FIELDFORGE: A PRETRAINING AND POST-TRAINING FRAMEWORK FOR FLOW-BASED CRYSTAL GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Generating thermodynamically competitive crystals while retaining novelty remains a central challenge in materials discovery. To address this challenge, we present FieldForge, a flow-based training framework that unifies pretraining (PT), supervised fine-tuning (SFT), and reinforcement learning (RL) within a single crystal generator. Pretraining combines a Transformer supporting variable atom counts with flow matching, while respecting crystal geometry throughout the generative process. SFT then adapts the pretrained generative distribution through controlled data selection, preparing the generator for RL without introducing additional training structures. RL uses a GRPO-inspired critic-free policy update to refine the same velocity field using stability, novelty, and diversity rewards, with stability estimates provided by an interatomic potential. On the MP-20 track of LeMat-GenBench, the pretrained FieldForge model achieves state-of-the-art stable-unique-novel (SUN) and metastable-unique-novel (MSUN) rates, combining thermodynamic quality with uniqueness and novelty. SFT followed by RL further improves generation performance on both MP-20 and Alex-MP-20, demonstrating the effectiveness of the complete training framework.

1 INTRODUCTION

Discovering new inorganic materials requires generating crystals that are both thermodynamically competitive and distinct from known structures. Deep generative models can learn complex composition–structure distributions (Xie et al., 2022; Zeni et al., 2025), while machine-learned interatomic potentials enable efficient screening (Batzner et al., 2022). Yet a trade-off between thermodynamic quality and novelty can limit the generation of crystals that satisfy both criteria.

Addressing this trade-off requires an expressive pretrained generator that captures the training distribution, followed by post-training refinement toward discovery objectives. Recent work demonstrates reward-based refinement of pretrained crystal generators (Park & Walsh, 2026; Höllmer & Martiniani, 2026; Su et al., 2026). A coordinated training framework should connect these capabilities, providing a strong generative foundation and a systematic approach to jointly improving thermodynamic quality and novel coverage.

We introduce FieldForge, which integrates pretraining (PT), supervised fine-tuning (SFT), and reinforcement learning (RL) within a single Transformer velocity field. Our pretraining incorporates periodicity and lattice geometry into flow paths and crystal-periodic rotary attention, providing geometric inductive biases within a standard Transformer backbone. A null-species token supports variable atom counts and novel coverage (Okhotin et al., 2025). Trained through flow matching (Liu et al., 2023; Lipman et al., 2023), this backbone provides a shared foundation for SFT and RL without architectural changes.

SFT bridges generative pretraining and reward optimization through a controlled subset of the original corpus, without introducing additional training structures. Quality-balanced SFT (QB-SFT) adjusts the stable–metastable balance through the on-hull fraction of its training subset, providing a better initialization for subsequent RL optimization. RL reuses the pretraining noise design for stochastic rollout exploration and applies critic-free group-relative updates to the same velocity field (Shao et al., 2024). A composite terminal reward combines interatomic-potential-based stability estimates

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

with rule- and descriptor-based signals for chemical validity, novelty, and diversity. Starting from the fine-tuned checkpoint, RL increases novelty while largely preserving metastability, moving beyond the observed pretraining trade-off.

At inference, anti-annealed sampling (Miller et al., 2024) amplifies late-stage coordinate updates, increasing the fraction of generated crystals that are simultaneously metastable, unique, and novel across the evaluated checkpoints without retraining.

On the MP-20 track of LeMat-GenBench, pretrained FieldForge achieves state-of-the-art stable-unique-novel (SUN) and metastable-unique-novel (MSUN) rates. Post-training further improves generation performance on both MP-20 and Alex-MP-20, and our flow-based pretraining approach also achieves competitive performance in crystal structure prediction (CSP).

Our main contributions are:

- **A unified training framework.** Coordinated PT, SFT, and RL connecting crystal modeling with discovery-oriented distribution refinement.
- **Geometry-aware flow pretraining.** A strong Transformer generator that captures periodic geometry and supports variable atom counts, providing a shared foundation for post-training.
- **Controlled SFT data selection.** A data-selection strategy that adjusts the stable–metastable balance and provides a suitable starting distribution for RL.
- **Reinforcement learning with shared flow dynamics.** Reusing the pretraining noise design for exploration and optimizing the same velocity field. Applied to the fine-tuned checkpoint, this increases novelty while largely preserving metastability, moving beyond the observed pretraining trade-off.

2 Related work

Backbones: from equivariant graphs to Transformers. Equivariant message passing has been the dominant way to encode Euclidean symmetries while accounting for periodic boundaries in crystal generators, as in MatterGen (Zeni et al., 2025). Beyond backbone equivariance, DiffCSP++ (Jiao et al., 2024) and SymmCD (Levy et al., 2025) explicitly enforce crystallographic symmetry through Wyckoff constraints or asymmetric-unit representations. Recent work also explores Transformer backbones: CrystalDiT (Yi et al., 2026) and ADiT (Joshi et al., 2025) use diffusion Transformers (DiTs), while Crystalite (Hadži Veljković et al., 2026) combines continuous element tokens with periodic attention biases. FieldForge uses a Transformer without imposing full equivariance or explicit space-group constraints, instead incorporating periodicity and lattice geometry through flow paths and relative attention. This retains geometric inductive biases without symmetry-specific generation mechanisms, providing a shared backbone for pretraining and post-training.

Generative objectives: from VAEs to diffusion and flow matching. CDVAE (Xie et al., 2022) combines a composition VAE with coordinate Langevin dynamics, and DiffCSP (Jiao et al., 2023) replaces that decomposition with joint equivariant diffusion over fractional coordinates and lattice. FlowMM (Miller et al., 2024) formulates crystal generation as Riemannian flow matching; CrystalFlow (Luo et al., 2025) uses conditional flow matching, OMatG (Höllmer et al., 2025) uses stochastic interpolants (Albergo et al., 2023), and KLDM (Cornet et al., 2025) lifts torus diffusion into an augmented Euclidean space. Our formulation is a rectified flow (Liu et al., 2023; Lipman et al., 2023) in which element types, periodic coordinates, and lattice share one velocity objective, with lattice-aware image selection built into the coordinate path.

Supervised fine-tuning. LLM pipelines use supervised fine-tuning on curated data between pretraining and RL (Ouyang et al., 2022; Zhang et al., 2023; DeepSeek-AI, 2025). Representative crystal generators instead optimize a single generative objective on their training data (Xie et al., 2022; Jiao et al., 2023; Miller et al., 2024). In materials discovery, Alexandria (Schmidt et al., 2023) provides stability-curated training data, while GNoME (Merchant et al., 2023) couples stability filtering with iterative active-learning curation; neither is a generator fine-tuning method. Our SFT continues training with the original flow-matching loss on a selected subset of the original training split, rather than newly constructed instruction data. This adapts the pretrained distribution for RL without adding data or an auxiliary model.

Reward-based post-training. DDPO (Black et al., 2024) and DPOK (Fan et al., 2023) formulate iterative generation as a multi-step policy, while group-relative policy optimization (GRPO) (Shao et al., 2024) replaces a learned critic with group-relative advantages. Flow-GRPO (Liu et al., 2025) and DanceGRPO (Xue et al., 2025) extend online policy optimization to deterministic flow models by constructing stochastic training policies. DRaFT (Clark et al., 2024) and AlignProp (Prabhudesai et al., 2024) instead backpropagate through the sampling trajectory for differentiable rewards; this does not directly support our relaxation, convex-hull, novelty, and validity operations. In crystal generation, Chemeleon2 (Park & Walsh, 2026) is the closest precedent for de novo generation (DNG), applying GRPO with stability, diversity, novelty, and validity rewards to a pretrained latent diffusion model. OMatG-IRL (Höllmer & Martiniani, 2026) constructs stochastic surrogate transitions for flow policies, and CrystalGRPO (Su et al., 2026) reinforces coordinate and lattice flows with target- and coverage-aware CSP rewards. Unlike Chemeleon2, FieldForge inserts a controlled SFT data-selection stage before RL and post-trains the same direct-space flow that jointly generates composition, coordinates, and lattice; we also report every stage of the framework under one evaluation protocol, so the contribution of each stage can be read off directly.

3 METHOD

Figure 1 summarizes the training pipeline; Appendix A gives all dimensions and hyperparameters. We organize the method by pipeline stage, in the order a sample is actually produced. Section 3.1 covers *pretraining*, which rests on two independent choices: a Transformer parameterizes the velocity field (Section 3.1.1), and a flow-matching strategy prescribes the paths that field is regressed against (Section 3.1.2). Section 3.2 covers *post-training* and Section 3.3 *inference*.

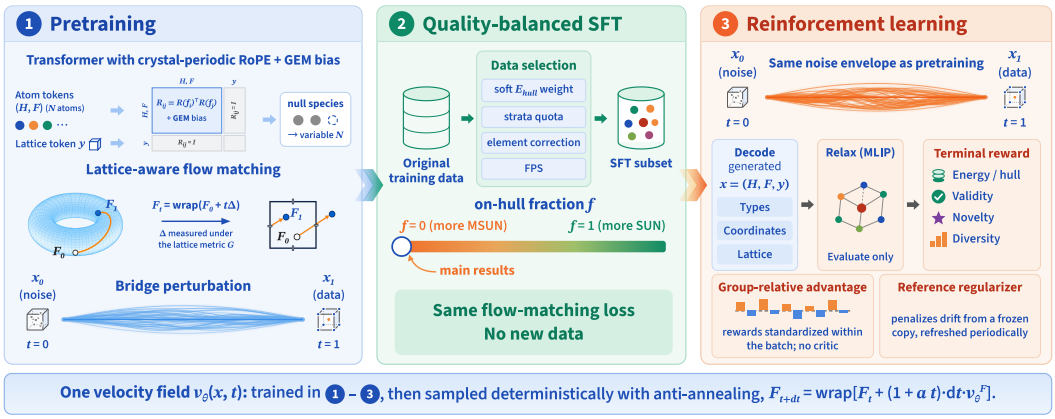


Figure 1: Overview of the FieldForge training framework. A single crystal velocity field is pretrained by flow matching, supervised on a selected subset of the original training corpus, and then optimized with group-relative rewards. Crystal geometry is preserved throughout the pipeline, and the inter-atomic potential is used to evaluate reinforcement-learning rollouts.

3.1 PRETRAINING: ARCHITECTURE AND FLOW MATCHING

3.1.1 ARCHITECTURE

The architecture parameterizes the velocity field: it maps a noisy state and a flow time to one velocity per channel, and is specified independently of the paths those velocities are regressed against. A crystal with N atoms is represented as $x = (H, F, y)$: type features $H \in \mathbb{R}^{N \times d_H}$, fractional coordinates $F \in [0, 1)^{N \times 3}$, and a six-dimensional lattice latent y encoding three cell lengths and three angles. Each atom is one token, and the cell contributes one further token. Elements are embedded as continuous descriptor vectors compressed using PCA (Hadži Veljković et al., 2026), and decoded by cosine nearest-neighbour snapping to element prototypes; this keeps the type channel a genuine regression target compatible with a continuous flow.

Variable atom count. A fixed set of input tokens cannot otherwise change the number of atoms in a generated cell. We pad crystals to the maximum cell size and represent the *null species* (Okhotin et al., 2025) with a learned unit-norm prototype in the element vocabulary. Padded tokens target this prototype and uniformly sampled coordinates; type loss covers every token, whereas coordinate loss covers only real atoms. Tokens decoded as null are removed, so atom count is generated jointly with composition rather than supplied to the sampler.

The pre-norm trunk uses DiT-style adaptive LayerNorm conditioning on flow time, zero-initialized modulation and residual gates, and a parameter-free LayerNorm before each residual gate (Appendix A). Zero initialization makes each new block initially approximate the identity, while the additional normalization prevents residual scale from drifting over the long training schedule.

Crystal-periodic geometric attention. We combine a crystal-periodic extension of rotary position embedding (RoPE) (Su et al., 2024) with Crystalite's Geometry Enhancement Module (GEM) (Hadži Veljković et al., 2026). For one attention head,

$$\mathrm{Attn}_i(Q,K,V) = \sum_j \mathrm{softmax}_j\left(\frac{q_i^\top R_{ij} k_j}{\sqrt{d_h}} + (B_{\mathrm{GEM}})_{ij}\right) v_j, \tag{1}$$

where q_i, k_j, v_j are query, key, and value vectors, and d_h is the head dimension. For atom–atom pairs, $R_{ij} = R(f_i)^\top R(f_j)$; otherwise, $R_{ij} = I$. RoPE uses dyadic integer multiples of 2π to encode relative positions independently of cell origin and periodic images, while GEM adds lattice-metric pairwise geometry.

3.1.2 Lattice-aware flow matching

Algorithm 1 constructs the joint flow-matching target: element features and lattice latents follow Euclidean paths, while fractional coordinates follow a lattice-metric periodic path. We use $t = 0$ for noise and $t = 1$ for data.

Algorithm 1 Joint flow-matching training objective.

Require: $x_1 = (H_1, F_1, y_1)$; $G = LL^\top$; $\mathcal{C} = \{H, F, y\}$; (μ_y, σ_y); $\{\eta_c, \lambda_c\}_{c\in\mathcal{C}}$; N_{steps}

$t \sim \mathrm{LogitNormal}(0,1)$; $H_0, \tilde{y}_0 \sim \mathcal{N}(0, I)$; $F_0 \sim \mathcal{U}(-\frac{1}{2}, \frac{1}{2})^3$

Euclidean paths *(y is whitened)*

$(H_t, v_H^*) \leftarrow \big((1-t)H_0 + tH_1,\ H_1 - H_0\big)$; $\tilde{y}_1 \leftarrow (y_1 - \mu_y) \oslash \sigma_y$

$(\tilde{y}_t, v_y^*) \leftarrow \big((1-t)\tilde{y}_0 + t\tilde{y}_1,\ \tilde{y}_1 - \tilde{y}_0\big)$

Lattice-aware periodic path

$\Delta_i \leftarrow \underset{\delta \in F_{1,i} - F_{0,i} + \mathbb{Z}^3}{\arg\min} \delta^\top G \delta, \quad \forall i$

$F_t \leftarrow \mathrm{wrap}(F_0 + t\Delta)$

$\mathrm{mask}_i \leftarrow \mathbf{1}[i \text{ is a real atom}]$; $\bar{\Delta} \leftarrow \frac{\sum_i \mathrm{mask}_i m_i \Delta_i}{\sum_i \mathrm{mask}_i m_i}$; $v_F^* \leftarrow \Delta - \bar{\Delta}$

Bridge perturbation and prediction

$\varepsilon_c \sim \mathcal{N}(0, I)$; $\tilde{x}_t^c \leftarrow x_t^c + 4\eta_c t(1-t) N_{\mathrm{steps}}^{-1/2} \varepsilon_c$, $c \in \mathcal{C}$; $\tilde{F}_t \leftarrow \mathrm{wrap}(\tilde{F}_t)$

$\hat{v} \leftarrow v_\theta\big(\tilde{H}_t, \mathrm{mod}_1(\tilde{F}_t + \frac{1}{2}), \sigma_y \odot \tilde{y}_t + \mu_y, t\big)$

Shared velocity-matching loss

$\mathrm{mask}_F = \mathrm{mask}$, $\mathrm{mask}_H = \mathrm{mask}_y = \mathbf{1}$; $\mathcal{L}_c \leftarrow \|\mathrm{mask}_c \odot (\hat{v}_c - v_c^*)\|_2^2$, $c \in \mathcal{C}$; $\mathcal{L}_{\mathrm{FM}} \leftarrow \sum_{c\in\mathcal{C}} \lambda_c \mathcal{L}_c$

return $\mathcal{L}_{\mathrm{FM}}$

The periodic path follows the full minimum-image displacement Δ, whereas the velocity target subtracts the mass-weighted mean over real atoms, $v_F^* = \Delta - \bar{\Delta}$, where m_i denotes atomic mass. This removes an arbitrary translation of the entire crystal. The bridge perturbation leaves both endpoints unchanged and provides a shared noise design connecting pretraining to RL exploration.

3.2 Post-training: supervised fine-tuning, then reinforcement learning

Flow matching learns the training distribution but does not directly optimize discovery objectives. We therefore combine supervised fine-tuning to establish a suitable starting distribution with reinforcement learning to improve thermodynamic quality, novelty, and diversity. Both stages refine the

same velocity field without changing the architecture. An interatomic potential provides stability estimates for the terminal reward, while policy updates require no learned value function.

3.2.1 SUPERVISED FINE-TUNING

QB-SFT acts as a controlled distribution intervention that adapts the pretrained generative prior and provides a suitable initialization for subsequent RL optimization (Appendix C). During SFT, the model and flow-matching loss remain unchanged, while the training data are replaced by a subset selected from the original training split. Selecting the top-k structures with the lowest energy above hull (E_{hull}) can reduce chemical and structural coverage, whereas random selection provides no explicit preference for thermodynamic quality. Our procedure moves probability mass toward low-energy structures while retaining chemically and geometrically informative variation.

Concretely, QB-SFT first constructs low-energy candidate pools, with separate on-hull and near-hull allocations to control the stable–metastable balance. Within each pool, it allocates tempered quotas over element-count, atom-count, and crystal-system strata. Candidates receive soft energy weights with a clipped full-set/core element-incidence correction. Weighted farthest-point sampling then preserves structural coverage using lattice features and the radial distribution function (RDF). We vary the on-hull fraction while holding the other selection settings fixed to isolate its effect. Appendix C details the selection procedure, including thresholds, descriptors, and parameter ranges; Appendix D reports the on-hull scan and targeted comparisons of random SFT, core-energy threshold, and subset size.

For the main results, SFT uses subsets of 5,000 structures for MP-20 and 115,000 for Alex-MP-20, with a core threshold $T = 0.025\,\text{eV/atom}$. SFT runs for 50,000 optimization steps on both datasets. For each run, we report results from the best checkpoint.

3.2.2 REINFORCEMENT LEARNING

Terminal objective. Starting from the SFT model, we further optimize the flow using physical feedback evaluated on the final generated crystal. The objective is

$$J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}\left[R(C(x_{t_T}))\right], \tag{2}$$

where t_T is the terminal rollout time, C decodes the terminal state into a crystal, and R is evaluated after relaxation. This enables optimization of non-differentiable objectives such as E_{hull}, chemical validity, and structural novelty.

Stochastic flow policy. To connect pretraining with RL exploration, we reuse the noise design of the pretraining bridge perturbation to construct stochastic flow transitions. The learned velocity field $v_\theta(x_t, t)$ determines the transition mean, while the added noise enables exploration during RL. For channel c,

$$x^c_{t_{k+1}} = \underbrace{x^c_{t_k} + \Delta t_k\, v^c_\theta(x_{t_k}, t_k)}_{\mu^c_{\theta,k}} + \sigma^c_k\, \epsilon^c_k, \qquad \sigma^c_k = \eta_c\, 4t_k(1-t_k)\sqrt{|\Delta t_k|}, \tag{3}$$

where $\epsilon^c_k \sim \mathcal{N}(0, I)$. Nonzero exploration variance gives each transition a tractable probability density. Fractional-coordinate transitions are wrapped to the fundamental domain, and their likelihood is evaluated using the nearest periodic image. At inference, we use the deterministic flow sampler of Section 3.3.

Critic-free terminal policy optimization. For a rollout batch of B trajectories, we standardize the terminal returns within the batch,

$$A_i = \frac{R_i - \overline{R}}{s_R + \epsilon_A}, \tag{4}$$

following the group-relative normalization of Shao et al. (2024), and do not train a separate value function. The same trajectory-level advantage A_i is assigned to the selected transitions of trajectory i. The resulting critic-free policy objective is

$$\mathcal{L}_{\text{PG}} = -\frac{1}{B|\mathcal{S}|}\sum_{i=1}^{B}\sum_{k\in\mathcal{S}} A_i\, \log \pi_\theta\left(x_{i,t_{k+1}} \mid x_{i,t_k}, t_k\right), \tag{5}$$

where $\mathcal{S}$ is a stratified subset of flow transitions with nonzero exploration variance used for gradient computation.

The full trajectory is always retained for terminal decoding and reward evaluation; transition subsampling only reduces the memory cost of backpropagation.

To constrain policy drift during terminal reward optimization, we regularize each update against a periodically refreshed frozen reference policy,

$$\mathcal{L}_{\mathrm{RL}} = \mathcal{L}_{\mathrm{PG}} + \beta\, \mathcal{L}_{\mathrm{ref}}. \tag{6}$$

The reference term is

$$\mathcal{L}_{\mathrm{ref}} = \frac{1}{B|\mathcal{S}|} \sum_{i=1}^{B} \sum_{k \in \mathcal{S}} K_{i,k}, \qquad K_{i,k} = \frac{1}{D_i} \sum_{c} \frac{\left\| \mu^{c}_{\theta,k} - \mu^{c}_{\mathrm{ref},k} \right\|^2_{\mathcal{A}_c}}{2(\sigma^{c}_{k})^2}, \tag{7}$$

where D_i is the number of active scalar degrees of freedom for trajectory i and $\mathcal{A}_c$ denotes the active entries of channel c. The reference constrains policy drift while allowing reward-guided refinement.

Terminal reward. The reward balances thermodynamic quality with chemical validity and distributional coverage. Thermodynamic quality is represented by a dense energy-above-hull term together with bonuses for near-hull structures, $E_{\mathrm{hull}} \leq 0.1\,\mathrm{eV/atom}$, and stable structures, $E_{\mathrm{hull}} \leq 0$. We further include SMACT-based chemical validity (Davies et al., 2019) and composition- and structure-level novelty and diversity (Park & Walsh, 2026). Active reward components are standardized over finite-energy structures before aggregation.

The interatomic potential is used only to relax and score decoded terminal structures and is not differentiated through. The transition-likelihood construction, reference regularization, reward definition, relaxation protocol, and training hyperparameters are given in Appendix E.

3.3 Inference: deterministic anti-annealed sampling

Generation uses deterministic integration. Following Miller et al. (2024), we apply inference anti-annealing to amplify the fractional-coordinate velocity later in the trajectory,

$$F_{t+\mathrm{d}t} = \mathrm{wrap}\left[F_t + \left(1 + a\,t\right) \mathrm{d}t\; v^{F}_{\theta}(x_t, t)\right], \tag{8}$$

where a is the anti-annealing coefficient tuned for each task.

4 Experiments

4.1 Setup

De novo generation. In DNG, the model generates both composition and geometry. We evaluate on MP-20 (Jain et al., 2013; Xie et al., 2022) and Alex-MP-20, derived from Alexandria (Schmidt et al., 2023). Each FieldForge model is sampled with ten seeds, producing 250 crystals per seed and 2,500 in total. We relax these structures with NequIP (Batzner et al., 2022; Tan et al., 2025) and evaluate them using LeMat-GenBench (Betala et al., 2025). Evaluation covers validity, uniqueness, novelty, and thermodynamic stability. Based on E_{hull}, stable crystals satisfy $E_{\mathrm{hull}} \leq 0$, while metastable crystals satisfy $0 < E_{\mathrm{hull}} \leq 0.1\,\mathrm{eV/atom}$. Combining each stability class with uniqueness and novelty gives SUN and MSUN, respectively.

Crystal structure prediction. In CSP, the composition is fixed and the model predicts its geometry. We adopt the same pretraining approach with composition conditioning (Appendix A). We evaluate the complete MP-20 and MPTS-52 (Jiao et al., 2023) test splits and report means over four seeds. MPTS-52 is harder by construction: its training cells contain fewer atoms on average than its chronologically later test cells, so its absolute match rates should not be compared with those on MP-20. We report StructureMatcher (Ong et al., 2013) match rate (MR) and normalized RMSE over matched structures: MR measures how often a candidate matches, while RMSE measures geometric error among the matches.

4.2 DE NOVO GENERATION

On MP-20 (Table 1), pretrained FieldForge outperforms the evaluated baselines in SUN, MSUN, validity, and uniqueness. SFT further improves novel coverage and MSUN on MP-20, whereas on Alex-MP-20 (Table 2), it primarily improves thermodynamic quality. Subsequent RL increases novelty, SUN, and MSUN on both datasets while maintaining high validity and uniqueness. PT-SFT-RL also achieves a higher combined SUN and MSUN rate than PT-RL on both datasets, highlighting the benefit of SFT initialization and supporting the effectiveness of the complete training framework. FieldForge results in Tables 1 and 2 use coordinate anti-annealing with $a = 4$ (Section 3.3).

Table 1: LeMat-GenBench (Betala et al., 2025) MP-20 results. Published rows are leaderboard entries. * denotes non-pre-relaxed models. Our rows show PT and its post-training variants.

Model	Valid (%)↑	Unique (%)↑	Novel (%)↑	Stable (%)↑	Metastable (%)↑	SUN (%)↑	MSUN (%)↑	E_{hull} (eV/at.)↓	E_{form} (eV/at.)	RMSD (Å)↓
ADiT* (Joshi et al., 2025)	90.60	87.80	26.00	0.40	36.50	0.00	1.00	0.3333	−0.7339	0.3794
SymmCD* (Levy et al., 2025)	73.40	73.00	47.00	1.40	18.60	0.10	2.40	0.8761	−0.0195	0.8720
DiffCSP++* (Jiao et al., 2024)	95.30	95.10	62.00	1.00	26.40	0.20	5.00	0.4093	−0.5189	0.6933
MatterGen (Zeni et al., 2025)	95.70	95.10	70.50	2.00	33.40	0.20	15.00	0.1834	−0.7020	0.3878
MiAD (Okhotin et al., 2025)	96.20	94.30	40.20	6.40	**62.00**	1.00	16.60	**0.0804**	−0.8436	0.2494
MCFlow (Seong et al., 2026)	97.20	96.30	52.20	11.90	49.30	0.70	18.90	0.0987	−0.8513	0.1696
Chemeleon2* (Park & Walsh, 2026)	95.20	88.10	71.60	0.00	39.80	0.00	21.20	0.1557	−0.7167	0.4226
Crystalite (Hadži Veljković et al., 2026)	97.20	95.80	53.20	**12.70**	51.60	1.50	22.60	0.0905	−0.8916	0.1322
FieldForge (PT)	**97.68**	96.56	59.12	11.28	57.08	3.12	29.88	0.0809	−0.9203	**0.0732**
FieldForge (PT-RL)	97.20	96.60	66.88	11.04	54.44	**3.72**	33.88	0.0912	−0.9278	0.0994
FieldForge (PT-SFT)	97.40	**97.20**	63.56	9.00	57.00	1.88	33.96	0.0873	−0.9074	0.0911
FieldForge (PT-SFT-RL)	**97.68**	96.72	**75.48**	7.08	57.32	2.72	**41.12**	0.0872	−0.8972	0.1087

Table 2: LeMat-GenBench Alex-MP-20 track. Published rows are leaderboard entries. * denotes non-pre-relaxed models. Our rows show PT and its post-training variants.

Model	Valid (%)↑	Unique (%)↑	Novel (%)↑	Stable (%)↑	Metastable (%)↑	SUN (%)↑	MSUN (%)↑	E_{hull} (eV/at.)↓	E_{form} (eV/at.)	RMSD (Å)↓
CrystaLLM-π* (Bone et al., 2025)	93.40	93.20	23.90	4.80	**71.30**	0.30	12.50	0.0941	−0.5811	0.1914
Chemeleon* (Park et al., 2025)	94.00	93.90	59.80	4.00	54.40	1.00	28.00	0.1120	−0.6344	0.3089
Chemeleon2* (Park & Walsh, 2026)	97.40	93.50	**93.70**	0.30	67.40	0.20	**62.40**	0.0959	−0.6086	0.2678
FieldForge (PT)	95.52	95.52	67.80	6.48	63.40	2.92	42.04	0.0755	−0.6642	0.0689
FieldForge (PT-RL)	**98.12**	**96.76**	74.64	4.00	71.28	1.32	52.32	0.0688	−0.4332	0.0950
FieldForge (PT-SFT)	95.72	95.72	62.84	**11.12**	67.24	4.92	42.36	**0.0562**	−0.8447	**0.0538**
FieldForge (PT-SFT-RL)	95.76	95.48	77.40	9.48	66.44	**5.44**	53.20	0.0662	−0.5049	0.1010

4.3 CRYSTAL STRUCTURE PREDICTION

FieldForge achieves competitive CSP performance across both datasets and sampling budgets (Table 3). At CSP@1, it attains the highest match rates among the compared models on MP-20 and MPTS-52, together with the lowest RMSE on MP-20. At CSP@20, match rates remain competitive, although FieldForge does not consistently lead in either match rate or geometric accuracy.

5 DISCUSSION

We distinguish the trade-off between stable/metastable generation and unique-and-novel (UN) coverage from the balance between stable and metastable outputs, and examine how training and inference control these preferences.

5.1 PRETRAINING TRADE-OFFS AND INFERENCE GAINS

Our pretraining experiments reveal a trade-off between thermodynamic quality and novelty. Figure 2 (top) shows that, with architecture and sampler fixed, continued training generally increases metastability while reducing novelty, leaving MSUN relatively unchanged. Improved thermodynamic quality therefore does not necessarily translate into better joint generation performance.

Table 3: CSP results on the full MP-20 and MPTS-52 test splits with one and twenty samples per target. Baseline results are taken from the original papers, except for Crystalite's CSP@20 results, which we evaluate ourselves. FieldForge results are averaged over four sampling seeds. A dash denotes an unreported result.

	MP-20				MPTS-52			
	CSP@1		CSP@20		CSP@1		CSP@20	
Model	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓	MR (%) ↑	RMSE ↓
CDVAE (Xie et al., 2022)	33.90	0.1045	—	—	5.34	0.2106	—	—
DiffCSP (Jiao et al., 2023)	51.49	0.0631	77.93	0.0492	12.19	0.1786	34.02	0.1749
FlowMM (Miller et al., 2024)	61.39	0.0566	—	—	17.54	0.1726	—	—
CrystalFlow (Luo et al., 2025)	62.02	0.0710	**78.34**	0.0577	21.00	0.1613	37.81	0.1584
MCFlow (Seong et al., 2026)	64.08	0.0561	76.08	0.0383	27.16	0.1401	**41.45**	0.1296
Crystalite (Hadži Veljković et al., 2026)	66.05	0.0329	71.94	**0.0319**	31.49	**0.0701**	37.46	**0.0794**
FieldForge	**66.31**	**0.0322**	74.68	0.0434	**31.55**	0.0991	39.61	0.0952

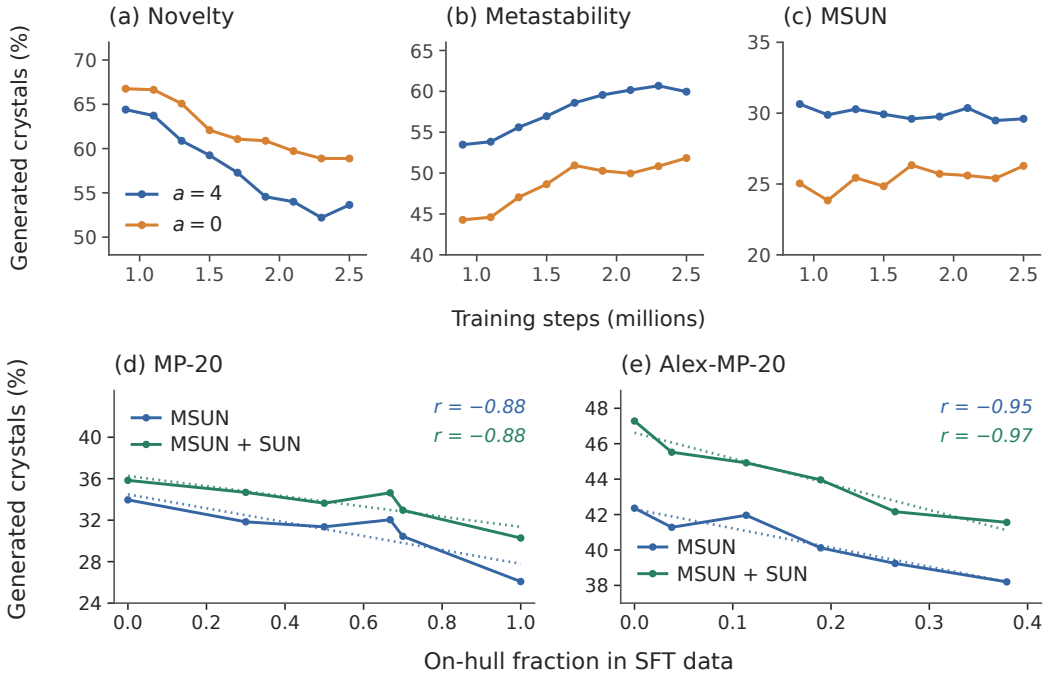


Figure 2: Effects of pretraining, inference, and SFT data selection. Top: from 0.9M to 2.5M training steps on LeMat-GenBench, metastability rises while novelty falls; anti-annealing ($a = 4$) improves MSUN over $a = 0$ throughout. Bottom: increasing the fraction of on-hull structures in the SFT subset lowers MSUN and MSUN+SUN in the 5,000-structure MP-20 (left) and 115,000-structure Alex-MP-20 ($T = 0.025$, right) scans. Dotted lines and r denote linear fits and Pearson correlations across the displayed points.

Anti-annealing changes this balance at inference. At matched checkpoints, $a = 4$ achieves higher metastability but lower novelty than $a = 0$. Despite this reduction in novelty, MSUN remains higher throughout the sweep, showing that anti-annealing improves the yield of unique and novel metastable crystals without retraining. Further pretraining ablations are reported in Appendix B.

5.2 SFT CONTROLS STABLE VERSUS METASTABLE GENERATION

We vary the *on-hull fraction*—the proportion of SFT structures with approximately zero energy above hull—while keeping the model and training objective fixed. On MP-20, increasing this fraction generally raises SUN but lowers MSUN; the SUN gains do not offset the MSUN losses. In the 115,000-

structure Alex-MP-20 scan at $T = 0.025$ shown in Figure 2 (bottom), SUN declines alongside MSUN and MSUN+SUN (Appendix D). The response to on-hull allocation therefore depends on the dataset and selection settings, but in both scans it changes the balance of the two joint rates.

In our pretraining runs, stable and metastable generation rates tend to vary together, with higher rates accompanied by lower UN coverage. SFT, in contrast, provides a separate control over the relative preference for stable versus metastable outputs, reflected in the balance between SUN and MSUN. By making this preference adjustable through data selection alone, SFT provides a controllable starting distribution for subsequent RL.

5.3 RL ADVANCES THE METASTABILITY–NOVELTY FRONTIER

RL induces an outward shift in the empirical metastability–novelty frontier. Starting from the fine-tuned checkpoint, RL increases novelty and MSUN on both MP-20 and Alex-MP-20 while leaving the metastable fraction approximately unchanged (Tables 1 and 2). In contrast to the checkpoint trend in Section 5.1, these gains in novelty are not accompanied by a comparable loss of metastability.

The effect of initialization differs across datasets. Applying RL directly to the pretrained model increases novelty at the expense of metastability on MP-20, but improves both on Alex-MP-20. Nevertheless, PT-SFT-RL achieves higher novelty and MSUN than PT-RL on both datasets, supporting the value of SFT initialization. SFT controls the stable–metastable balance, while subsequent RL improves novelty and MSUN beyond the observed pretraining trade-off.

6 CONCLUSION

FieldForge unifies flow-based pretraining, supervised fine-tuning, and reinforcement learning within a single Transformer generator that incorporates periodic crystal geometry. The pretrained model achieves state-of-the-art SUN and MSUN rates on the MP-20 track of LeMat-GenBench, and post-training further improves generation performance on both MP-20 and Alex-MP-20. Our flow-based pretraining approach also achieves competitive performance in composition-conditioned crystal structure prediction.

The stages provide distinct forms of control. During pretraining, increased metastability comes at the cost of unique-and-novel coverage, limiting gains in joint yield. Through controlled data selection and without adding new data, SFT adjusts the balance between stable and metastable generation and provides an effective starting distribution for RL. RL from the fine-tuned checkpoint then advances this empirical metastability–novelty frontier, increasing novelty while largely preserving metastability. Anti-annealing further improves MSUN without retraining. These results demonstrate the value of coordinating model design, supervised data selection, reward optimization, and inference within a unified framework for crystal generation.

Limitations. Following the LeMat-GenBench protocol, we assess stability using interatomic-potential relaxation and hull-based evaluation. The observed SFT trends may vary with dataset composition and subset size. Further validation with DFT and evaluation across broader chemical spaces and larger unit cells remain directions for future work.

REPRODUCIBILITY STATEMENT

Appendices A–F provide the implementation details and ablations for pretraining, SFT, and RL. Appendix G defines the evaluation protocol, and Appendix H reports additional analyses and qualitative results.

THE USE OF LARGE LANGUAGE MODELS

In this work we used large language models (LLMs) for two purposes. First, to aid and polish writing: the authors wrote the technical content, and LLM tools were used to improve grammar, wording, and paragraph flow, to tighten and reorganize author-written text, and to resolve LaTeX formatting issues. Second, for retrieval and discovery: LLM tools helped locate and summarize related work

during the literature search. Every reference was subsequently verified by the authors against the original publication. All research ideas, methods, experiments, data, analyses, and conclusions are the authors' own; LLMs were not used for research ideation, experimental design or execution, data generation, or mathematical claims, and the authors take full responsibility for the content of the paper.

REFERENCES

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions, 2023. arXiv:2303.08797.

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, et al. A foundation model for atomistic materials chemistry. *The Journal of Chemical Physics*, 163(18): 184110, 2025. doi: 10.1063/5.0297006. URL https://doi.org/10.1063/5.0297006.

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*, 13:2453, 2022. doi: 10.1038/s41467-022-29939-5.

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, 2025. arXiv:2512.04562. Code: https://github.com/LeMaterial/lemat-genbench.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. In *International Conference on Learning Representations (ICLR)*, 2024. arXiv:2305.13301.

Cyprien Bone, Matthew Walker, Bradley A. A. Martin, 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, 2025. arXiv:2511.21299.

Kevin Clark, Paul Vicol, Kevin Swersky, and David J. Fleet. Directly fine-tuning diffusion models on differentiable rewards. In *International Conference on Learning Representations (ICLR)*, 2024. arXiv:2309.17400.

François R. J. Cornet, Federico Bergamin, Arghya Bhowmik, Juan Maria Garcia-Lastra, Jes Frellsen, and Mikkel N. Schmidt. Kinetic Langevin diffusion for crystalline materials generation. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 11215–11244, 2025. URL https://proceedings.mlr.press/v267/cornet25a.html.

Daniel W. Davies, Keith T. Butler, Adam J. Jackson, Jonathan M. Skelton, Kazuki Morita, and Aron Walsh. SMACT: Semiconducting materials by analogy and chemical theory. *Journal of Open Source Software*, 4(38):1361, 2019. doi: 10.21105/joss.01361.

DeepSeek-AI. DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning, 2025. arXiv:2501.12948.

Ying Fan, Olivia Watkins, Yuqing Du, Hao Liu, Moonkyung Ryu, Craig Boutilier, Pieter Abbeel, Mohammad Ghavamzadeh, Kangwook Lee, and Kimin Lee. DPOK: Reinforcement learning for fine-tuning text-to-image diffusion models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. arXiv:2305.16381.

Tin Hadži Veljković, Joshua Rosenthal, Ivor Lončarić, and Jan-Willem van de Meent. Crystalite: A lightweight transformer for efficient crystal modeling, 2026. arXiv:2604.02270.

Philipp Höllmer and Stefano Martiniani. Open materials generation with inference-time reinforcement learning, 2026. arXiv:2602.00424.

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 *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 23417–23450, 2025. URL https://proceedings.mlr.press/v267/hollmer25a.html.

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, 2013. doi: 10.1063/1.4812323.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal structure prediction by joint equivariant diffusion. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. arXiv:2309.04475.

Rui Jiao, Wenbing Huang, Yu Liu, Deli Zhao, and Yang Liu. Space group constrained crystal generation. In *International Conference on Learning Representations (ICLR)*, 2024. URL https://openreview.net/forum?id=jkvZ7v4OmP.

Chaitanya K. Joshi, Xiang Fu, Yi-Lun Liao, Vahe Gharakhanyan, Benjamin Kurt Miller, Anuroop Sriram, and Zachary W. Ulissi. All-atom diffusion transformers: Unified generative modelling of molecules and materials. In *International Conference on Machine Learning (ICML)*, 2025. arXiv:2503.03965.

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. In *International Conference on Learning Representations (ICLR)*, 2025. URL https://openreview.net/forum?id=xnssGv9rpW.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations (ICLR)*, 2023. arXiv:2210.02747.

Jie Liu, Gongye Liu, Jiajun Liang, Yangguang Li, Jiaheng Liu, Xintao Wang, Pengfei Wan, Di Zhang, and Wanli Ouyang. Flow-GRPO: Training flow matching models via online RL. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025. arXiv:2505.05470.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations (ICLR)*, 2023. arXiv:2209.03003.

Xiaoshan Luo, Zhenyu Wang, Qingchang Wang, Xuechen Shao, Jian Lv, Lei Wang, Yanchao Wang, and Yanming Ma. CrystalFlow: A flow-based generative model for crystalline materials. *Nature Communications*, 2025. doi: 10.1038/s41467-025-64364-4.

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, 2023. doi: 10.1038/s41586-023-06735-9.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: Generating materials with Riemannian flow matching. In *International Conference on Machine Learning (ICML)*, 2024. arXiv:2406.04713.

Andrey Okhotin, Maksim Nakhodnov, Nikita Kazeev, Mikhail Lazarev, Andrey E. Ustyuzhanin, and Dmitry Vetrov. MiAD: Mirage atom diffusion for de novo crystal generation, 2025. arXiv:2511.14426.

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, 2013. doi: 10.1016/j.commatsci.2012.10.028.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Hyunsoo Park and Aron Walsh. Guiding generative models to uncover diverse and novel crystals via reinforcement learning. *Nature Machine Intelligence*, 2026. doi: 10.1038/s42256-026-01262-4. arXiv:2511.07158.

Hyunsoo Park, Anthony Onwuli, and Aron Walsh. Exploration of crystal chemical space using text-guided generative artificial intelligence. *Nature Communications*, 16:4379, 2025. doi: 10.1038/s41467-025-59636-y.

Mihir Prabhudesai, Anirudh Goyal, Deepak Pathak, and Katerina Fragkiadaki. Aligning text-to-image diffusion models with reward backpropagation, 2024. arXiv:2310.03739.

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

Jonathan Schmidt, Noah Hoffmann, Hai-Chen Wang, Pedro Borlido, Pedro J. M. A. Carriço, Tiago F. T. Cerqueira, Silvana Botti, and Miguel A. L. Marques. Machine-learning-assisted determination of the global zero-temperature phase diagram of materials. *Advanced Materials*, 35(22):2210788, 2023. doi: 10.1002/adma.202210788.

Kiyoung Seong, Sungsoo Ahn, Sehui Han, and Changyoung Park. Multimodal crystal flow: Any-to-any modality generation for unified crystal modeling, 2026. arXiv:2602.20210.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. DeepSeekMath: Pushing the limits of mathematical reasoning in open language models, 2024. arXiv:2402.03300.

Jianlin Su, Yu Lu, Shengfeng Pan, Ahmed Murtadha, Bo Wen, and Yunfeng Liu. RoFormer: Enhanced transformer with rotary position embedding. *Neurocomputing*, 568:127063, 2024. arXiv:2104.09864.

Kaixiang Su, Hongfei Xue, and Qiang Zhu. CrystalGRPO: Target-aligned and coverage-preserving reinforcement learning for flow-based crystal structure prediction, 2026. arXiv:2608.06582.

Chuin Wei Tan, Marc L. Descoteaux, Mit Kotak, Gabriel de Miranda Nascimento, Seán R. Kavanagh, Laura Zichi, Menghang Wang, Aadit Saluja, Yizhong R. Hu, Tess Smidt, Anders Johansson, William C. Witt, Boris Kozinsky, and Albert Musaelian. High-performance training and inference for deep equivariant interatomic potentials, 2025. arXiv:2504.16068.

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. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/bbf23e81b0ad7637fe9a731d0b676ca6-Abstract-Conference.html.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. In *International Conference on Learning Representations (ICLR)*, 2022. arXiv:2110.06197.

Zeyue Xue, Jie Wu, Yu Gao, Fangyuan Kong, Lingting Zhu, Mengzhao Chen, Zhiheng Liu, Wei Liu, Qiushan Guo, Weilin Huang, and Ping Luo. DanceGRPO: Unleashing GRPO on visual generation, 2025. arXiv:2505.07818.

Xiaohan Yi, Guikun Xu, Zhong Zhang, Liu Liu, Yatao Bian, Xi Xiao, and Peilin Zhao. CrystalDiT: Simple diffusion transformers for crystal generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 1462–1470, 2026. doi: 10.1609/aaai.v40i2.37121.

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*, 639:624–632, 2025. doi: 10.1038/s41586-025-08628-5.

Shengyu Zhang, Linfeng Dong, Xiaoya Li, Sen Zhang, Xiaofei Sun, Shuhe Wang, Jiwei Li, Runyi Hu, Tianwei Zhang, Fei Wu, and Guoyin Wang. Instruction tuning for large language models: A survey, 2023. arXiv:2308.10792.

A PRETRAINING AND SAMPLING DETAILS

Table 4 summarizes the DNG pretraining configuration. Each pretraining ablation changes one component. Table 5 reports the composition-conditioned CSP configuration, which uses fixed atom types without null species, type loss, or type-channel perturbation.

For the pretraining bridge perturbation in Algorithm 1, we fix $N_{\text{steps}} = 150$, matching the sampling step count. During training, this is a reference scale for the perturbation amplitude, not a number of integration steps.

Table 4: PT configuration.

Architecture		Training and sampling	
Trainable parameters	≈67M	Global batch	128 (2×64)
Transformer width	512	Learning rate	2×10^{-4}, 2000 warmup
Layers / heads	14 / 16	Optimizer	AdamW $(0.9, 0.95)$, wd 10^{-2}
Block	AdaLN-zero + sandwich norm	Training steps	2.5×10^6
Atom tokenization	continuous, PCA d_H=16	Precision / EMA	`bf16` / 0.9999
Coordinate encoding	rotary (3D, dyadic)	Gradient guard	skip non-finite steps
Velocity heads	scaled $\tanh$; coord. fractional	$(\lambda_H, \lambda_F, \lambda_y)$	$(2, 50, 20)$
GEM	shared across layers	Coordinate loss	wrapped-frac. MSE
Distance / edge bias	enabled / enabled	Time distribution	logit-normal $(0, 1)$
Edge-bias dims	256 hid., 12 freq., 32 RBF	(η_H, η_F, η_y)	$(0.05, 0.20, 0.40)$
Noise gate	enabled	Whitening / frac. path	lattice / torus
PBC search radius	1	Sampling	150 Euler steps
Max atoms per cell	20	Sampling weights	EMA
Variable atom count	null species	Training perturbation N_{steps}	150

B PRETRAINING AND SAMPLING ABLATIONS

We first isolate the contributions of null species, rotary attention, and the lattice metric under the base sampler, then examine how anti-annealing modifies these effects.

B.1 EVALUATION SETUP

FieldForge (PT) and three single-component ablations are each trained for 2.5M steps (Table 4). *No null species* samples atom counts from an external distribution instead of generating them through null tokens. *No rotary* replaces rotary attention with absolute Fourier coordinate embeddings. *Identity metric* replaces the lattice Gram matrix G by the identity in minimum-image selection while retaining the periodic path (Algorithm 1).

Internal evaluation. We compare 18 matched checkpoints (0.8M–2.5M steps) using the same sampling seed, with $a = 0$ and $a = 4$, on up to 2,500 relaxed structures per checkpoint. Table 6 reports mean paired differences from PT, controlling for training progress. For MSUN, $\pm$ is the checkpoint standard deviation divided by $\sqrt{18}$; it summarizes variation along a correlated training trajectory, not independent-run uncertainty.

External evaluation. Table 7 reports ten-seed means on LeMat-GenBench: matched 2.0M-step checkpoints with $a = 0$, and each model's best checkpoint selected using internal MSUN and SUN

Table 5: Structure prediction configuration. Shared settings span both dataset columns. CSP@1 and CSP@20 use their respective anti-annealing optima.

Parameter	MP-20	MPTS-52
Architecture		
Trainable parameters	≈269M	
Transformer width	1024	
Layers / heads	14 / 16	
Block	AdaLN-zero + sandwich norm	
Atom tokenization	one-hot, given by target	
Coordinate encoding	rotary (3D, dyadic)	
Velocity heads	scaled $\tanh$; coord. fractional	
GEM	per layer	shared across layers
Distance / edge bias	enabled / enabled	
Edge-bias dims	256 hid., 12 freq., 32 RBF	
Noise gate	enabled	
PBC search radius	1	
Max atoms per cell	20	52
Variable atom count	none (composition given)	
Training and sampling		
Global batch	256 (4×64)	256 (2×128)
Learning rate	2×10^{-4}, 5000 warmup, cosine	
Optimizer	AdamW $(0.9, 0.95)$	
Weight decay	10^{-1}	10^{-2}
Training steps	2.5×10^6	
Evaluated checkpoint	final	best val. match rate (1.96×10^6)
Precision / EMA	`bf16` / 0.9999	
Gradient guard	skip non-finite steps	
$(\lambda_H, \lambda_F, \lambda_y)$	$(0, 50, 20)$	
Coordinate loss	wrapped-frac. MSE	
Time distribution	logit-normal $(0, 1)$	
(η_H, η_F, η_y)	$(0, 0.20, 0.40)$	$(0, 0.10, 0.20)$
Whitening / frac. path	lattice / torus	
Sampling	150 Heun steps, clean readout at $t = 0.98$	
Anti-annealing a, CSP@1	10	14
Anti-annealing a, CSP@20	0	12
Sampling weights	EMA	
Training perturbation N_{steps}	150	

results with $a = 4$. Cross-block differences therefore include both checkpoint and sampler effects. Internal and external evaluation procedures differ, so their metric values need not coincide.

Table 6: Matched-checkpoint ablations under internal DNG evaluation. Values are mean percentage-point differences from FieldForge (PT); parentheses count checkpoints with lower MSUN than PT.

	$a = 0$	$a = 4$				
Ablation	ΔMSUN	ΔMSUN	Δmetastable	ΔU.N.	Δstable	Δcomp. val.
No null species	-2.82 ± 0.44 (17/18)	-4.26 ± 0.50 (18/18)	$+6.89$	-10.78	$+1.89$	$+2.48$
No rotary	-2.40 ± 0.39 (17/18)	-0.78 ± 0.57 (11/18)	-2.99	$+2.34$	-0.85	-0.02
Identity metric	-1.07 ± 0.21 (15/18)	0.00 ± 0.34 (11/18)	$+1.01$	-0.63	$+0.15$	$+0.26$

B.2 Component contributions under the base sampler

With $a = 0$, every ablation lowers MSUN in both evaluations. The largest loss comes from removing null species, followed by rotary attention and the lattice metric (Tables 6 and 7).

No null species reduces novel coverage. At the matched external checkpoint, stable and metastable rates increase, but uniqueness and novelty fall, reducing MSUN. Null species therefore contribute

Table 7: Pretraining ablations on LeMat-GenBench. All generated structures are pre-relaxed before evaluation.

Variant	Valid (%)↑	Unique (%)↑	Novel (%)↑	Stable (%)↑	Metastable (%)↑	SUN (%)↑	MSUN (%)↑	E_{hull} (eV/at.)↓	E_{form} (eV/at.)	RMSD (Å)↓
Anti-annealed sampling (a=4), best checkpoints										
FieldForge (PT)	97.68	96.56	59.12	11.28	57.08	3.12	29.88	0.0809	−0.9203	0.0732
No null species	98.28	97.84	47.88	14.56	61.44	2.92	25.40	0.0625	−0.8990	0.0718
No rotary	97.28	97.16	52.80	12.36	60.32	2.96	28.40	0.0712	−0.9187	0.0704
Identity metric	97.48	97.40	61.32	9.56	54.20	2.96	28.48	0.0901	−0.9173	0.0777
Base sampler (a=0), matched at 2.0M steps										
FieldForge (PT)	97.28	97.04	59.68	10.08	51.80	2.04	26.44	0.0928	−0.8518	0.1115
No null species	96.64	96.32	53.44	12.28	52.72	2.16	21.92	0.0834	−0.8602	0.1249
No rotary	96.68	95.96	60.12	8.72	49.56	1.44	24.52	0.1020	−0.8378	0.1230
Identity metric	97.24	97.16	57.92	9.68	52.16	1.76	24.76	0.0915	−0.8586	0.1159

primarily through coverage rather than thermodynamic quality alone. Since the type-loss weight was tuned jointly for null decisions and element classification, this comparison evaluates the configured mechanism, not the token in isolation.

No rotary lowers thermodynamic quality. Its external novelty rate remains close to PT, whereas stable and metastable rates decrease, lowering both SUN and MSUN.

Identity metric isolates the value of lattice-aware paths. The decrease in MSUN across both evaluations, together with lower external SUN, shows the benefit of selecting periodic images by physical distance rather than fractional distance alone.

B.3 Interaction with anti-annealing

Anti-annealing raises mean MSUN for all four runs, but compensates for the geometric ablations more than for the loss of null species (Figure 3a).

Geometric deficits narrow, but coverage remains limiting. In the matched internal comparison, $a = 4$ reduces the No rotary deficit and nearly closes the mean Identity metric gap, while widening the No null species deficit (Table 6). Stronger coordinate updates can thus offset much of the MSUN loss from the geometric ablations, but do not recover the coverage supplied by null species. The external $a = 4$ results likewise show the largest deficit for No null species.

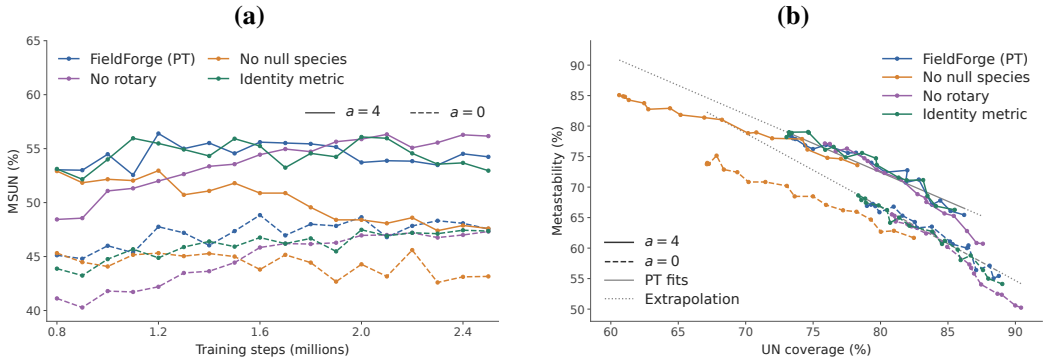


Figure 3: Pretraining ablations. **(a)** MSUN versus training steps. **(b)** Metastability versus pre-relaxation UN coverage. Colours identify models; solid/dashed lines denote $a = 4/0$. Grey lines fit PT separately for each sampler, with dotted extrapolations.

Rotary attention improves training efficiency. A smaller average gap does not eliminate the difference in learning speed: without rotary attention, high MSUN is reached later in training (Figure 3a).

Why anti-annealing helps. Larger late-stage coordinate updates (Equation (8)) may improve local geometric refinement, partly compensating for weaker geometric modeling. Empirically, $a = 4$ increases metastability while reducing UN coverage, with a net gain in MSUN (Figure 3b).

C SUPERVISED FINE-TUNING DETAILS

QB-SFT selects a subset of K structures from the original training corpus, balancing thermodynamic quality with chemical and structural coverage. We first construct candidate pools, allocate stratified quotas, compute candidate priorities from energy weights and element-incidence correction, and select structurally diverse examples. When an on-hull fraction is specified, quota allocation and selection are performed separately within the two pools before merging the results.

C.1 SFT SELECTION PROCEDURE

Step 1: candidate pools and on-hull allocation. We construct a low-energy core using the threshold $E_{\text{hull}} \leq T$. The main SFT results use $T = 0.025\,\text{eV/atom}$ on both datasets, as does the Alex-MP-20 on-hull scan in Figure 2.

For a requested on-hull fraction f, candidates are divided into an on-hull pool, $|E_{\text{hull}}| \leq \epsilon_{\text{hull}}$, and a near-hull pool, $\epsilon_{\text{hull}} < E_{\text{hull}} \leq T$, where $\epsilon_{\text{hull}} = 10^{-8}\,\text{eV/atom}$. The two pools receive target sizes of $\text{round}(fK)$ and $K - \text{round}(fK)$, respectively. Without a specified f, selection uses the low-energy core as a single pool. This explicit allocation makes f a controlled variable rather than a by-product of energy weighting; Section 5.2 varies it.

If a pool is undersized, all available structures are retained and the remaining quota is filled by priority-based backfilling using w, defined in Step 4. The realized on-hull fraction is recorded for each run.

Step 2: stratified quota allocation. To maintain coverage across element counts, cell sizes, and crystal systems, we group candidates by element-count bin $b_e \in \{1, 2, 3, 4, 5+\}$, atom-count bin $b_n \in \{1\text{–}4, 5\text{–}8, 9\text{–}12, 13\text{–}16, 17\text{–}20, 21+\}$, and crystal system b_c (triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, or cubic). For stratum $s = (b_e, b_n, b_c)$, containing N_s candidates in the current pool, its quota is

$$q_s \propto N_s^{\alpha}, \qquad \alpha = 0.7. \tag{9}$$

A largest-remainder correction makes the quotas sum to the target size of that pool. Thus, $\alpha = 1$ preserves empirical stratum proportions, whereas $\alpha < 1$ increases the relative allocation to less common strata.

Step 3: soft quality weighting. Lower-energy candidates receive higher priority. Let $e_x = \max(E_{\text{hull}}(x), 0)$. We assign

$$w_Q(x) = \exp\left(-e_x/\tau\right), \qquad \tau = 0.02\,\text{eV/atom}. \tag{10}$$

The threshold T determines candidate eligibility, while the weighting temperature τ controls how rapidly priority decreases with energy within the candidate set.

Step 4: element-incidence correction. Low-energy filtering can underrepresent particular elements. Let $p_{\text{full}}(Z)$ and $p_{\text{core}}(Z)$ denote the fractions of structures containing element Z in the full training set and low-energy core, respectively. For the element set $Z(x)$ of a candidate, we apply

$$w_E(x) = \text{clip}\left[\exp\left\{\gamma\,\frac{1}{|Z(x)|}\sum_{Z\in Z(x)}\log\frac{p_{\text{full}}(Z)+\epsilon}{p_{\text{core}}(Z)+\epsilon}\right\}, w_{\min}, w_{\max}\right], \tag{11}$$

with $(\gamma, w_{\min}, w_{\max}) = (0.5, 0.5, 2.0)$ and $\epsilon = 10^{-6}$. The final priority is $w(x) = w_Q(x)w_E(x)$. This clipped correction mitigates element-frequency shifts without forcing exact agreement with the full training distribution.

Step 5: diversity-aware structural selection. Within each stratum, we retain approximately $4q_s$ high-priority candidates and apply weighted farthest-point sampling (FPS). Each structure is represented by scale-free lattice ratios, normalized cell angles, $\log(V/N_{\text{atoms}})$, clipped atom count, and a 32-bin histogram of periodic minimum-image pair distances up to 8 Å. The histogram uses 27 periodic

image shifts and upper-triangular atom pairs, and is normalized to unit sum. Descriptor dimensions are standardized over the complete input set.

The first selected structure has the highest priority. Given the selected set S_k, subsequent structures are chosen by

$$x_{k+1} = \arg\max_{x \notin S_k} \left[\frac{\min_{x_i \in S_k} d(x, x_i)}{\max_{x' \notin S_k} \min_{x_i \in S_k} d(x', x_i)} \right] \left[\frac{w(x)}{\max_{x'} w(x')} \right]^{\lambda}, \qquad \lambda = 0.25, \tag{12}$$

where d is Euclidean descriptor distance and the maximizations are over the retained candidates in the current stratum. This balances structural distance from previously selected examples against candidate priority. Setting $\lambda = 0$ gives pure FPS after the first selection; disabling FPS replaces it with priority ranking. Selected structures are merged across strata and, when applicable, across the two energy pools.

Implementation and audit. Data selection is performed independently of generator training. For each selected subset, we record the realized on-hull fraction, mean energy, low-hull fraction, categorical distributions, and element-incidence Jensen–Shannon divergence. Table 8 summarizes the selector parameters.

Table 8: Parameters of the SFT data selector. N denotes the number of structures in the training split and K the requested subset size. Recommended ranges are reported when the implementation accepts a broader numeric domain.

Parameter	Default	Recommended range / choices	Meaning
`input`	required	train split	Input MP-20-style training table; the test split is never used for SFT
`mode`	`qb`	`random`, `on_hull`, `low_hull`, `qb`	Selection strategy or ablation control
`target-size`	5000	$1 \le K \le N$	Number of selected structures
`seed`	42	integer	Random seed for sampling and stratum order
`jobs`	CPU count	integer ≥ 1	Parallel CIF parsing workers
`low-hull-threshold`	0.025	0–0.1 eV/atom	Upper edge of low-energy core / near-hull pool
`on-hull-eps`	10^{-8}	10^{-12}–10^{-6} eV/atom	Tolerance for $E_{\mathrm{hull}} \approx 0$
`on-hull-fraction`	`None`	$[0, 1]$	Explicit fraction assigned to on-hull pool; QB only
`tau`	0.02	0.005–0.05 eV/atom	Temperature in $w_Q = \exp(-E_{\mathrm{hull}}/\tau)$
`alpha`	0.7	0–1	Stratum quota exponent in $q_s \propto N_s^{\alpha}$
`element-gamma`	0.5	0–1	Strength of element-incidence correction
`weight-clip-low`	0.5	0.1–1	Lower bound of element correction
`weight-clip-high`	2.0	1–5	Upper bound of element correction
`min-per-stratum`	0	integer ≥ 0	Optional minimum quota per occupied stratum
`rdf-bins`	32	8–128	Number of periodic pair-distance histogram bins
`rdf-rmax`	8.0	4–12 Å	Maximum pair distance used by the descriptor
`candidate-factor`	4.0	1–10	Candidate pool multiplier relative to q_s
`fps-quality-power`	0.25	0–1	Quality exponent in weighted FPS; 0 is pure FPS after first point
`no-fps`	false	boolean flag	Disable FPS and rank directly by selection priority

D Supervised fine-tuning analysis and ablations

D.1 Targeted SFT ablations

We evaluate random SFT, the core-energy threshold, and subset size (Table 9). The random rows control for additional supervised updates without a directed selection rule. The size comparisons use matched thresholds and on-hull fractions; the core-threshold comparisons use matched sizes and on-hull fractions within each dataset.

Table 9 shows that SFT performance depends on both data selection and subset size. On MP-20, quality-based selection outperforms random SFT at the same subset size, whereas the matched-size comparison on Alex-MP-20 yields similar performance. The effect of the core threshold is dataset-dependent: increasing T improves MSUN on Alex-MP-20 but reduces it on MP-20. Larger subsets likewise do not consistently improve performance, benefiting Alex-MP-20 in the matched comparison but not MP-20. The 200,000-structure Alex-MP-20 scan also departs from the approximately linear trend observed in the smaller scans.

Table 9: SFT selection ablations. K is the subset size, T the energy threshold (eV/atom), and f_{on} the realized on-hull fraction. Rates are percentages; SUN+MSUN is the sum of the two disjoint joint rates.

Setting	K	T	f_{on}	SUN	MSUN	SUN+MSUN
MP-20						
Main-result SFT	5,000	0.025	0.0000	1.88	33.96	35.84
f_{on} scan	5,000	0.025	0.3000	2.84	31.84	34.68
f_{on} scan	5,000	0.025	0.5000	2.28	31.36	33.64
f_{on} scan	5,000	0.025	0.6676	2.60	32.04	34.64
f_{on} scan	5,000	0.025	0.7000	2.52	30.44	32.96
f_{on} scan	5,000	0.025	1.0000	4.20	26.08	30.28
Random SFT	5,000	—	0.4380	3.00	30.12	33.12
T scan	5,000	0.100	0.0000	1.96	32.56	34.52
K scan	10,000	0.100	0.0000	3.04	31.40	34.44
Alex-MP-20						
Random SFT	5,000	—	0.0818	2.60	41.84	44.44
Random SFT	87,130	—	0.0807	2.16	43.44	45.60
T scan	43,565	0.025	0.0000	2.52	40.00	42.52
T scan	43,565	0.100	0.0000	2.32	40.72	43.04
f_{on} scan	87,130	0.100	0.0000	1.88	43.60	45.48
f_{on} scan	87,130	0.100	0.0500	2.12	43.32	45.44
f_{on} scan	87,130	0.100	0.1500	1.76	43.64	45.40
f_{on} scan	87,130	0.100	0.2500	1.76	42.28	44.04
f_{on} scan	87,130	0.100	0.3500	2.28	42.72	45.00
f_{on} scan	87,130	0.100	0.5000	2.68	40.64	43.32
Main-result SFT	115,000	0.025	0.0000	4.92	42.36	47.28
f_{on} scan	115,000	0.025	0.0379	4.24	41.28	45.52
f_{on} scan	115,000	0.025	0.1137	2.96	41.96	44.92
f_{on} scan	115,000	0.025	0.1894	3.84	40.12	43.96
f_{on} scan	115,000	0.025	0.2652	2.92	39.24	42.16
f_{on} scan	115,000	0.025	0.3788	3.36	38.20	41.56
T scan	115,000	0.100	0.0000	2.04	43.24	45.28
f_{on} scan	115,000	0.100	0.0379	2.04	43.16	45.20
f_{on} scan	115,000	0.100	0.1137	2.04	42.28	44.32
f_{on} scan	115,000	0.100	0.1894	2.08	42.92	45.00
f_{on} scan	115,000	0.100	0.2652	2.04	41.96	44.00
f_{on} scan	115,000	0.100	0.3788	2.48	41.28	43.76
f_{on} scan	200,000	0.100	0.0000	1.96	43.84	45.80
f_{on} scan	200,000	0.100	0.0218	1.84	43.00	44.84
f_{on} scan	200,000	0.100	0.0654	1.88	42.24	44.12
f_{on} scan	200,000	0.100	0.1089	1.80	42.36	44.16
f_{on} scan	200,000	0.100	0.1525	2.08	42.92	45.00
f_{on} scan	200,000	0.100	0.2178	2.20	43.28	45.48

D.2 Control correlations

For each selected subset, we collect the on-hull fraction, mean E_{hull}, low-hull fraction, three-element share, 9–12-atom share, cubic share, and element-incidence divergence:

$$u = (f_{\text{on}}, \overline{E}_{\text{hull}}, f_{\text{low}}, f_{\text{3elem}}, f_{\text{9–12}}, f_{\text{cubic}}, D_{\text{JS}}). \tag{13}$$

Table 10 correlates these statistics with downstream generation metrics. Composition, atom-count, symmetry, and element-incidence statistics co-vary with the scan, but are not independently varied; these correlations do not establish separate causal effects.

D.3 Control scans

The SFT scan diagnostics are placed here to keep the dataset-specific normalization explicit. Figure 4 reports both the MP-20 5,000 and 115,000 Alex-MP-20 ($T = 0.025$) QB-SFT scans side by side. In each panel, both axes are normalized by the maximum observed on the corresponding scan.

Table 10: Pearson correlations between selected-subset statistics and SUN, MSUN, and MSUN+SUN. Each cell gives $(r_{\text{MP20}}, r_{\text{Alex-115,000}})$; the Alex scan uses $T = 0.025$. The on-hull-fraction row uses the six points shown in each lower panel of Figure 2. A dash denotes a constant factor.

Subset factor	SUN	MSUN	MSUN+SUN
On-hull fraction	$(+0.78, -0.68)$	$(-0.88, -0.95)$	$(-0.88, -0.97)$
Mean E_{hull}	$(-0.75, +0.66)$	$(+0.91, +0.95)$	$(+0.94, +0.97)$
Low-hull fraction	$(+0.05, -)$	$(-0.09, -)$	$(-0.11, -)$
3-element share	$(+0.89, -0.72)$	$(-0.94, -0.91)$	$(-0.92, -0.96)$
9–12-atom share	$(+0.77, +0.84)$	$(-0.89, +0.72)$	$(-0.91, +0.85)$
Cubic share	$(+0.93, -0.78)$	$(-0.98, -0.89)$	$(-0.96, -0.96)$
Element-incidence D_{JS}	$(-0.02, +0.46)$	$(-0.08, -0.21)$	$(-0.12, +0.01)$

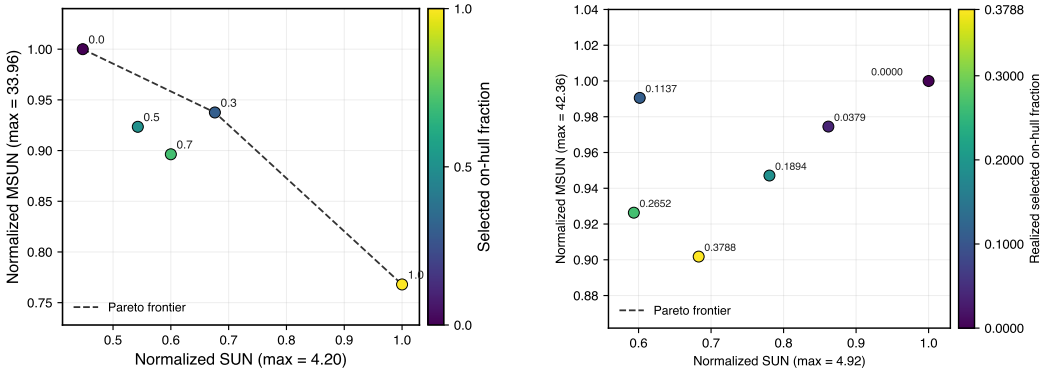


Figure 4: Normalized SUN–MSUN trade-off over the QB-SFT scans. Left: MP-20 5,000; right: 115,000 Alex-MP-20 ($T = 0.025$). Each point corresponds to one selected on-hull fraction (labels); the left panel contains five MP-20 points at $f_{\text{on}} = 0, 0.3, 0.5, 0.7, 1.0$ and the right panel contains six Alex-MP-20 points at realized fractions $f_{\text{on}} = 0, 0.0379, 0.1137, 0.1894, 0.2652, 0.3788$. The additional MP-20 point at $f_{\text{on}} = 0.6676$ is shown in Figure 2 but is not included in this five-point diagnostic. SUN and MSUN are divided by their respective scan maxima (4.20/33.96 and 4.92/42.36). Dashed lines connect the non-dominated points under simultaneous maximization; the Alex scan has a single non-dominated point at $f_{\text{on}} = 0$. Each panel is a within-dataset diagnostic, not a cross-dataset calibration.

Section 5.2 reports the scan-level correlations and interprets them as shifts in the stable–metastable generation balance. The selector is deliberately separated from the generator backbone: the same procedure can be paired with another crystal generator that consumes equivalent structure fields, with thresholds and descriptor ranges re-estimated from its training distribution.

E Reinforcement learning details

E.1 Transition density

The stochastic policy of Equation (3) is Gaussian within each output channel. For the non-periodic type and lattice channels, the log density over active scalars is

$$\log \pi_\theta^c(a_k^c \mid s_k) = -\frac{1}{2} \sum_{j \in \mathcal{A}_c} \left[\frac{(a_{k,j}^c - \mu_{\theta,k,j}^c)^2}{(\sigma_k^c)^2} + \log\bigl(2\pi(\sigma_k^c)^2\bigr) \right], \tag{14}$$

where $\mathcal{A}_c$ denotes the active entries of channel c. Type and fractional-coordinate likelihoods are evaluated only over active atom slots, whereas the lattice likelihood always contains its six scalar degrees of freedom.

Fractional coordinates are periodic. The corresponding exact transition density is therefore a wrapped Gaussian,

$$p_{\mathbb{T}}(a \mid \mu, \sigma) = \sum_{n \in \mathbb{Z}^d} \mathcal{N}(a + n; \mu, \sigma^2 I). \tag{15}$$

Rather than evaluating the infinite image sum, we use the nearest periodic residual

$$\delta_k^{\mathrm{frac}} = \mathrm{wrap}\left(x_{t_{k+1}}^{\mathrm{frac}} - \mu_{\theta,k}^{\mathrm{frac}}\right), \qquad \mathrm{wrap}(z) = z - \mathrm{round}(z), \tag{16}$$

inside Equation (14). This approximation is appropriate when the transition mass is concentrated around a single periodic image and, critically, treats actions separated only by a unit-cell boundary as nearby rather than maximally distant.

The likelihood uses a componentwise torus wrap rather than a lattice-metric search over periodic images. The log likelihood of one flow transition is

$$\ell_{\theta,k} = \sum_c \log \pi_\theta^c(a_k^c \mid s_k). \tag{17}$$

E.2 Objective and safeguards

We use the policy objective and total loss in Equations (5) and (6), with $\log \pi_\theta$ given by the transition log likelihood $\ell_{\theta,k}$ in Equation (17). For unconditional generation, the full rollout batch forms the group for advantage normalization in Equation (4). The current and reference policies share the same channel-wise transition variances; their reference penalty is defined in Equation (7).

Let $\overline{K}_q$ denote the mean reference displacement at RL iteration q, and let $\widetilde{g}_q$ denote the gradient after global norm clipping. An update is skipped if the reference-displacement safeguard exceeds κ,

$$\theta_{q+1} = \begin{cases} \mathrm{AdamW}(\theta_q, \widetilde{g}_q), & \overline{K}_q \le \kappa, \\ \theta_q, & \overline{K}_q > \kappa, \end{cases} \tag{18}$$

and the frozen reference policy is refreshed periodically,

$$\pi_{\mathrm{ref}}^{q+1} = \begin{cases} \pi_{\theta_{q+1}}, & q+1 \equiv 0 \pmod H, \\ \pi_{\mathrm{ref}}^{q}, & \text{otherwise.} \end{cases} \tag{19}$$

Here H is the reference-refresh period, set to five RL iterations (Table 11).

Each rollout contains $T = 50$ flow transitions, whereas gradients are evaluated on $K = 25$ transitions selected by stratified temporal subsampling. This reduces the memory cost of backpropagation while maintaining coverage of early, middle, and late portions of the flow trajectory. The $t = 0$ transition is excluded because the exploration variance vanishes at the endpoint.

Reference-penalty interpretation. For the fractional-coordinate channel, $\mu_{\theta,k}^F$ and $\mu_{\mathrm{ref},k}^F$ in Equation (7) denote the pre-wrap Euler–Maruyama transition means,

$$\mu_{\theta,k}^F = F_k + \Delta t_k\, v_\theta^F(x_k, t_k), \qquad \mu_{\mathrm{ref},k}^F = F_k + \Delta t_k\, v_{\mathrm{ref}}^F(x_k, t_k). \tag{20}$$

Because both policies are evaluated from the same current state F_k, their displacement is

$$\mu_{\theta,k}^F - \mu_{\mathrm{ref},k}^F = \Delta t_k \left[v_\theta^F(x_k, t_k) - v_{\mathrm{ref}}^F(x_k, t_k)\right]. \tag{21}$$

Consequently, crossing a unit-cell boundary by the common transition mean does not introduce a discontinuous order-one displacement: the shared coordinate origin cancels before the Euclidean norm is evaluated. Periodic wrapping is applied only after sampling the next fractional-coordinate state, whereas its action likelihood is evaluated using the nearest-image residual of Equation (16).

Equation (7) should therefore be interpreted as a local regularizer on changes in the fractional-coordinate velocity field, rather than as the exact KL divergence between two wrapped transition distributions. This choice matches the implementation used for all reported RL results.

E.3 Terminal reward

Let e_i denote the energy above hull obtained from the in-loop structural relaxation and let

$$f_i = \mathbf{1}\{e_i \text{ is finite}\} \tag{22}$$

indicate whether physical evaluation succeeds.

For the dense energy term, a failed evaluation is assigned the upper clipping value,

$$\bar{e}_i = \begin{cases} e_i, & f_i = 1, \\ 3, & f_i = 0, \end{cases} \tag{23}$$

and the corresponding component is

$$r_i^{\text{energy}} = -\operatorname{clip}(\bar{e}_i, 0, 3). \tag{24}$$

Two additional terms identify the thermodynamic regions emphasized during training,

$$r_i^{\text{meta}} = f_i \, \mathbf{1}\{e_i \le 0.1\}, \qquad r_i^{\text{stable}} = f_i \, \mathbf{1}\{e_i \le 0\}. \tag{25}$$

The dense energy term preserves ordering across finite-energy structures, whereas the two indicator terms increase the relative weight assigned to near-hull and on-hull candidates.

Chemical validity is represented by

$$r_i^{\text{valid}} = f_i \, \mathbf{1}\{\text{the decoded composition passes the SMACT test}\}, \tag{26}$$

using a SMACT-based composition-validity test (Davies et al., 2019). The test searches for chemically consistent oxidation states satisfying charge balance together with the corresponding electronegativity constraints.

Creativity. The training-time creativity proxy combines composition novelty, within-batch uniqueness, and a continuous structural distance. Let q_i be the reduced composition of sample i, let $\mathcal{Q}_{\text{train}}$ be the set of reduced compositions in the training reference set, and define

$$n_i = \mathbf{1}\{q_i \notin \mathcal{Q}_{\text{train}}\}. \tag{27}$$

The reference set is sampled once without replacement from the training split with pseudorandom seed 0 and kept fixed across rollouts. For the inexpensive uniqueness proxy, we form the key

$$\text{key}_i = (q_i, \text{round}_{0.1\,\text{Å}}(a_i, b_i, c_i)) \tag{28}$$

from the reduced composition and lattice lengths. We set $u_i = 1$ if key_i has not appeared earlier in the current rollout batch and $u_i = 0$ otherwise.

Let $h_i = \text{AMD}_{100}(x_i) \in \mathbb{R}^{100}$ be the Average Minimum Distance descriptor formed from the first 100 periodic-neighbor distances. For samples with the same reduced composition, define the comparison set

$$\mathcal{C}_i = \{h_j : j \neq i, \; q_j = q_i\} \cup \mathcal{H}_{\text{ref}}(q_i) \tag{29}$$

using the other rollout samples and the fixed training reference. The boundary distance is the nearest Chebyshev distance,

$$d_i = \min_{h \in \mathcal{C}_i} \|h_i - h\|_\infty, \tag{30}$$

with $d_i = 1$ when $\mathcal{C}_i$ is empty. The creativity reward is

$$r_i^{\text{creative}} = f_i \begin{cases} 1, & n_i = 1 \text{ and } u_i = 1, \\ 0, & n_i = 0 \text{ and } u_i = 0, \\ \min\{d_i, 1\}, & n_i \neq u_i \text{ and } h_i \text{ is defined}, \\ 0, & n_i \neq u_i \text{ and } h_i \text{ is undefined}. \end{cases} \tag{31}$$

Thus the binary novel-and-unique endpoint receives unit reward, the non-novel-and-duplicate endpoint receives zero reward, and the two boundary cases receive a continuous geometry-aware score.

Composition and structural diversity. For composition diversity, each structure is represented by a 95-dimensional element-fraction vector g_i^{comp}, whose entry for atomic number Z is the fraction of atoms of element Z. For structural diversity, we use $h_i = \text{AMD}_{100}(x_i)$ and standardize each AMD coordinate using the mean and standard deviation of the structural reference set,

$$g_i^{\text{struct}} = \frac{h_i - \mu_{\text{AMD}}}{\sigma_{\text{AMD}} + 10^{-6}}. \tag{32}$$

Both feature spaces use the degree-two polynomial kernel

$$K(x, y) = (x^\top y + 1)^2. \tag{33}$$

For each feature type $t \in \{\text{comp}, \text{struct}\}$, the implementation forms a fixed-length generated feature multiset $G^t = \{g_i^t\}_{i=1}^B$ over all rollout positions and a fixed reference feature set $H^t = \{h_j^t\}_{j=1}^{N_t}$. When a feature vector cannot be constructed, a zero vector of the corresponding dimension is used in the kernel computation; that sample is subsequently excluded from the relevant normalization set below. Thus, in the reported configuration, the generated-set size in the diversity calculation is $M = B = 96$, independently of the size of the normalization set $\mathcal{V}_t$.

For compactness, write $G = \{g_i\}_{i=1}^M$ and $H = \{h_j\}_{j=1}^N$. The implementation uses the generated-set-dependent part of the negative MMD score,

$$\begin{aligned} \mathcal{D}_{\text{gen}}(G, H) &= -\frac{1}{M(M-1)} \sum_{i \neq j} K(g_i, g_j) + \frac{2}{MN} \sum_{i=1}^{M} \sum_{p=1}^{N} K(g_i, h_p), \quad M \geq 2,\ N \geq 1, \\ \mathcal{D}_{\text{gen}}(G, H) &= 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad M < 2 \text{ or } N = 0. \end{aligned} \tag{34}$$

For $M, N \geq 2$, the full negative unbiased squared maximum mean discrepancy is

$$\mathcal{D}_{\text{gen}}(G, H) - \frac{1}{N(N-1)} \sum_{p \neq q} K(h_p, h_q).$$

The second term depends only on the fixed reference set and cancels from every leave-one-out difference. The zero-valued branch in Equation (34) is the implementation convention for an empty or singleton generated set and makes the boundary case explicit.

The raw per-sample leave-one-out utility is therefore

$$\widehat{r}_i^{\text{div}} = \mathcal{D}_{\text{gen}}(G, H) - \mathcal{D}_{\text{gen}}(G \setminus \{g_i\}, H). \tag{35}$$

The composition and structural utilities are computed independently in their respective feature spaces. Let $\mathcal{V}_{\text{comp}} = \mathcal{V}$ and $\mathcal{V}_{\text{struct}} = \{i \in \mathcal{V} : h_i \text{ is defined}\}$. These sets are used only for component-wise min–max normalization and do not redefine the generated feature multiset G^t. For $t \in \{\text{comp}, \text{struct}\}$, the normalized diversity component is

$$r_i^{t\text{div}} = f_i \frac{\widehat{r}_i^{t\text{div}} - \min_{m \in \mathcal{V}_t} \widehat{r}_m^{t\text{div}}}{\max_{m \in \mathcal{V}_t} \widehat{r}_m^{t\text{div}} - \min_{m \in \mathcal{V}_t} \widehat{r}_m^{t\text{div}}}. \tag{36}$$

If $|\mathcal{V}_t| < 2$, the required feature or reference set is unavailable, or the denominator is below 10^{-12}, the corresponding normalized component is set to zero. This gives $r_i^{\text{compdiv}}, r_i^{\text{structdiv}} \in [0, 1]$.

For every active component j, statistics are computed over

$$\mathcal{V} = \{i : f_i = 1\}, \tag{37}$$

with population statistics

$$\mu_j = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} r_i^j, \qquad s_j = \left[\frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} (r_i^j - \mu_j)^2 \right]^{1/2}. \tag{38}$$

The component is standardized as

$$z_i^j = \frac{r_i^j - \mu_j}{s_j + 10^{-6}}. \tag{39}$$

The provisional aggregate reward used in the reported single-stage RL procedure is

$$\begin{aligned} \widetilde{R}_i = z_i^{\text{energy}} + 2z_i^{\text{meta}} + z_i^{\text{stable}} + z_i^{\text{valid}} \\ + 1.5z_i^{\text{creative}} + z_i^{\text{compdiv}} + 0.25z_i^{\text{structdiv}}. \end{aligned} \tag{40}$$

A reward component that is constant over the finite-energy subset has zero standardized variance and therefore contributes no ranking signal in that rollout. For $|\mathcal{V}| \geq 2$, the final trajectory return is

$$R_i = \begin{cases} \widetilde{R}_i, & i \in \mathcal{V}, \\ \min_{m \in \mathcal{V}} \widetilde{R}_m - 1, & i \notin \mathcal{V}. \end{cases} \tag{41}$$

Thus a failed physical evaluation is exactly one reward unit below the worst successfully evaluated trajectory in the rollout. It is retained in the policy-gradient batch and participates in the whole-batch advantage normalization of Equation (4). When $|\mathcal{V}| < 2$, the implementation falls back to the unstandardized weighted component sum.

Only the seven components in Equation (40) contribute to the reported single-stage RL policy updates.

The novelty, uniqueness, and diversity quantities used during RL are training-time proxies.

E.4 Reinforcement-learning schedule

The reported method uses a single reinforcement-learning stage initialized from the EMA weights of the best quality-balanced SFT checkpoint. No intermediate reinforcement-learning stage or intermediate RL-selected checkpoint is used.

Table 11 lists the training hyperparameters. The 50-step stochastic rollouts reduce RL cost, whereas checkpoint selection and reported DNG evaluation use the finer 150-step deterministic Euler sampler. Both integrate the same learned velocity field; the finer evaluation discretization does not require a separately trained policy.

The regular policy itself is optimized, checkpointed, selected, and evaluated. EMA weights are used only to initialize reinforcement learning from the SFT checkpoint; no separate RL EMA policy is maintained for reporting.

Table 11: Reported reinforcement-learning configuration on MP-20.

Parameter	Setting	Parameter	Setting
Initialization	Best SFT checkpoint; EMA weights	Reference coefficient β	5×10^{-4}
RL iterations	40	Reference skip threshold κ	5.0
Batch size	96	Reference refresh period	every 5 iterations
Rollout transitions	50	Global gradient-norm bound	2.0
DNG evaluation steps	150, deterministic	Shallow / deep relaxation	50 / 200 steps
Fractional residual	nearest periodic image	Energy sanitization floor	$-0.5\,\mathrm{eV/atom}$
Exploration	$(0.15,\ 0.15,\ 0.15)$	Deep-relaxation cutoff	$0.15\,\mathrm{eV/atom}$
Gradient transitions	25, stratified	Dense energy / near-hull / stable	$(1,\ 2,\ 1)$
Epochs per rollout	1	Creativity / composition diversity	$(1.5,\ 1)$
Learning rate	8×10^{-5}	Structure diversity / chemical validity	$(0.25,\ 1)$

E.5 Terminal physical evaluation

Two-pass relaxation. The terminal evaluator first applies numerical and geometric checks before calling the interatomic potential. A decoded structure is rejected if it contains non-finite coordinates or lattice parameters, has $V < 10^{-3}\,\text{Å}^3$ or $V/N < 10^{-3}\,\text{Å}^3$, has a smallest lattice singular value below $10^{-3}\,\text{Å}$, has lattice condition number above 10^6, has normalized Gram determinant below 10^{-8}, or, for $N \geq 2$, has a minimum interatomic distance below $0.5\,\text{Å}$.

Structures that pass these checks first undergo a 50-step NequIP relaxation (Batzner et al., 2022). Candidates with a provisional $E_{\text{hull}} \leq 0.15\,\mathrm{eV/atom}$ are then relaxed for 200 steps, and the longer relaxation result replaces the initial one. If the second relaxation fails, the earlier value is not restored. This two-pass procedure concentrates the more expensive relaxation on structures close enough to the convex hull for additional optimization to affect their ranking.

Post-relaxation structures with a minimum interatomic distance below 0.5 Å are rejected. Energies below $-0.5\,\mathrm{eV/atom}$ are also treated as invalid, as such values indicate pathological relaxation outcomes rather than physically meaningful structures.

F Reinforcement learning diagnostics and ablations

F.1 RL checkpoint selection and final evaluation

We select RL checkpoints using a reduced internal evaluation and report final performance through a separate, full-budget evaluation. Regular-policy checkpoints saved after iterations 10, 20, 30, and 40 are ranked by SUN+MSUN, with ties resolved by higher MSUN, higher validity, and then the earlier iteration.

For checkpoint selection, we use 150 deterministic flow steps, coordinate anti-annealing $a_{\mathrm{coord}} = 4$, and sampling seed 21793. Each checkpoint generates 2,500 structures, from which 650 are sampled uniformly without replacement using the fixed seed and subjected to 200 steps of NequIP pre-relaxation. The fixed-seed procedure then constructs a set of 625 successfully exported relaxed CIFs for MACE hull evaluation. An evaluation is valid only if exactly 625 CIFs are available and at least 80% of the benchmark-valid structures receive finite MACE hull scores. All selection rates use the fixed denominator of 625, including invalid or failed benchmark samples.

After selection, the chosen checkpoint is evaluated independently using the 2,500-structure LeMat-GenBench protocol in Appendix G. All metrics are recomputed from the generated structures rather than taken from the training reward. Tables 1 and 2 report this final evaluation, not the reduced checkpoint-selection results.

F.2 Single-stage reward ablations

We perform matched leave-one-component-out diagnostics on the reported single-stage RL procedure. Each variant retains the SFT initialization, RL schedule, inference setting, and checkpoint-selection rule of the full method, but sets the indicated reward-component weight to zero. For every variant, the checkpoint is selected from iterations 10, 20, 30, and 40 using the protocol above. Table 12 reports the resulting $a_{\mathrm{coord}} = 4$ selection diagnostic.

Table 12: MP-20 reward-component ablations under the 625-structure internal checkpoint-selection diagnostic with MACE-hull scoring. Rates are percentages; Combined denotes MSUN+SUN; Δ is its percentage-point difference from the full reward.

Reward configuration	MSUN	SUN	Combined	Δ (pp)
Full reward	**41.44**	4.96	**46.40**	—
Without dense energy reward	39.52	4.80	44.32	−2.08
Without near-hull bonus	37.92	**6.24**	44.16	−2.24
Without structural diversity	38.24	4.64	42.88	−3.52
Without chemical validity	38.08	4.64	42.72	−3.68
Without composition diversity	38.24	4.32	42.56	−3.84
Without stability bonus	36.64	4.96	41.60	−4.80
Without creativity	35.84	4.16	40.00	−6.40

Every tested removal lowers the combined score under the matched reduced diagnostic. Removing creativity produces the largest decrease (-6.40 percentage points), followed by the stability bonus (-4.80), composition diversity (-3.84), and chemical validity (-3.68). Removing the near-hull bonus increases SUN from 4.96% to 6.24%, but the corresponding decrease in MSUN yields a net reduction of 2.24 percentage points in the combined score.

G Evaluation protocol

We evaluate two tasks. In de novo generation (DNG) the model produces composition and geometry together and is scored against a reference corpus; in crystal structure prediction (CSP) the composition

of a held-out crystal is supplied and the model must recover its structure. This appendix defines each reported quantity and the settings behind it.

G.1 Benchmark and sample budget

Cross-model DNG numbers are taken from LeMat-GenBench (Betala et al., 2025), a standardized implementation of the validity, uniqueness, novelty, and post-relaxation stability suite established with MatterGen (Zeni et al., 2025), together with a public leaderboard. Its value here is that the hull reference, novelty corpus, structure-matching settings, and sample budget are fixed across submissions, so leaderboard rows are comparable with one another in a way that independently run pipelines are not. We relax generated structures locally and submit the relaxed structures; the benchmark computes every reported metric. We therefore report external LeMat-GenBench values whenever models are compared against each other, and use our internal pipeline only for matched ablations, where every variant shares one protocol.

Each DNG evaluation uses ten sampling seeds and generates 250 crystals per seed, giving $N_{\text{gen}} = 2{,}500$ structures in total. All 2,500 structures are submitted to relaxation and thermodynamic scoring; they are not pre-filtered by novelty or uniqueness, which would bias the benchmark's own computation of those metrics.

G.2 De novo generation

Validity. Composition validity is decided by SMACT (Davies et al., 2019), which tests whether the generated stoichiometry admits a charge-neutral oxidation-state assignment. Structure validity requires the decoded cell to be constructible as a periodic structure and its minimum non-self interatomic distance to satisfy $d_{\min}(C) \geq 0.5\,\text{Å}$. Overall validity requires both checks, with the remaining structural gates and the treatment of rejected samples as implemented in LeMat-GenBench. These rates are reported for interpretability; they are not the eligibility filter for the metrics below.

Uniqueness, novelty, and UN. Structure comparisons use pymatgen's StructureMatcher (Ong et al., 2013); two crystals are duplicates when the matcher returns a valid alignment. Uniqueness is the fraction of evaluated structures remaining once duplicates are collapsed, novelty the fraction matching no structure in the reference corpus, and UN the rate of structures that are both. The matcher tolerances, the reference corpus of each track, and the construction of the unique-and-novel set are those of LeMat-GenBench (Betala et al., 2025).

Thermodynamic competitiveness. LeMat-GenBench evaluates the submitted structures against MLIP-specific convex hulls constructed from LeMat-Bulk (Betala et al., 2025). The two stability events are defined to be disjoint,

$$\mathcal{S} = \{C : E_{\text{hull}}(C) \leq 0\}, \qquad \mathcal{M} = \{C : 0 < E_{\text{hull}}(C) \leq 0.1\,\text{eV/atom}\}. \tag{42}$$

Writing $\mathcal{U}$ for the unique-and-novel subset of the N_{gen} generated crystals,

$$\text{SUN} = \frac{|\mathcal{U} \cap \mathcal{S}|}{N_{\text{gen}}}, \qquad \text{MSUN} = \frac{|\mathcal{U} \cap \mathcal{M}|}{N_{\text{gen}}}, \tag{43}$$

which are joint rates over one set rather than products of independently measured marginals, and are disjoint from each other. Their sum is therefore the fraction of generated crystals that are unique, novel, and within $0.1\,\text{eV/atom}$ of the hull, counted once. The hull snapshot, energy corrections, and treatment of failed relaxations follow LeMat-GenBench.

Stability calipers. Post-relaxation and as-generated stability are never interchanged. For the main DNG results, our generated structures are pre-relaxed locally with NequIP; the cross-potential analyses in Appendix H use the specified MLIP back-ends.

G.3 Crystal structure prediction

CSP@k draws k candidates per test composition from independent initial-noise seeds using the same deterministic sampler within each setting. A target is counted as matched if any candidate matches its ground-truth structure under pymatgen's `StructureMatcher` (version 2024.11.13), with `stol`$= 0.5$, `ltol`$= 0.3$, and `angle_tol`$= 10°$; all other arguments retain their defaults.

These tolerances follow CDVAE and DiffCSP (Xie et al., 2022; Jiao et al., 2023), rather than the DNG novelty and uniqueness protocol. Match rate is the fraction of test targets with at least one match. For each matched target, we take the minimum normalized RMS distance among matching candidates, using the first value returned by `get_rms_dist`; the reported RMSE is the mean of these minima over matched targets only. The distances are dimensionless, normalized by $(V/N)^{1/3}$.

Our CSP@1 and CSP@20 results use separately configured samplers (Appendix A), rather than nested subsets of a shared candidate pool.

Complete test splits. CSP is scored on complete test splits. The MPTS-52 test file is ordered by year, so scoring only a prefix selects the oldest and easiest entries and inflates the match rate; all reported CSP numbers use the complete split.

G.4 Seeds and reporting

Sampling seed is the dominant source of variance in both tasks, so comparisons use four-seed means for CSP and ten-seed pooled statistics for DNG.

H Additional results and analyses

H.1 Stability comparison

Figure 5 examines the energy distributions, relaxation displacements, and compositional complexity of the generated crystals, providing a more detailed view than the aggregate generation rates. FieldForge (PT) is concentrated at lower E_{hull} and more negative formation energies. Relaxation RMSD provides complementary structural information: it measures how far a generated crystal moves during relaxation, with smaller values indicating proximity to a relaxed local minimum.

The composition-count panels add a separate perspective on the candidates that satisfy the stability criteria. FieldForge (PT) retains a larger fraction of stable and metastable structures with four or more distinct elements than the other plotted generators. Its thermodynamically competitive candidates therefore extend beyond chemically simple compositions, which is useful for inverse design in complex multicomponent systems.

H.2 Robustness across interatomic potentials

Figure 6 examines whether the stability distributions persist across MACE, UMA, and ORB (Batatia et al., 2025; Wood et al., 2025; Rhodes et al., 2025). For each generator, the three potentials produce similar overall distributional shapes. The comparison covers both energetic quantities and the structural displacement during relaxation, rather than relying on a single stability summary. FieldForge (PT) retains its concentration at low E_{hull} and small relaxation displacements across these evaluations. Although the detailed distributions vary with the potential, these recurring features support the stability observations above and indicate that they are not specific to a single evaluator.

H.3 Crystallographic symmetry distributions

Figure 7 compares the symmetry distributions of generated crystals with MP-20 as a reference. The space-group panel shows coverage across crystallographic classes, while the Wyckoff panels provide complementary views of the generated site patterns. Although FieldForge (PT) does not enforce crystallographic symmetry, its samples retain broad coverage across these descriptors rather than concentrating on a single structural class. The distributions also reveal different degrees of concentration among the generators. This is a structural-distribution analysis rather than a symmetry-conditioned benchmark: it complements the joint generation metrics in Table 1 by showing the variety of crystallographic arrangements represented in the generated set.

H.4 Element distributions

Figure 8 compares elemental coverage in the training data and generated structures. Whereas the composition-count analysis above measures how many distinct elements occur in each crystal,

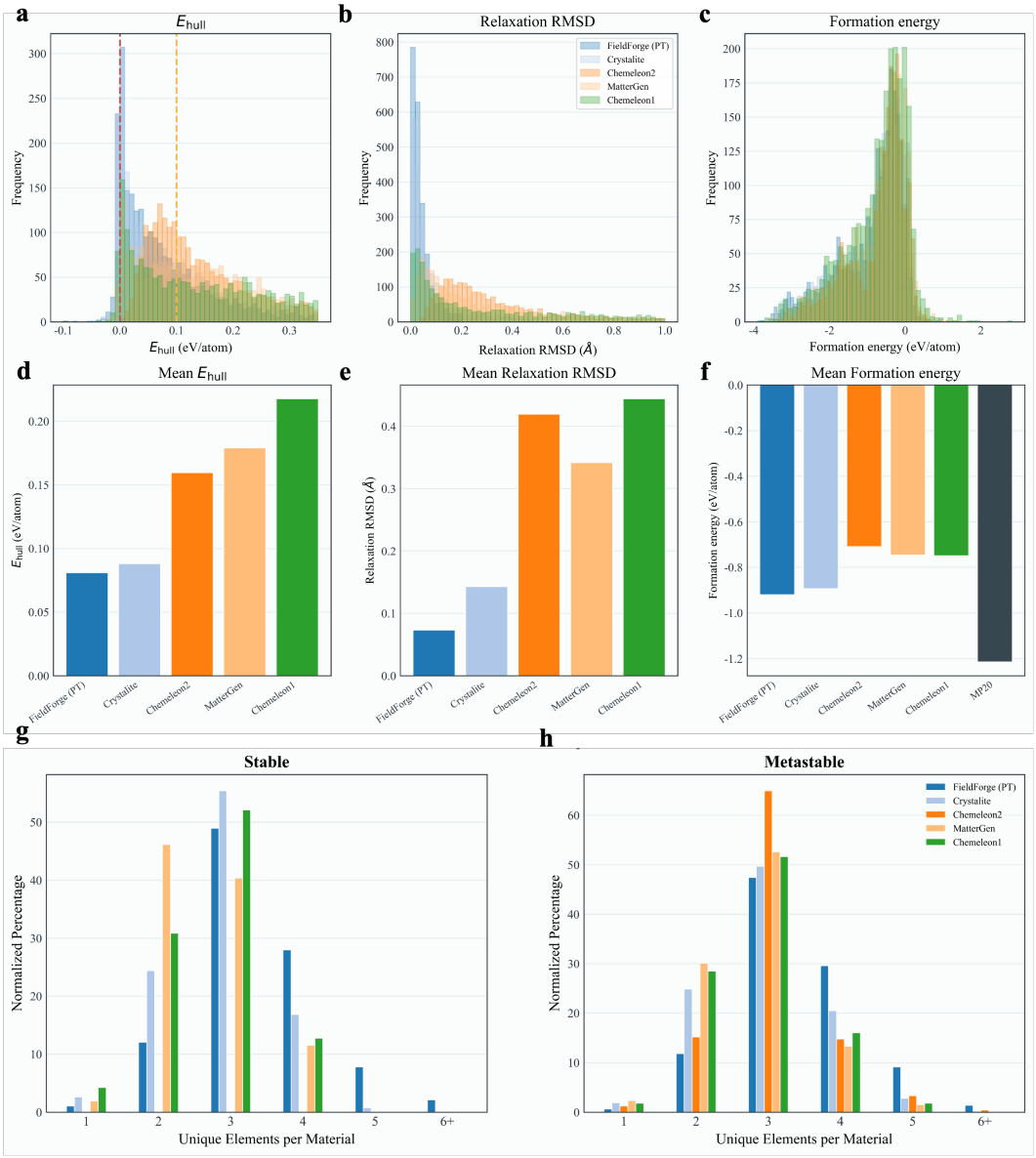


Figure 5: **Relaxed stability and compositional complexity.** The comparison uses 2,500 structures per model: FieldForge (PT), Crystalite, Chemeleon2, MatterGen, and Chemeleon1. MP-20 is shown as a training-data reference where indicated. All quantities follow the standard LeMat-GenBench protocol, including its prescribed relaxation, reference-hull, and formation-energy treatment. **(a–c)** Distributions of E_{hull}, relaxation RMSD, and formation energy per atom. **(d–f)** Corresponding means. **(g,h)** Unique-element counts among stable and metastable samples.

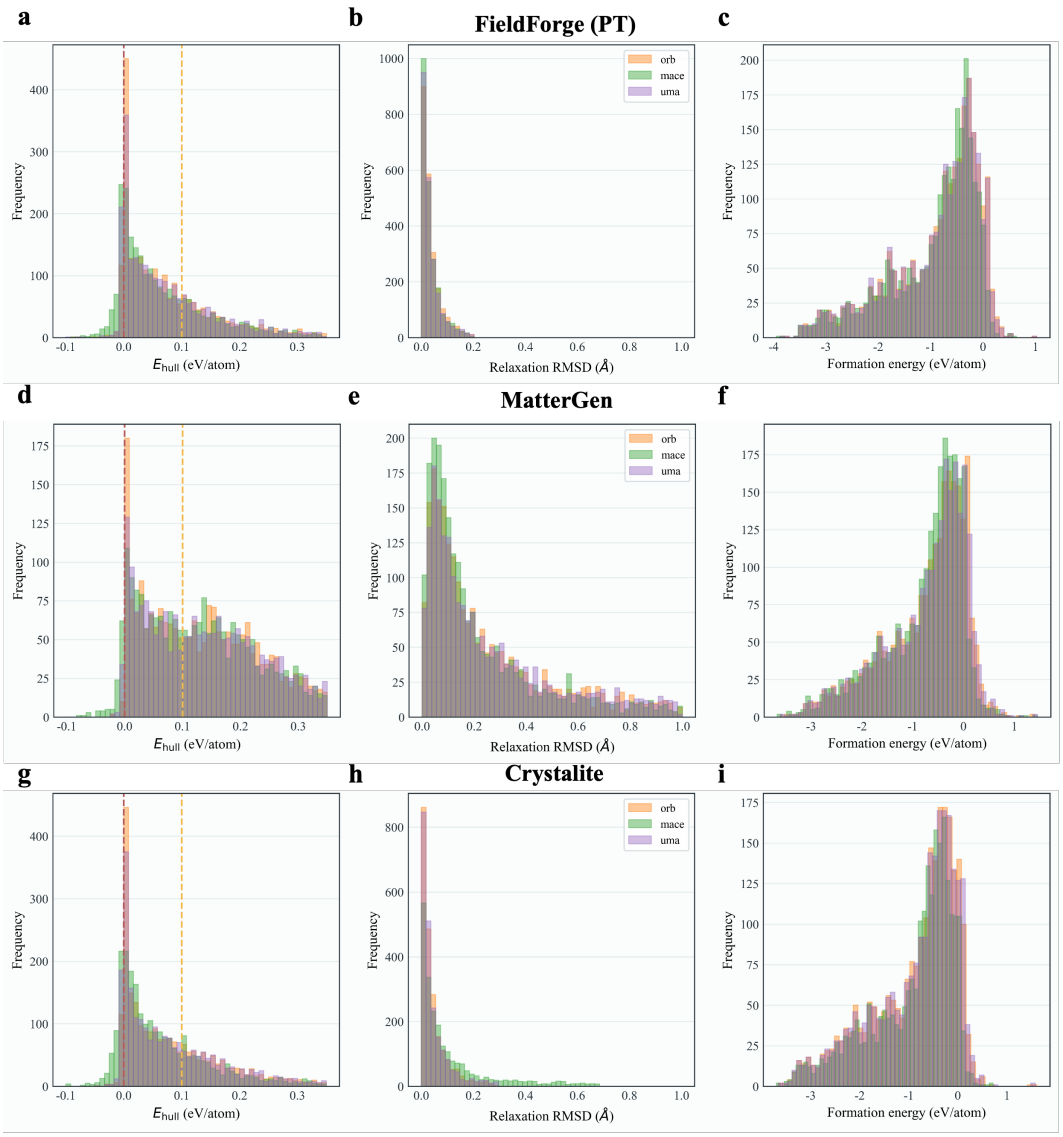


Figure 6: **Stability distributions under three MLIP back-ends.** Each potential uses its corresponding standard LeMat-GenBench relaxation, reference-hull, and formation-energy protocol. Rows show **(a–c)** FieldForge (PT), **(d–f)** MatterGen, and **(g–i)** Crystalite; columns report E_{hull}, relaxation RMSD, and formation energy per atom. Colours distinguish MACE, UMA, and ORB.

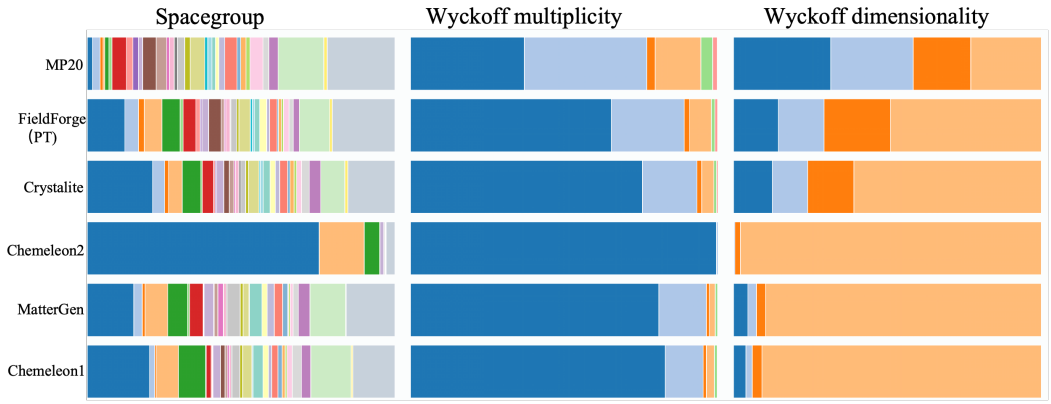


Figure 7: **Crystallographic symmetry distributions on** 2,500 **structures per model.** MP-20 is the reference dataset. FieldForge (PT) and Crystalite are relaxed, whereas MatterGen and Chemeleon1/2 are evaluated as generated. From left to right, the panels report detected space groups, Wyckoff multiplicities, and Wyckoff dimensionalities. Assignments use pymatgen's `SpaceGroupAnalyzer` with a distance tolerance of 0.1 Å. Unless a suffix is shown, Chemeleon denotes Chemeleon1 throughout this appendix.

this comparison identifies which elements receive greater representation across the generated set. FieldForge (PT) places more mass on several transition metals, including Co, Ni, and Cu, than MP-20 and Crystalite, while assigning less mass to the metalloid Te. Thus, the generated composition profile does not simply reproduce the elemental frequencies of the training corpus. The differences between generators reveal distinct elemental preferences and preferential exploration of different chemical subspaces.

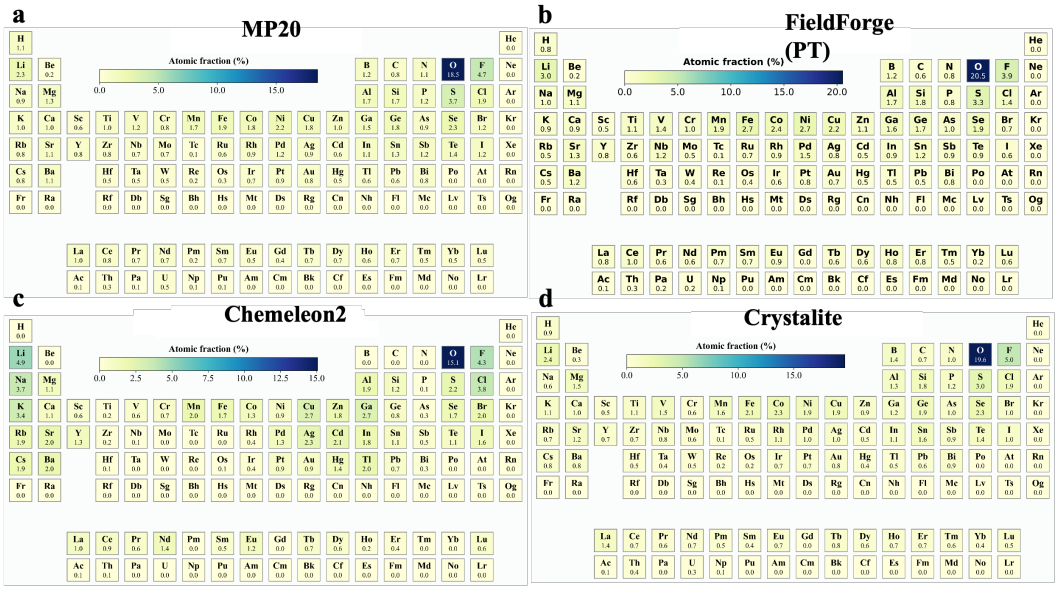


Figure 8: **Element distributions over the periodic table.** The plot uses the first 2,500 lexically sorted CIFs per model. Atomic fraction is an element's total atom count divided by the total number of parsed atoms across valid structures, rather than an equal-weight average of per-structure fractions. **(a–d)** MP-20, FieldForge (PT), Chemeleon2, and Crystalite. MP-20 is the reference dataset; FieldForge (PT) and Crystalite are relaxed, whereas Chemeleon2 is unrelaxed. Panels use different colour-scale ranges; cross-model comparisons should use the numerical atomic fractions.

H.5 REPRESENTATIVE GENERATED STRUCTURES

Figures 9 and 10 illustrate FieldForge (PT) candidates that simultaneously satisfy the novelty and uniqueness filters while meeting the stable or metastable energy criterion after relaxation. The selected examples span multiple compositions and space groups, including multicomponent crystals with different atomic arrangements. They provide a qualitative view of the candidates underlying the aggregate SUN and MSUN rates and connect the statistical coverage analyses above to individual generated structures.

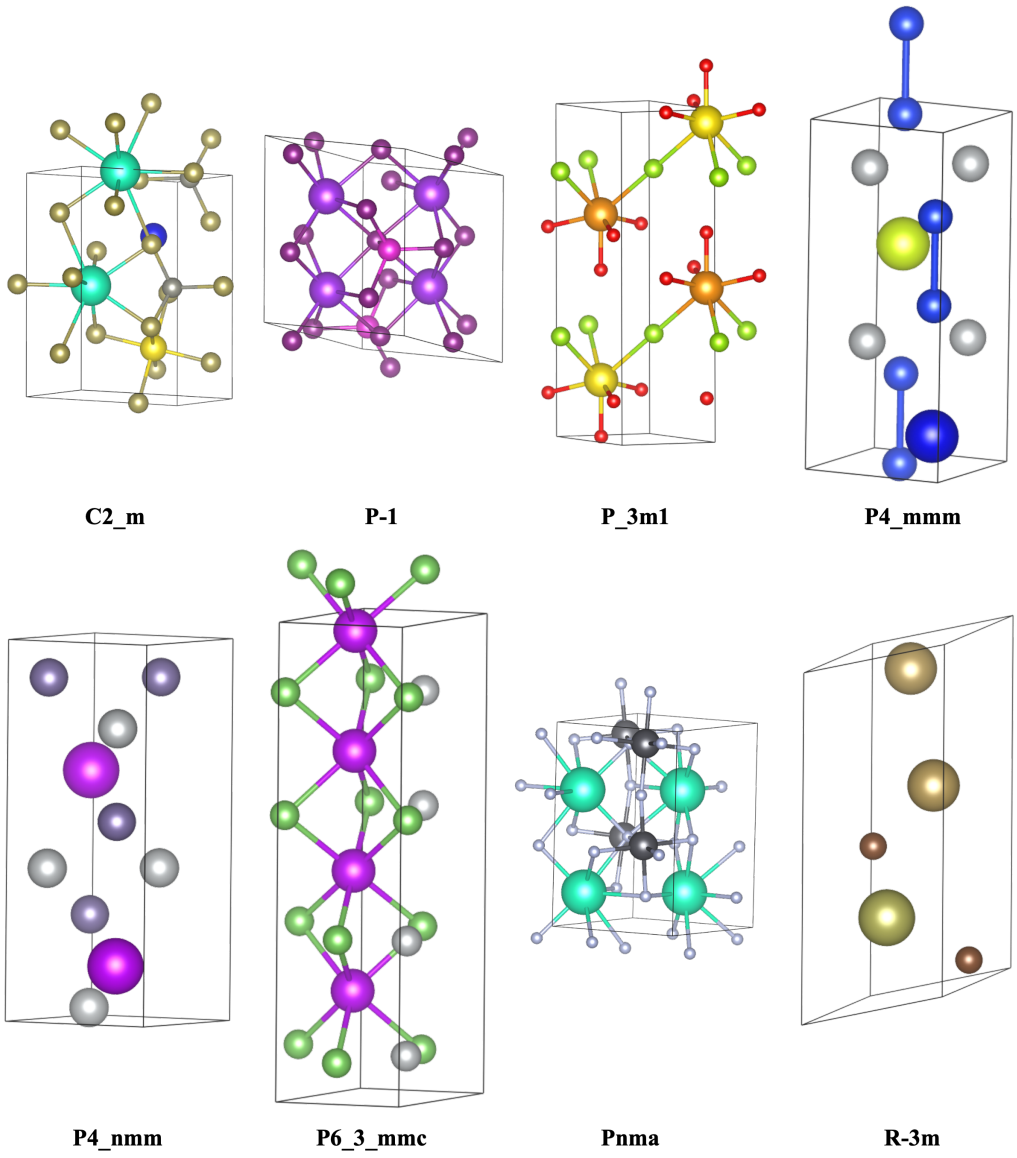


Figure 9: **Representative FieldForge (PT) SUN structures.** Selected generated crystals that are stable, unique, and novel after relaxation.

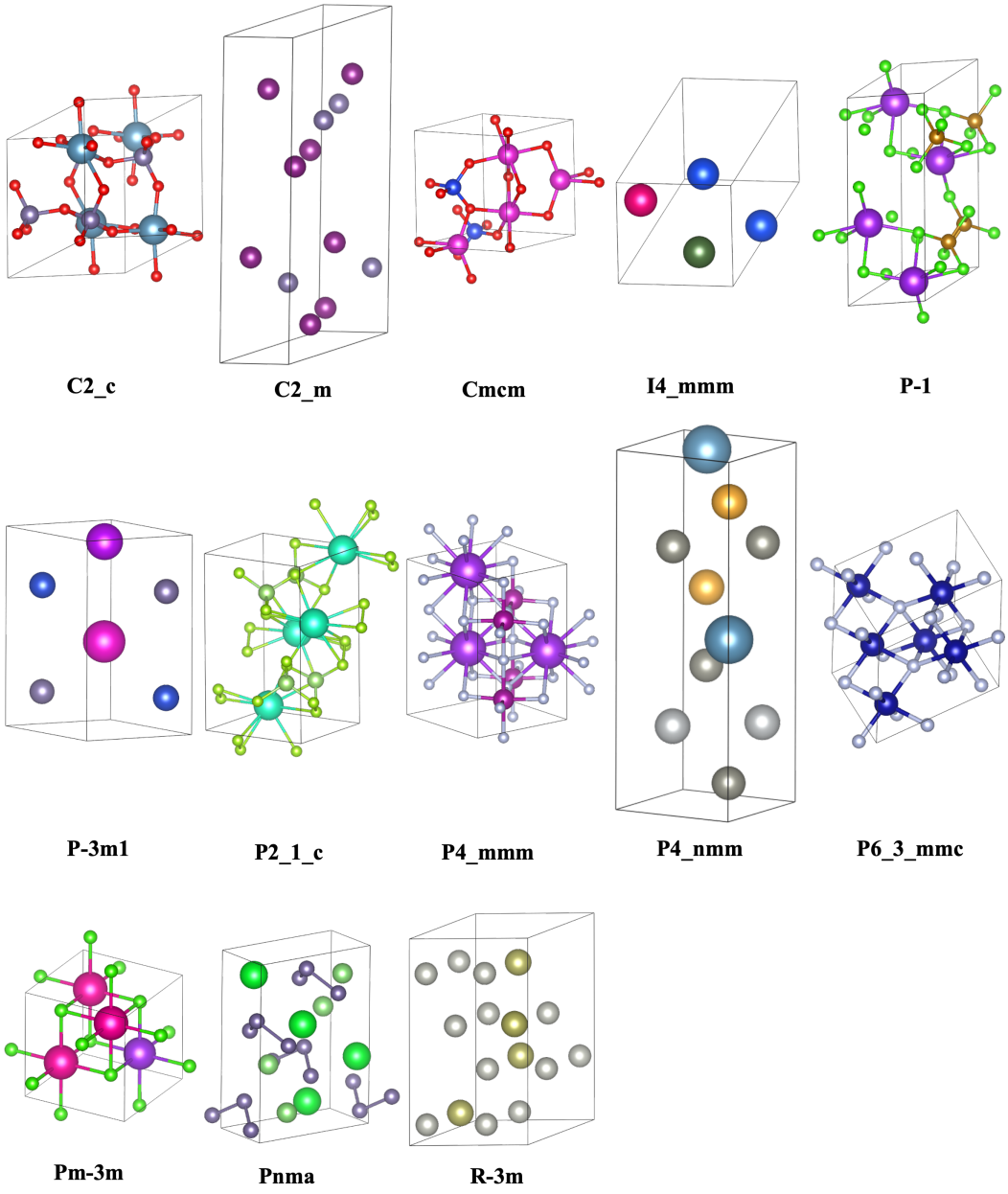


Figure 10: **Representative FieldForge (PT) MSUN structures.** Selected unique and novel crystals with $0 < E_{\text{hull}} \leq 0.1\,\text{eV/atom}$ after relaxation.