical compounds. For each structure pair, we first determine whether the frameworks match using `StructureMatcher` as described in Sec. 5.1; if so, we record the fraction of shared elements. Since a single framework can admit multiple plausible compositions, we use a "match-everyone-to-reference" approach Martirossyan et al. (2025), retaining only the highest compositional overlap among all matching frameworks for each reference structure.

6.3 Stability, Uniqueness, and Novelty

We use the benchmarking framework LeMat-GenBench Betala et al. (2026) to compare OMatG-FC performance on standardized metrics against other recent generative models. This benchmarking suite has the advantage of a public leaderboard for ease of comparison and unified benchmarks, and uses three machine-learned interatomic potentials—MACE Batatia et al. (2025), UMA Wood et al. (2025), and Orb-v3 Rhodes et al. (2025)—as a surrogate for DFT relaxation to compute stability rate, the energy above hull, and the root-mean-squared-distance (RMSD with Å units) between the generated and relaxed structures. A validity metric screens structures based on a number of parameters measuring physical plausibility, including CIF readability, structural, density, and lattice parameter validity, symmetry, and charge neutrality. Uniqueness and novelty are assessed using the disorder-adapted `StructureMatcher` fingerprint method as developed by MatterGen Zeni et al. (2025), which treats chemically-disordered versions of the same structure as being equivalent. Generated structures

Table 1: Distribution-based metrics from *de novo* generation of 10 000 structures using models trained on the MP-20 dataset with the structural frameworks present in the MP-20 test dataset. We report comparisons to other *de novo* generation models: DiffCSP, FlowMM, and OMatG-EncDecODE. Best scores in each category are bolded and second-best are underlined.

Method	**Integration steps**	**Validity** (% ↑)			**Property** (↓)				
		Structural	Composition	Combined	wdist (ρ)	wdist (Nary)	wdist ($\langle CN \rangle_S$)	wdist (elem)	wdist (sg)
DiffCSP	1000	<u>99.91</u>	82.46	82.46	0.3133	0.3193	0.3916	1.6300	20.40
FlowMM	1000	92.26	76.87	76.86	1.0712	0.1130	0.6603	0.9255	13.02
OMatG-EncDecODE	840	97.25	84.16	84.16	<u>0.1155</u>	**0.0553**	**0.1026**	<u>0.8249</u>	24.80
OMatG-FC-LinearODE	610	**100.0**	**87.38**	**87.38**	0.6045	<u>0.0719</u>	0.1890	**0.4038**	**4.413**
OMatG-FC-LinearSDE	620	**100.0**	<u>86.33</u>	<u>86.32</u>	**0.0860**	0.0856	<u>0.1163</u>	0.8483	<u>5.888</u>

Table 2: Benchmarks for *de novo* generation as provided by the LeMat-GenBench software Betala et al. (2026) on 2500 generated structures. Stability is defined as having ≤ 0 eV/atom above hull and metastability is defined as ≤ 0.1 eV/atom above hull. Uniqueness and novelty results are combined with (meta)stability assessments to produce (M)SUN rates. For OMatG and OMatG-FC models, structures are subsampled from those in Tab. 1 and benchmarks are computed, otherwise benchmarks from other models are reproduced from the LeMat-GenBench leaderboard. Models marked with a † utilize pre-relaxation before assessment, and therefore their results cannot directly be compared. Best scores in each category (excluding pre-relaxed structures) are bolded and second-best are underlined.

Method	Validity Rate (%, ↑)	Unique Rate (%, ↑)	Novelty Rate (%, ↑)	Stable Rate (%, ↑)	Metastable Rate (%, ↑)	S.U.N. Rate (%, ↑)	M.S.U.N. Rate (%, ↑)	$\langle E \rangle / N$ (eV, ↓)	RMSD (Å, ↓)
ADiT	90.6	87.8	26.0	0.4	36.5	0.0	1.0	0.33	0.38
DiffCSP	95.7	<u>94.8</u>	**66.2**	<u>2.3</u>	29.8	0.1	**8.5**	<u>0.27</u>	0.59
DiffCSP++	95.3	**95.1**	<u>62.0</u>	1.0	26.4	<u>0.2</u>	5.0	0.41	0.69
SymmCD	73.4	73.0	47.0	1.4	18.6	0.1	2.4	0.88	0.87
OMatG-EncDecODE	92.4	91.6	55.1	0.4	30.3	0.1	3.6	0.31	0.61
OMatG-FC-LinearODE	<u>95.9</u>	92.2	30.4	2.0	<u>49.8</u>	**0.4**	5.9	<u>0.27</u>	<u>0.31</u>
OMatG-FC-LinearSDE	**96.5**	93.4	34.9	**3.6**	**56.4**	<u>0.2</u>	<u>6.8</u>	**0.17**	**0.30**
MatterGen†	95.7	95.1	70.5	2.0	33.4	0.2	15.0	0.18	0.39
PLaID++†	96.0	77.8	24.2	12.4	60.7	1.0	7.6	0.085	0.13

are compared against a reference dataset, where structures are unique if they are non-redundant within the generated set, and structures are novel if they are absent from the reference dataset. Stability is defined in terms of the energy above the convex hull (E_{hull}), where stable structures satisfy $E_{\text{hull}} \leq 0$ eV/atom and metastable structures satisfy $0 < E_{\text{hull}} \leq 0.1$ eV/atom. Combined, these metrics define S.U.N (stable, unique, and novel) and M.S.U.N. (metastable, unique, and novel) rates which are the union of the three properties for each structure. We follow the LeMat-GenBench leaderboard standard and compute these metrics on 2,500 structures sampled from our generated test set.

7 RESULTS

We present distributional metrics for generated structures in Table 1. OMatG-FC performs competitively on most distributional benchmarks, with notable improvement on the space group Wasserstein distance (see Appendix C for more detail). This improvement arises because fixed positions and lattice parameters yield accurate symmetries by construction, whereas non-constrained models produce positional noise that requires relaxed symmetry tolerances for space group identification. For the same reason, structural validity is 100%.

De novo generation metrics from LeMat-GenBench are shown in Table 2. OMatG-FC achieves state-of-the-art performance in SUN rate, average energy above hull $\langle E \rangle / N$, and RMSD between generated and relaxed structures. This is expected, as structural frameworks from the dataset guide generated crystals closer to stable atomic packings. Our models also achieve the highest validity scores, which here include structural validity but also the metrics described in Sec. 6.3. Finally, we demonstrate high stability and metastability rates, which compensate for lower novelty rates (similarly to the PLaID++ model Xu et al. (2025)). Compositional overlap results are included in Appendix A.

8 Discussion and Conclusion

We introduce a new type of generative model, OMatG-FC, which allows for the exploration of novel chemical compositions on structural frameworks. While counterintuitive for the prospect of discovering novel types of materials, we show that our method is a state-of-the-art *de novo* generative model. This is accomplished through simplification of the learning task, as well as the inductive bias imposed by the structural framework constraint, which pushes the learned distribution towards crystals that are closer to the convex hull (*i.e.*, more stable). Ultimately, OMatG-FC paves the way for discovery *via* structural frameworks—limited in our study to those from the dataset, but in practice applicable to any theorized point pattern.

Acknowledgments

The authors would like to thank Shenglong Wang at NYU IT HPC for supporting this work. The authors acknowledge funding from NSF Grant OAC-2311632. P. H. and S. M. also acknowledge support from the Simons Center for Computational Physical Chemistry (Simons Foundation grant 839534, MT). The authors gratefully acknowledge the computational resources, consultation support, and core personnel that have contributed to the research results reported in this publication, provided by the IT High Performance Computing at New York University and NYU Langone High Performance Computing.

REFERENCES

Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A Next-generation Hyperparameter Optimization Framework, July 2019. URL http://arxiv.org/abs/1907.10902.

Michael S. Albergo and Eric Vanden-Eijnden. Building Normalizing Flows with Stochastic Interpolants, March 2023. URL http://arxiv.org/abs/2209.15571.

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic Interpolants: A Unifying Framework for Flows and Diffusions, November 2023. URL http://arxiv.org/abs/2303.08797.

Dmytro Antypov, Chris M. Collins, Matthew S. Dyer, John B. Claridge, and Matthew J. Rosseinsky. Classification and statistical analysis of structural disorder in crystalline materials. *Journal of Applied Crystallography*, 58(3):659–677, June 2025. ISSN 1600-5767. doi: 10.1107/S1600576725003000. URL https://journals.iucr.org/paper?S1600576725003000.

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Filippo Bigi, Samuel M. Blau, Vlad Cărare, Michele Ceriotti, Sanggyu Chong, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Edvin Fako, Fabio Falcioni, Andrea C. Ferrari, John L. A. Gardner, Mikołaj J. Gawkowski, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Jonas Grandel, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Cheuk Hin Ho, Stephan Hofmann, Christian Holm, Jad Jaafar, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Panagiotis Kourtis, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Chen Lin, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Louise A. M. Rosset, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Christopher Sutton, Viktor Svahn, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Santiago Vargas, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Thomas Wolf, Fabian Zills, and Gábor Csányi. A foundation model for atomistic materials chemistry. *J. Chem. Phys.*, 163(18):184110, November 2025. ISSN 0021-9606. doi: 10.1063/5.0297006. URL https://doi.org/10.1063/5.0297006.

Siddharth Betala, Samuel P. Gleason, Ali Ramlaoui, Andy Xu, Georgia Channing, Daniel Levy, Clémentine Fourrier, Nikita Kazeev, Chaitanya K. Joshi, Sékou-Oumar Kaba, Félix Therrien, Alex Hernandez-Garcia, Rocío Mercado, N. M. Anoop Krishnan, and Alexandre Duval. LeMat-GenBench: A Unified Evaluation Framework for Crystal Generative Models, January 2026. URL http://arxiv.org/abs/2512.04562.

Lilia Boeri, Richard Hennig, Peter Hirschfeld, Gianni Profeta, Antonio Sanna, Eva Zurek, Warren E Pickett, Maximilian Amsler, Ranga Dias, Mikhail I Eremets, Christoph Heil, Russell J Hemley, Hanyu Liu, Yanming Ma, Carlo Pierleoni, Aleksey N Kolmogorov, Nikita Rybin, Dmitry Novoselov, Vladimir Anisimov, Artem R Oganov, Chris J Pickard, Tiange Bi, Ryotaro Arita, Ion Errea, Camilla Pellegrini, Ryan Requist, E K U Gross, Elena Roxana Margine, Stephen R Xie, Yundi Quan, Ajinkya Hire, Laura Fanfarillo, G R Stewart, J J Hamlin, Valentin Stanev, Renato S Gonnelli, Erik Piatti, Davide Romanin, Dario Daghero, and Roser Valenti. The 2021 room-temperature superconductivity roadmap. *J. Phys.: Condens. Matter*, 34(18):183002, March 2022. ISSN 0953-8984. doi: 10.1088/1361-648X/ac2864. URL https://dx.doi.org/10.1088/1361-648X/ac2864.

Cyprien Bone, Matthew Walker, Kuangdai Leng, Luis M. Antunes, Ricardo Grau-Crespo, Amil Aligayev, Javier Dominguez, and Keith T. Butler. Discovery and recovery of crystalline materials with property-conditioned transformers, November 2025. URL http://arxiv.org/abs/2511.21299.

Avishek Joey Bose, Tara Akhound-Sadegh, Guillaume Huguet, Kilian Fatras, Jarrid Rector-Brooks, Cheng-Hao Liu, Andrei Cristian Nica, Maksym Korablyov, Michael Bronstein, and Alexander Tong. SE(3)-Stochastic Flow Matching for Protein Backbone Generation, April 2024. URL http://arxiv.org/abs/2310.02391.

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. URL http://arxiv.org/abs/2402.04997.

Zhendong Cao, Xiaoshan Luo, Jian Lv, and Lei Wang. Space Group Informed Transformer for Crystalline Materials Generation. *Science Bulletin*, 70(21):3522–3533, November 2025. ISSN 20959273. doi: 10.1016/j.scib.2025.09.035. URL http://arxiv.org/abs/2403.15734.

Rees Chang, Angela Pak, Alex Guerra, Ni Zhan, Nick Richardson, Elif Ertekin, and Ryan P. Adams. Space Group Equivariant Crystal Diffusion, May 2025. URL http://arxiv.org/abs/2505.10994.

Anthony K. Cheetham and Ram Seshadri. Artificial Intelligence Driving Materials Discovery? Perspective on the Article: Scaling Deep Learning for Materials Discovery. *Chem. Mater.*, 36 (8):3490–3495, April 2024. ISSN 0897-4756. doi: 10.1021/acs.chemmater.4c00643. URL https://doi.org/10.1021/acs.chemmater.4c00643.

Stefano Curtarolo, Gus L. W. Hart, Marco Buongiorno Nardelli, Natalio Mingo, Stefano Sanvito, and Ohad Levy. The high-throughput highway to computational materials design. *Nat. Mater.*, 12(3):191–201, March 2013. ISSN 1476-4660. doi: 10.1038/nmat3568. URL https://www.nature.com/articles/nmat3568.

Pierre-Paul De Breuck, Hai-Chen Wang, Gian-Marco Rignanese, Silvana Botti, and Miguel A. L. Marques. Generative AI for crystal structures: a review. *npj Computational Materials*, 11(1): 370, December 2025. ISSN 2057-3960. doi: 10.1038/s41524-025-01881-2. URL https://www.nature.com/articles/s41524-025-01881-2.

Hagen Eckert, Simon Divilov, Michael J. Mehl, David Hicks, Adam C. Zettel, Marco Esters, Xiomara Campilongo, and Stefano Curtarolo. The AFLOW Library of Crystallographic Prototypes: Part 4. *Computational Materials Science*, 240:112988, May 2024. ISSN 09270256. doi: 10.1016/j.commatsci.2024.112988. URL http://arxiv.org/abs/2401.06875.

Itai Gat, Tal Remez, Neta Shaul, Felix Kreuk, Ricky T. Q. Chen, Gabriel Synnaeve, Yossi Adi, and Yaron Lipman. Discrete Flow Matching, July 2024. URL http://arxiv.org/abs/2407.15595.

Bernd Gludovatz, Anton Hohenwarter, Keli V. S. Thurston, Hongbin Bei, Zhenggang Wu, Easo P. George, and Robert O. Ritchie. Exceptional damage-tolerance of a medium-entropy alloy CrCoNi at cryogenic temperatures. *Nat. Commun.*, 7(1):10602, February 2016. ISSN 2041-1723. doi: 10.1038/ncomms10602. URL https://www.nature.com/articles/ncomms10602.

Geoffroy Hautier, Chris Fischer, Virginie Ehrlacher, Anubhav Jain, and Gerbrand Ceder. Data Mined Ionic Substitutions for the Discovery of New Compounds. *Inorg. Chem.*, 50(2):656–663, January 2011. ISSN 0020-1669. doi: 10.1021/ic102031h. URL https://doi.org/10.1021/ic102031h.

David Hicks, Cormac Toher, Denise C. Ford, Frisco Rose, Carlo De Santo, Ohad Levy, Michael J. Mehl, and Stefano Curtarolo. AFLOW-XtalFinder: A reliable choice to identify crystalline prototypes. *npj Computational Materials*, 7(1):30, February 2021. ISSN 2057-3960. doi: 10.1038/s41524-020-00483-4. URL https://www.nature.com/articles/s41524-020-00483-4.

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. URL https://openreview.net/forum?id=gHGrzxFujU.

William Hume-Rothery. *The Structure of Metals and Alloys*. Chemical Publishing Company, 1939.

Anubhav Jain, Geoffroy Hautier, Charles Moore, Shyue Ong, Chris Fischer, Tim Mueller, Kristin Persson, and Gerbrand Ceder. A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.*, 50:2295–2310, June 2011. doi: 10.1016/j.commatsci.2011.02.023.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, July 2013. doi: 10.1063/1.4812323. URL https://aip.scitation.org/doi/10.1063/1.4812323.

Konstantin S. Jakob, Aron Walsh, Karsten Reuter, and Johannes T. Margraf. Learning Crystallographic Disorder: Bridging Prediction and Experiment in Materials Discovery. *Advanced Materials*, 38(5):e14226, January 2026. ISSN 0935-9648, 1521-4095. doi: 10.1002/adma.202514226. URL https://advanced.onlinelibrary.wiley.com/doi/10.1002/adma.202514226.

Ho-Young Jang, Donggun Eum, Jiung Cho, Jun Lim, Yeji Lee, Jun-Hyuk Song, Hyeokjun Park, Byunghoon Kim, Do-Hoon Kim, Sung-Pyo Cho, Sugeun Jo, Jae Hoon Heo, Sunyoung Lee, Jongwoo Lim, and Kisuk Kang. Structurally robust lithium-rich layered oxides for high-energy and long-lasting cathodes. *Nature Communications*, 15(1):1288, February 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-45490-x. URL https://www.nature.com/articles/s41467-024-45490-x.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal Structure Prediction by Joint Equivariant Diffusion, July 2023. URL https://arxiv.org/abs/2309.04475v2.

Rui Jiao, Wenbing Huang, Yu Liu, Deli Zhao, and Yang Liu. Space Group Constrained Crystal Generation, April 2024. URL http://arxiv.org/abs/2402.03992.

Noriaki Kamaya, Kenji Homma, Yuichiro Yamakawa, Masaaki Hirayama, Ryoji Kanno, Masao Yonemura, Takashi Kamiyama, Yuki Kato, Shigenori Hama, Koji Kawamoto, and Akio Mitsui. A lithium superionic conductor. *Nature Materials*, 10(9):682–686, September 2011. ISSN 1476-1122, 1476-4660. doi: 10.1038/nmat3066. URL https://www.nature.com/articles/nmat3066.

Daniel Levy, Siba Smarak Panigrahi, Sékou-Oumar Kaba, Qiang Zhu, Mikhail Galkin, Santiago Miret, and Siamak Ravanbakhsh. SymmCD: Symmetry-Preserving Crystal Generation with Diffusion Models. In *AI for Accelerated Materials Design - NeurIPS 2024*, November 2024. URL https://openreview.net/forum?id=V7x2KZQn2v.

Richard Liaw, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E Gonzalez, and Ion Stoica. Tune: A research platform for distributed model selection and training. *arXiv preprint arXiv:1807.05118*, 2018.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow Matching for Generative Modeling, February 2023. URL http://arxiv.org/abs/2210.02747.

Maya Martirossyan, Thomas Egg, Philipp Höllmer, George Karypis, Mark Transtrum, Adrian Roitberg, Mingjie Liu, Richard Hennig, Ellad B. Tadmor, and Stefano Martiniani. All that structure matches does not glitter. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2025. URL https://openreview.net/forum?id=ig9ujp50D4.

Amil Merchant, Simon Batzner, Samuel S. Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624:80–85, November 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06735-9. URL https://www.nature.com/articles/s41586-023-06735-9.

Houssam Metni, Laura Ruple, Lauren N. Walters, Luca Torresi, Jonas Teufel, Henrik Schopmans, Jona Östreicher, Yumeng Zhang, Marlen Neubert, Yuri Koide, Kevin Steiner, Paul Link, Lukas Bär, Mariana Petrova, Gerbrand Ceder, and Pascal Friederich. Generative models for crystalline materials, November 2025. URL http://arxiv.org/abs/2511.22652.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: Generating Materials with Riemannian Flow Matching, June 2024. URL http://arxiv.org/abs/2406.04713.

Kenji Nomura, Hiromichi Ohta, Akihiro Takagi, Toshio Kamiya, Masahiro Hirano, and Hideo Hosono. Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors. *Nature*, 432(7016):488–492, November 2004. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature03090. URL https://www.nature.com/articles/nature03090.

Ryotaro Okabe, Mouyang Cheng, Abhijatmedhi Chotrattanapituk, Manasi Mandal, Kiran Mak, Denisse Córdova Carrizales, Nguyen Tuan Hung, Xiang Fu, Bowen Han, Yao Wang, Weiwei Xie, Robert J. Cava, Tommi S. Jaakkola, Yongqiang Cheng, and Mingda Li. Structural constraint integration in a generative model for the discovery of quantum materials. *Nature Materials*, pp. 1–8, September 2025. ISSN 1476-4660. doi: 10.1038/s41563-025-02355-y. URL https://www.nature.com/articles/s41563-025-02355-y.

Chris J Pickard and R J Needs. *Ab Initio* random structure searching. *J. Phys.: Condens. Matter*, 23(5): 053201, February 2011. ISSN 0953-8984, 1361-648X. doi: 10.1088/0953-8984/23/5/053201. URL https://iopscience.iop.org/article/10.1088/0953-8984/23/5/053201.

Prof C. J. Pickard. AIRSS, October 2016. URL https://www.mtg.msm.cam.ac.uk/Codes/AIRSS.

Pawan Prakash, Jason B. Gibson, Zhongwei Li, Gabriele Di Gianluca, Juan Esquivel, Eric Fuemmeler, Benjamin Geisler, Jung Soo Kim, Adrian Roitberg, Ellad B. Tadmor, Mingjie Liu, Stefano Martiniani, Gregory R. Stewart, James J. Hamlin, Peter J. Hirschfeld, and Richard G. Hennig. Guided Diffusion for the Discovery of New Superconductors, September 2025. URL http://arxiv.org/abs/2509.25186.

Benjamin Rhodes, Sander Vandenhaute, Vaidotas Šimkus, James Gin, Jonathan Godwin, Tim Duignan, and Mark Neumann. Orb-v3: Atomistic simulation at scale, April 2025. URL http://arxiv.org/abs/2504.06231.

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. URL http://arxiv.org/abs/2011.13456.

Nathan J. Szymanski and Christopher J. Bartel. Establishing baselines for generative discovery of inorganic crystals. *Materials Horizons*, 12(19):8000–8011, 2025. doi: 10.1039/D5MH00010F. URL https://pubs.rsc.org/en/content/articlelanding/2025/mh/d5mh00010f.

Atsushi Togo, Kohei Shinohara, and Isao Tanaka. Spglib: A software library for crystal symmetry search, March 2024. URL http://arxiv.org/abs/1808.01590.

Tien-Sinh Vu, Minh-Quyet Ha, Duong-Nguyen Nguyen, Viet-Cuong Nguyen, Yukihiro Abe, Truyen Tran, Huan Tran, Hiori Kino, Takashi Miyake, Koji Tsuda, and Hieu-Chi Dam. Towards understanding structure–property relations in materials with interpretable deep learning. *npj Computational Materials*, 9(1):215, December 2023. ISSN 2057-3960. doi: 10.1038/s41524-023-01163-9. URL https://www.nature.com/articles/s41524-023-01163-9.

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. URL http://arxiv.org/abs/2506.23971.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal Diffusion Variational Autoencoder for Periodic Material Generation, March 2022. URL http://arxiv.org/abs/2110.06197.

Andy Xu, Rohan Desai, Larry Wang, Gabriel Hope, and Ethan Ritz. PLaID++: A Preference Aligned Language Model for Targeted Inorganic Materials Design, September 2025. URL http://arxiv.org/abs/2509.07150.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, Ryota Tomioka, and Tian Xie. MatterGen: A generative model for inorganic materials design, December 2023. URL http://arxiv.org/abs/2312.03687.

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*, pp. 1–3, January 2025. ISSN 1476-4687. doi: 10.1038/s41586-025-08628-5. URL https://www.nature.com/articles/s41586-025-08628-5.

A Composition Overlap Results

We evaluate compositional overlap between matching frameworks by plotting the distribution of compositional overlap percentages in Figure 4 using the method described in Sec. 6.2. Matching structural frameworks typically share some chemical species, but many generated structures have compositions absent from matching frameworks in the test set, suggesting that the model explores novel chemical space.

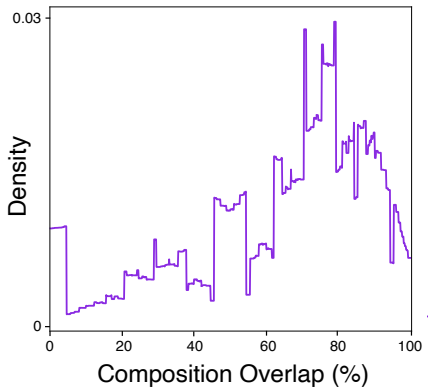

Figure 4: Kernel density estimate (using a tophat kernel) of the composition overlap between generated structures using the OMatG-FC (with LinearODE) model and the reference dataset in a match-everyone-to-reference fashion. Partial overlap with best-matching reference structures indicates partial adherence to known compositions as well as a propensity for compositional novelty.

B Hyperparameter tuning

We perform hyperparameter tuning of the models based on the linear interpolant for both ODE and SDE sampling schemes, using the `Ray Tune` package Liaw et al. (2018) with the Optuna library Akiba et al. (2019). For training, we tune the relative loss weights λ, the batch size, the learning rate, the weight decay and, $\gamma(t)$ (for SDE). For inference, we tune the number of integration steps, the species noise η, the time-dependent noise $\varepsilon(t)$ (for SDE), and the magnitude of the velocity annealing parameter s for the volume scale. We minimize with respect to the following evaluation metric:

$$\begin{aligned}\text{eval}_{\text{FCP}} = \text{avg}\Big[&1 - \text{combined validity},\\ &\text{wdist}(\rho), \text{wdist}(N\text{ary}), \text{wdist}(\langle CN\rangle_S),\\ &\text{wdist}(\text{elem}), \text{wdist}(\text{space group})\Big].\end{aligned} \tag{7}$$

C Space Group Identification

In Table 1, the space group Wasserstein distances are evaluated *via* space group identification with `spglib` Togo et al. (2024). Model performance is evaluated using a variable precision protocol introduced by OMatG or by setting `symprec`=1×10^{-2}. The results in Table 1 reflect in the lowest wdist(sg) values for each model from the precision protocol choices. For each fixed choice of tolerance handling, OMatG-FC models outperform others.

D Structure Matcher

PyMatGen's `StructureMatcher` is a commonly used tool to compute match rates between generated and reference structures. The algorithm determines whether two structures match using three tolerances: the lattice (`ltol`) and angle tolerance (`angle_tol`), which control the allowed

deviation in lattice lengths and angles, respectively; and the site tolerance, `stol`, which controls the allowed volume-normalized atomic-site displacements based on the average free length per atom, $(V/N)^{\frac{1}{3}}$, where V is the unit cell volume and N is the number of atoms. Matching typically proceeds by reducing structures to primitive cells (we however, set this to False), aligning lattice vectors within tolerances, and computing a normalized per-atom root-mean-square-distance (RMSD) between atomic positions.