CrysTune: Crystal Generation via Fine-Tuning of Large Language Models on Wyckoff Representations

Prashant Govindarajan [1 2 3] **Kishalay Das** [4] **Sarath Chandar** [1 2 3 5]

Abstract

The discovery of novel materials is essential for driving scientific and technological breakthroughs. Recent work has explored fine-tuning large language models (LLMs) for autoregressive crystal generation, but the ideal representation and training strategies for symmetry-based inductive biases remain unclear. We propose CrysTune, a class of LLMs fine-tuned on Wyckoff representations of crystals with two auxiliary tasks: canonicalization and template prediction. CrysTune shows competitive performance and improved stability-related metrics relative to LLMs trained on standard string-encoded representations. We further use these models as initial policies for reinforcement learning (RL) fine-tuning to optimize stability, validity, uniqueness, novelty, and diversity. RL-trained policies produce more valid and metastable crystals, while introducing novelty and diversity trade-offs. We also explore crystal system conditioning, showing that RL-trained policies produce a higher proportion of crystals matching the target condition.

1. Introduction

The search for novel and desirable crystalline materials in the exponentially large 3-dimensional atomic space is extremely challenging. Generative and autoregressive models and the unique representation schemes that have been proposed in the last five years have significantly advanced the state of AI-guided material discovery. The primary goal for the material generation problem is to design a model that can generate valid, stable, unique, novel (S.U.N), and diverse crystals that might also have to satisfy desired properties (e.g., density and band gap). Yet, several challenges remain in terms of searching for candidates outside the distribution of known crystal structures. Given the distinct symmetric properties of crystal structures and the limited availability of materials data, it becomes essential for generative modeling to incorporate the inductive bias due to group symmetries and periodicity. Towards this, multiple works have focused on improving the crystal representation and model architectures to respect the symmetry of crystals (Levy et al., 2025; Cao et al., 2025).

Parameter-efficient fine-tuning of large language models (LLMs) offers a promising alternative for materials generation, reducing the need for specialized tokenizers, custom architectures, and prolonged training. With well-designed prompts and string-based crystal representations, crystal generation can be formulated as autoregressive next-token prediction, analogous to natural language generation. Crystal-Text-LLM (Gruver et al., 2024) fine-tunes LLaMA-2 (Touvron et al., 2023) models on prompts describing atom types, coordinates, lattice parameters, and associated properties, supporting both unconditional and property-conditional generation. However, it does not explicitly incorporate structural inductive biases such as space-group, site-symmetry, or crystal-symmetry information into the LLM training pipeline.

CrystalFormer (Cao et al., 2025) introduces a compressed linear crystal representation that encodes space groups and Wyckoff positions, including multiplicities and letter codes, to train a custom transformer. Its use of alphabetically ordered Wyckoff letters imposes a natural sequential structure, but the custom architecture does not leverage pre-trained LLM knowledge. Moreover, existing LLM-based approaches are trained primarily with next-token prediction, limiting their ability to capture the symmetry inherent in 3D crystal structures. In this work, we investigate fine-tuning LLMs with Wyckoff-based string representations and specialized strategies for learning crystalline symmetries. Our proposed method, **CrysTune**, makes the following contributions.

1. **LLMs trained on Wyckoff representation**: We fine-tune LLaMA-2 chat models (Touvron et al., 2023) with

[1]Chandar Research Lab [2]Mila-Quebec AI Institute [3]Polytechnique Montreal [4]Yale University [5]Canada CIFAR AI Chair. Correspondence to: Prashant Govindarajan <prashant.govindarajan@mila.quebec>.

Proceedings of the 43rd International Conference on Machine Learning, Seoul, South Korea. PMLR 306, 2026. Copyright 2026 by the author(s).

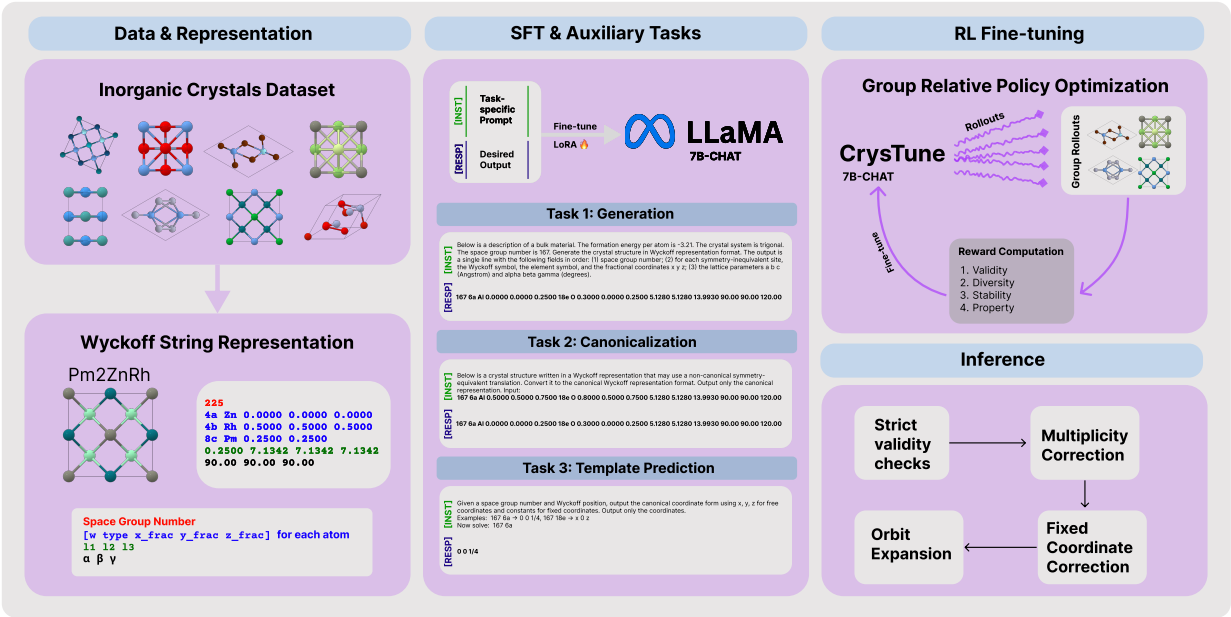

Figure 1. Overview of the proposed CrysTune pipeline, including Wyckoff-based crystal representation, supervised fine-tuning with auxiliary tasks, GRPO-based RL optimization, and post-processing steps for generating valid crystal structures.

a Wyckoff-based crystal representation adopted from CrystalFormer, directly relying on the LLM's tokenizer, similar to Crystal-Text-LLM.

2. **Auxiliary tasks for symmetry learning**: In addition to training on the generative task, we introduce two novel auxiliary tasks 1) *canonicalization*, and 2) *template prediction*, which help the model learn about crystal symmetries.

3. **RL-finetuning**: We further train the fine-tuned model with reinforcement learning with rewards encouraging higher validity, novelty, stability, uniqueness, and diversity, and desired property (e.g. crystal system and density). We show that RL-finetuned models excel in some key metrics while suffering from mode collapse.

We evaluate our baselines and ablations using key metrics for validity, novelty, diversity, and stability from LeMatGenBench (Betala et al., 2025), a recent evaluation framework for crystal generative models. Comparing SFT-trained models, we show that the model trained with both auxiliary tasks – *canonicalization* and *template prediction*— outperforms counterparts trained only on generation or on generation with *canonicalization*. To further align the SFT-trained models with reinforcement learning, we adopt the reward function from (Park & Walsh, 2025), which maps each generated crystal to a scalar reward by combining stability, validity, creativity (uniqueness and novelty), and diversity metrics. We find that RL fine-tuning for a moderate number of steps substantially improves validity, stability, and diversity. However, because RL naturally shifts the policy toward high-reward regions that likely resemble known crystals, we observe a major drop in novelty. Finally, we perform property-conditioned RL training and demonstrate that the RL policy generates a higher proportion of crystals with the desired crystal system. While important challenges remain, CrysTune represents a significant step toward symmetry-aware LLM fine-tuning for crystalline materials discovery.

2. Background

Related Work Traditional approaches to crystal generation rely on Density Functional Theory (DFT) simulations (Kohn & Sham, 1965), which, while accurate for electronic and structural properties, are too costly for large-scale exploration. Lately, data-driven approaches have gained considerable traction in crystal materials research, particularly in crystal generation and design (Xie et al., 2021; Luo et al., 2023; Jiao et al., 2023; Zeni et al., 2023) and property prediction (Xie & Grossman, 2018; Chen et al., 2019; Choudhary & DeCost, 2021; Das et al., 2023a; 2022; Yan et al., 2022; Das et al., 2023b; Mukherjee et al., 2026). Following CDVAE (Xie et al., 2021) and SyMat (Luo et al., 2023), equivariant diffusion models have become a popular paradigm for crystal generation. DiffCSP (Jiao et al., 2023), MatterGen (Zeni et al., 2023), and UniMat (Yang et al., 2023) jointly denoise continuous and discrete crystal variables. Text-conditioned models such as TGDMat (Das et al., 2025), CrysLLMGen (Khastagir et al., 2026) and Chemeleon (Park et al., 2025) enable generation from natural language descriptions, while DiffCSP++ (Jiao et al., 2024) and SymmCD (Levy et al., 2025) incorporate space-

group symmetry through Wyckoff positions. Flow matching (Lipman et al., 2023; Albergo & Vanden-Eijnden, 2023; Liu et al., 2023) has also emerged as a compelling alternative for crystal generation (Miller et al., 2024; Sriram et al., 2024). Customized transformer models such as CrystalFormer and WyFormer (Cao et al.; Kazeev et al., 2025), pre-trained on Wyckoff-based crystal representations, further demonstrate the importance of symmetry-aware generation. Recently, large language models (LLMs) have gained traction in the natural sciences, including materials generation (Van et al., 2025). CrystalLLM (Antunes et al., 2024) and Crystal-Text-LLM (Gruver et al., 2024) train token-based models on textual descriptions of unit cells and atomic arrangements, demonstrating promising ability to generate structurally valid and stable materials. Recent works highlight reinforcement learning as a promising strategy for generating crystals with improved properties (Govindarajan et al., 2025; 2024; Park & Walsh, 2025; Mohanty et al., 2026; Xu et al., 2025). CrysTune combines LLM fine-tuning with improved crystal representations, drawing inspiration from Crystal-Text-LLM (Gruver et al., 2024) and CrystalFormer (Cao et al., 2025). We also acknowledge PLaID++ (Xu et al., 2025), which fine-tunes LLMs on Wyckoff representations, while highlighting that our work differs significantly in its supervised and RL fine-tuning strategies.

Space Groups and Wyckoff Positions A crystal $\mathcal{C}$ with N atoms in its unit cell is characterized by lattice parameters (l_1, l_2, l_3) and angles (α, β, γ). Atomic species are denoted by $\mathbf{A} = \{a_i\}_{i=1}^N$, and positions by fractional coordinates $\mathbf{X} = \{\mathbf{x}_i; \mathbf{x}_i \in [0,1)^3\}_{i=1}^N$, where $\mathbf{x}_i = (x_i, y_i, z_i)$. The unit cell repeats along the three lattice vectors, so each atom has periodic images generated by integer cell translations. In three dimensions, crystals are classified into seven crystal systems: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic, based on their symmetries and shapes. They are further categorized into space groups G, indexed by $K \in \{1, \ldots, 230\}$, which describe the full set of symmetry operations compatible with the periodic lattice. Each element of G can be written as $g = (\mathbf{R}, \mathbf{t})$, where $\mathbf{R}$ is a lattice-preserving linear symmetry operation and $\mathbf{t}$ is a translation vector. It acts on a fractional coordinate $\mathbf{r}$ as $g \cdot \mathbf{r} = \mathbf{R}\mathbf{r} + \mathbf{t} \pmod{\mathbf{1}}$. Operations with $\mathbf{R} = \mathbf{I}$ are pure translations, while nontrivial $\mathbf{R}$ includes rotations, reflections, inversions, rotoinversions, screw rotations, and glide reflections.

Within each space group, there exist symmetry-related "slots" in the unit cell, formally orbits $\mathcal{O}_G(\mathbf{r}) = \{g \cdot \mathbf{r} : g \in G\}$. These slots are grouped by symmetry and labeled as Wyckoff positions. Each label, such as $1a$, $2b$, or $4c$, consists of a multiplicity m, indicating the number of equivalent copies per unit cell, and a letter ℓ identifying the slot type, assigned alphabetically from the most symmetric slot "a" to the least symmetric general position. Higher-symmetry slots produce fewer copies and impose coordinate constraints, whereas lower-symmetry slots allow more free coordinates and generate more copies. Given a space group, e.g., 174, and a Wyckoff position, e.g., $3j$, all copies of an atom can be generated, as shown in Table 6 and analogously for rhombohedral space group 167 in Table 7. Each Wyckoff position has a coordinate template $\tau(K, \ell)$ involving free parameters $\{x, y, z\}$ and rational constants, e.g., $\tau(174, 3j) = (x, y, 0)$ and $\tau(167, 12c) = (0, 0, z)$. Thus, a single representative per orbit suffices to reconstruct the remaining $m - 1$ elements by applying the operations of G. Consequently, Wyckoff-based representations are compact and expose a natural sequential structure, with atoms ordered from highest to lowest site symmetry.

LLM Fine-tuning Supervised fine-tuning (SFT) is commonly used to adapt pretrained LLMs to downstream tasks by continuing training on task-specific input–output examples. Given a prompt-response pair, SFT optimizes the standard causal language modeling objective over the response tokens, encouraging the model to imitate the target outputs. While SFT provides a strong task-adapted initialization, further alignment can be achieved through reinforcement learning. Group relative policy optimization (Shao et al., 2024) is a policy gradient algorithm for aligning LLMs without requiring a separate value-function critic. Instead of estimating an absolute baseline, GRPO computes the normalized advantage of each generated output relative to a group of outputs sampled from the same prompt. These group-relative advantages are then optimized using a PPO-style clipped surrogate objective, where the probability ratio $r_i(\theta)$ is clipped to $[1-\epsilon, 1+\epsilon]$ to constrain policy updates and improve training stability.

$$A_i = \frac{R(S_i) - \frac{1}{K}\sum_{j=1}^{K} R(S_j)}{\sigma_R} \tag{1}$$

$$J(\theta) = \mathbb{E}\left[\sum_{i=1}^{K} \min\left(r_i(\theta)A_i, \text{clip}(r_i(\theta), 1-\epsilon, 1+\epsilon)A_i\right)\right]. \tag{2}$$

3. Methods

3.1. Wyckoff Representation

To convert crystals from CIF format to a Wyckoff-based representation, we adopt the CrystalFormer pipeline (Cao et al., 2025). However, instead of treating non-numerical fields (e.g., Wyckoff positions) and numerical fields (e.g., fractional coordinates) separately, we combine the full representation into a single string suitable for LLM training. The string begins with the space group number n_G, which

serves as the premise for generating subsequent N_s Wyckoff sites and coordinates. For each distinct atom i, we then include three entities in sequence: 1) the Wyckoff position w_i (e.g., 4a), 2) the chemical symbol a_i (e.g., Cu), and 3) the fractional coordinates $\boldsymbol{x}_i$ (e.g., $0.5, 0.5, 0.5$). Atoms are ordered by site symmetry, from highest symmetry to lowest general position. Finally, we append the lattice parameters: lengths l_1, l_2, l_3 and angles α, β, γ.

$$s = \left[\, n_G \,\|\, \{w_i, a_i, x_i, y_i, z_i\}_{i=1}^{N_s} \,\|\, l_1, l_2, l_3, \alpha, \beta, \gamma \,\right]. \quad (3)$$

A single physical crystal can admit multiple symmetry-equivalent coordinate strings, since each atom may be represented by any position related through the space-group symmetry. We therefore adopt the canonical representation defined by the International Tables for Crystallography (Wondratschek et al., 2004), fixing two conventions: **(i)** the origin at the trivial translation $\mathbf{t} = \mathbf{0}$, e.g., selecting $(0, 0, 0)+$ over $(\frac{2}{3}, \frac{1}{3}, \frac{1}{3})+$ for space group 167, Table 7; and **(ii)** for each Wyckoff position, using only the first listed orbit representative, e.g., $(0, 0, \frac{1}{4})$ for 6a and $(0, 0, z)$ for 12c.

3.2. Supervised Fine-Tuning (SFT)

We train CrysTune on the MP-20 training set, a Materials Project (Jain et al., 2013) subset with at most 20 atoms per crystal and $\sim 27k$ crystals. We adopt the fine-tuning pipeline and hyperparameters from Crystal-Text-LLM (Gruver et al., 2024), and train LLaMA-2 chat models with low-rank adaptation (LoRA) (Hu et al., 2022). We rely entirely on LLaMA-2's tokenizer rather than adding custom tokens. At each step, we sample one of three tasks: 1) ***generation*** (default), 2) ***canonicalization***, and 3) ***template prediction***, with probabilities ($p_{\text{gen}} = 0.65, p_{\text{can}} = 0.25, p_{\text{temp}} = 0.10$). The space-separated representation provides field boundaries for identifying tokenizer spans and applying masks, as described in Section 3.2.1.

3.2.1. Task 1: Generation

The generation prompt $\mathcal{I}_{gen}$ instructs the model to generate a crystal in the canonical Wyckoff representation $\mathcal{C}$ and describes the desired output format. Similar to Crystal-Text-LLM, we randomly sample a subset of properties, or none, to include in the prompt, enabling both conditional and unconditional generation. These properties include the space group number, crystal system, density, chemical formula, energy above hull, and formation energy. During loss computation, we mask tokens corresponding to fixed coordinates; e.g., for position 18e in space group 167, whose template is $(x, 0, 1/4)$, we exclude the tokens for 0 and $1/4$. The loss is computed over tokens corresponding to the space group, Wyckoff position, atom type, lattice parameters, and free coordinates. For each token t_j in the response span, the generation loss is the masked causal language modeling objective.

$$\mathcal{L}_{\text{gen}}(\theta) \;=\; -\mathbb{E}\left[\, \frac{1}{\sum_j \mu_j} \sum_{j=1}^{L} \mu_j \,\log \pi_\theta\big(t_j \,\big|\, t_{<j}, \mathcal{I}_{\text{gen}}\big) \right]; \quad (4)$$

$$\mu_j = \begin{cases} 0, & t_j \text{ is fixed,} \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

Masking the fixed-coordinate tokens prevents the model from wasting capacity memorizing constants it could equivalently recover from the Wyckoff letter, and avoids artificially inflating per-token accuracy on positions that are trivial by construction.

3.2.2. Task 2: Canonicalization

The prompt for the *canonicalization* task, $\mathcal{I}_{can}$ instructs the model to transform a representation s that has undergone a non-trivial group symmetric translation over all the sites back to the canonical representation. The translation vector $\mathbf{v}$ is uniformly sampled from the finite set of non-trivial pure-translation symmetry operations $\mathcal{T}(G_{\mathcal{C}}) = \big\{\, \mathbf{v} \in [0,1)^3 \;:\; (\mathbf{I}, \mathbf{v}) \in G_{\mathcal{C}},\; \mathbf{v} \neq \mathbf{0} \,\big\}$, given the space group $G_{\mathcal{C}}$ of the crystal $\mathcal{C}$. As seen in Table 7, if $n_G = 167$, this set will be $\big\{\; (\frac{2}{3}, \frac{1}{3}, \frac{1}{3}),\; (\frac{1}{3}, \frac{2}{3}, \frac{2}{3}) \big\}$. The resultant translated string representation therefore becomes $\hat{s}_{\mathbf{v}} = \big[\, n_G \,\|\, \{w_i, a_i, (\mathbf{x}_i + \mathbf{v}) \bmod \mathbf{1}\}_{i=1}^{N_s} \,\|\, l_1, l_2, l_3, \alpha, \beta, \gamma \,\big]$, where $\mathbf{v} \sim \mathrm{Unif}_{\text{disc}}\big(\mathcal{T}(G_{\mathcal{C}})\big)$.

Each $\mathbf{v} \in \mathcal{T}(\mathcal{C})$ yields a symmetry-equivalent but generally non-canonical translated Wyckoff serialization. In this task, the model is trained to map this shifted form back to the canonical one: $s_{\mathbf{v}}(\mathcal{C}) \longmapsto s(\mathcal{C})$. The training objective is the causal language-modeling loss over the canonical target,

$$\mathcal{L}_{\text{canon}}(\theta) \;=\; -\mathbb{E}\left[\, \frac{1}{\sum_j \mu_j} \sum_{j=1}^{L} \mu_j \,\log \pi_\theta\big(t_j \,\big|\, t_{<j}, \mathcal{I}_{\text{can}}\big) \right], \quad (6)$$

where we mask fixed coordinates similar to Equation (4). Crystals for which $\mathcal{T}(\mathcal{C}) = \varnothing$ are skipped for this task and re-routed to the generation task. Conceptually, this task acts as an explicit symmetry-consistency objective: it encourages the model to map translated members of the same symmetry orbit onto a common canonical target, rather than leaving this invariance implicit in the data distribution.

3.2.3. Task 3: Template Prediction

Given a space group n_G and letter ℓ, the generator of the corresponding orbit is a deterministic coordinate template $\tau(n_G, \ell)$ expressed in the free parameters $\{x, y, z\}$ and constants – examples: $\tau(167, 6a) = (0,\, 0,\, \frac{1}{4})$ and

$\tau(167, 18e) = (x,\, 0,\, \frac{1}{4})$. Reliable generation requires the model to implicitly know τ for all (n_G, ℓ) combinations it encounters, but this knowledge is only supervised indirectly through the coordinate values that appear in training crystals. We therefore introduce a *template prediction* task that supervises τ explicitly. Given a prompt $p_{\text{tmpl}} = [\,\mathcal{I}_{\text{tmpl}} \,\|\, n_G \,\|\, w\,]$, which contains a fixed instruction, a space group number, and a Wyckoff symbol, the model must emit the three-component template string $\tau(n_G, \ell)$ with probability $\pi_\theta(\tau(n_G, \ell) \mid p_{\text{tmpl}})$. We then train with the template loss

$$\mathcal{L}_{\text{tmpl}}(\theta) = -\mathbb{E}\left[\sum_{j:\, t_j \in \tau(n_G,\ell)} \log \pi_\theta(t_j \mid t_{<j},\, p_{\text{tmpl}})\right]. \tag{7}$$

At training time, (n_G, ℓ) pairs are sampled from the Wyckoff sites present in the current crystal.

3.3. Inference

At inference time, the fine-tuned model π_θ samples a candidate string $\hat{s} \sim \pi_\theta(\cdot \mid p)$ from an unconditional or property-conditioned prompt p (Section 3.2.1). $\hat{s}$ is converted to a CIF file through a parser $\mathcal{P}(\hat{s})$, which treats it as a noisy proposal and projects it onto a structure consistent with the declared space group. Once $(\hat{n}_G, \hat{\ell}_i)$ are known, the site multiplicity, fixed coordinate axes, and generator orbit are deterministic; mismatches are handled through projection-based repair, while violations of the prior are treated as failures.

Hard validity checks Let $\hat{s} = \big[\,\hat{n}_G \,\|\, \{\hat{w}_i, \hat{a}_i, \hat{x}_i, \hat{y}_i, \hat{z}_i\}_{i=1}^{\hat{N}_s} \,\|\, \hat{l}_1, \hat{l}_2, \hat{l}_3, \hat{\alpha}, \hat{\beta}, \hat{\gamma}\,\big]$. The parser discards the full string if any global check fails: the sequence contains fewer than eight tokens; the site block cannot be partitioned into complete five-token site records; $\hat{n}_G \notin \{1, \ldots, 230\}$; or any coordinate $\hat{x}_i, \hat{y}_i, \hat{z}_i$ or lattice parameter $\hat{\boldsymbol{\ell}}$ cannot be parsed as a real number; or if the element symbol is unrecognized or satisfies $Z(\hat{a}_i) > 94$. For site-level checks, the parser drops only the offending Wyckoff site rather than the whole string. This occurs when $\hat{w}_i$ does not match the pattern "`<integer><letter>`" or its letter is outside the valid Wyckoff-label set $\{\mathrm{a}, \ldots, \mathrm{z}, \mathrm{A}\}$.

Symmetry-aware corrections The generation and auxiliary tasks (Sections 3.2.2 and 3.2.3) are designed to teach the model the Wyckoff tables: ideally, each rollout should produce multiplicities consistent with the declared space group and place symmetry-pinned coordinates at their required values. Supervised fine-tuning cannot guarantee this, since Wyckoff tables are discrete and combinatorial, and no smooth loss explicitly prevents incorrect multiplicities or slightly drifted pinned coordinates. Rather than discard such outputs, whose space group, elements, free coordinates, and lattice parameters may remain sensible, we project each string onto the nearest valid Wyckoff representation using the same tables used during training. The parser applies three correction steps, each targeting a predictable error class: **C1: Multiplicity correction**, where the multiplicity is recomputed from the declared space group and Wyckoff letter, i.e., $m_i = \mathrm{Mult}(\hat{n}_G, \hat{\ell}_i)$, so an invalid label such as "$6a$" in space group 229 is read as "$2a$" when the a site has multiplicity 2; **C2: Fixed-coordinate snapping**, where coordinates pinned by the Wyckoff template $\tau(\hat{n}_G, \hat{\ell}_i)$ are overwritten with their required constants, such as the pinned entries in $\tau(167, 6a) = (0, 0, \frac{1}{4})$ or $\tau(167, 18e) = (x, 0, \frac{1}{4})$, while free axes are left unchanged; and **C3: Orbit expansion**, where the corrected representative coordinate is expanded into all m_i symmetry-equivalent atoms before writing the CIF. This procedure preserves useful generated content while repairing residual grammatical violations.

Caveats While our parsing strategies and symmetry-aware corrections are design choices, they introduce some limitations. Because SFT encourages the model to learn canonical representations, during inference, **C2** enforces a canonical form even when generated coordinates could be interpreted through another template in the same orbit. For example, for space group 167, "`18e Cu 0.0 0.30 0.25`" is corrected to "`18e Cu 0.0 0.0 0.25`" under the canonical template $(x, 0, 1/4)$, although it matches the second orbit template $(0, x, 1/4)$. This may also introduce overlaps with other orbits. In Section E.3, we rerun inference with alternative parsing strategies, including space group correctness and flexible fixed-coordinate snapping, but these did not substantially improve key evaluation metrics.

3.4. RL Fine-Tuning with GRPO

The resulting SFT model is used as the initial policy for subsequent RL fine-tuning with GRPO for unconditional and conditional crystal generation. For unconditional generation, the prompt contains the simplest instruction from the SFT generation task, with no properties mentioned; hence, the input is the same throughout all episodes during training. For conditional generation, the RL dataset includes prompts corresponding to different target values. In our experiments, we mainly choose *crystal system* as the target property, and the RL fine-tuning dataset consists of seven prompts corresponding to the seven crystal systems. For each prompt p, the actor samples a group of $K = 5$ rollouts $\{\hat{s}^{(k)}\}_{k=1}^{K} \sim \pi_\theta(\cdot \mid p)$ with temperature $\tau = 1.0$. Each string is decoded through the parser $\mathcal{P}$ (Section 3.3) to a pymatgen structure $\mathcal{C}^{(k)}$, which is then scored by the reward function $R(\mathcal{C}^{(k)})$. The actor is trained with the clipped surrogate objective shown in Equation (2), together with a KL divergence $D_{\mathrm{KL}}^{\mathrm{lv}}$ penalty that penalizes large deviations

from the initial SFT policy.

Reward function We adopt the modular crystal-generation reward proposed in Chemeleon2 (Park & Walsh, 2025), which combines verifiable terms for thermodynamic stability, novelty/uniqueness, and distributional diversity: $R(\mathcal{C}) = \sum_{c \in \mathcal{R}} \omega_c \, \tilde{r}_c(\mathcal{C})$, where $\tilde{r}_c$ denotes the c-th component after its configured batch-level normalization. For unconditional RL fine-tuning, the individual reward terms correspond to the energy above hull (stability), creativity, composition diversity, and structure diversity. Energies are estimated with MACE-MP (Batatia et al., 2025) and compared against a pre-built Materials Project `PatchedPhaseDiagram`; creativity is based on whether a rollout is both unique within the batch and novel relative to the MP-20 reference set; and the two diversity terms compare generated and reference distributions in Chemeleon2 VAE latent spaces. For property-conditioned RL fine-tuning, the same Chemeleon2 reward can be augmented with a target-matching term. In our crystal system conditional run, each prompt specifies one of the seven crystal systems, and we add a binary reward for matching the target crystal system computed from `SpacegroupAnalyzer`. The active components, weights, and normalizations used in our runs are summarized in Table 4; explicit component definitions are provided in Section B.

4. Results

In this section, we perform ablations and comparisons with appropriate baselines to answer two major questions:

- **Q1:** *Does fine-tuning on Wyckoff representations and the inclusion of auxiliary tasks during SFT improve generative performance?*
- **Q2:** *Does reinforcement learning (RL) fine-tuning improve key metrics and enhance property-conditioned generation?*

To investigate **Q1**, we compare **CrysTune^{SFT}**, supervised fine-tuning on Wyckoff representations with two auxiliary tasks (Section 3.2), with **Crystal-Text-LLM (CrysTxtLLM^{SFT})** (Gruver et al., 2024). We also ablate CrysTune models trained only on *generation*, **CrysTune^{gen}**, and on *generation+canonicalization*, **CrysTune$^{gen+can}$**. The base LLM is LLaMA-2 7B HF. All models were trained for 4 epochs, since preliminary experiments showed a drop with longer training (Section E.1). For each model, we sample 2,500 structures with temperature 0.7 and top-$p = 1.0$. We discard samples that fail hard validity checks and resample until obtaining 2,500 CIF-exportable structures. For conditional generation, we sample 2,500 structures per property prompt. We follow LeMat-GenBench (Betala et al., 2025) and use Orb-v3 (Rhodes et al., 2025) for relaxation and stability analysis.

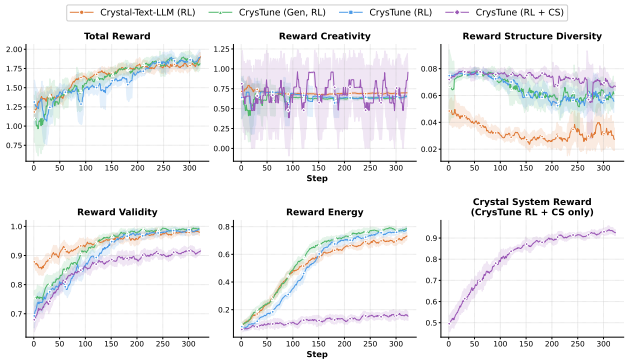

Figure 2. RL fine-tuning curves for each reward component. Energy and validity improve, while diversity and creativity show limited or inconsistent gains, indicating a trade-off between stability-related objectives and exploratory generation. In the crystal system conditioned run, the crystal system reward increases steadily.

4.1. SFT on Wyckoff Representations

In Table 1, CrysTune^{SFT} improves energy and stability metrics over CrysTxtLLM^{SFT}, while maintaining competitive diversity metrics. However, we observe decreases in validity, uniqueness, and novelty, with novelty exhibiting the most pronounced drop. We further note that the two methods differ in their parsing procedures: CrysTxtLLM uses a simple string-parsing process to directly convert generated text representations into CIF files, whereas CrysTune employs a multi-stage parsing pipeline with correction steps, as described in Section 3.3. This added complexity may allow some invalid and less novel structures to propagate into the evaluation pipeline. Furthermore, we present an ablation study over CrysTune variants with and without auxiliary tasks in Table 10. The full CrysTune^{SFT} model, which jointly incorporates both auxiliary tasks, achieves substantial improvements in validity, uniqueness, and energy- and stability-based metrics compared to CrysTune^{gen} and CrysTune$^{\text{gen+can}}$, which do not include the complete auxiliary-task setup. However, we do not observe a clear positive trend in novelty and diversity metrics with the addition of the auxiliary tasks.

4.2. GRPO Fine-tuning

Unconditional Generation Addressing **Q2**, we note that additional RL fine-tuning with GRPO in **CrysTune^{RL}** drives substantial gains in the validity (98.36%), while demonstrating improved metastability, energy and relaxation metrics compared to CrysTune^{SFT} (Table 1). We further show that our RL-trained model outperforms CrysTxtLLM^{RL}, obtained by applying the same GRPO fine-tuning pipeline to CrysTxtLLM^{SFT} – on all metrics except SUN, under-

Table 1. Comparison of CrysTxtLLM and CrysTune under SFT and RL fine-tuning. Best result per metric within each training setting is **bolded**. ↑/↓ denotes higher/lower is better.

	SFT		*RL Fine-tuning*	
Metric	**CrysTxtLLM^{SFT}**	**CrysTune^{SFT}**	**CrysTxtLLM^{RL}**	**CrysTune^{RL}**
Generation Quality				
Validity (%) ↑	**61.38**	52.21	**98.38**	98.36
Uniqueness (%) ↑	**98.78**	94.65	54.15	**83.78**
Novelty (%) ↑	**54.07**	29.40	16.58	**17.16**
Stability				
SUN (%) ↑	0.0	**0.223**	**0.041**	0.0
MSUN (%) ↑	0.360	**0.445**	0.287	**1.100**
Stability (%) ↑	0.144	**1.336**	0.041	**0.285**
Metastability (%) ↑	6.62	**30.29**	17.08	**42.62**
Energy/Relaxation				
E_{hull} (eV/atom) ↓	1.202	**1.049**	0.271	**0.167**
E_{form} (eV/atom) ↓	**0.328**	0.511	-0.452	**-0.865**
Relax. RMSE ↓	1.219	**0.500**	0.356	**0.147**
Diversity				
Element ↑	**0.712**	**0.712**	0.619	**0.678**
Space group ↑	**0.183**	0.130	**0.039**	0.017
Site number ↑	5.62	**6.40**	1.78	**2.06**
Physical size ↑	0.232	**0.314**	0.126	**0.215**

scoring the effectiveness of Wyckoff-based representations. However, RL optimization introduces clear mode collapse: novelty falls to 17.2%, space group diversity collapses from 0.183 to 0.017, and consequently, the SUN score drops to zero. While CrysTune^{RL} successfully steers generation toward low-energy metastable configurations, the lower stability indicates that the model likely overfits to the preference of MACE, the energy prediction model used in Chemeleon2's reward function we adopt – LeMatGenBench's evaluation follows Orb-v3 (Rhodes et al., 2025), a more recent and advanced interatomic potential.

Crystal system-conditioned generation To evaluate target-specific generation, we compare CrysTune^{SFT} and RL-tuned models with (CrysTune$^{RL}{}_{CS}$) and without (CrysTune^{RL}) the property-matching reward across all seven crystal systems, using identical prompts. Figure 3 shows the generated crystal system counts. The poor performance of the SFT baseline indicates weak influence of the conditional prompt. The unconditional RL model mostly improves over SFT, and the conditional RL model performs better in almost all cases. Gains are largest for rare systems such as trigonal, hexagonal, and tetragonal, where conditioned RL yields nearly an order-of-magnitude increase in on-target generation. Table 5 shows system-dependent trade-offs: cubic targets have the highest validity and strongest energy/relaxation metrics but lower novelty, whereas triclinic targets show the opposite pattern. Overall, RL fine-tuning improves controllability. Density-conditioned RL with a distance-based reward gives modest improvements, but Section E.2 shows that continuous property optimization remains difficult.

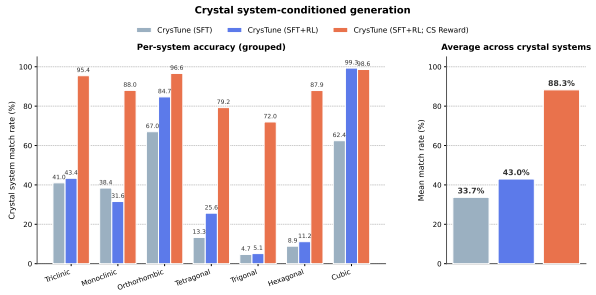

Figure 3. Generated crystal system counts across different targets. The model fine-tuned with RL and a property-matching reward generates more structures that satisfy the desired crystal system.

5. Discussion and Conclusion

Symmetry-aware LLMs offer a promising direction for automated materials discovery workflows. In this work, we introduced CrysTune, an approach for incorporating symmetry-based inductive biases into LLM-based crystal generation through auxiliary tasks during supervised fine-tuning. By training on Wyckoff-based representations, CrysTune improves the model's ability to capture crystallographic structure and leads to gains in several metrics of generated materials. We further show that reinforcement learning can be used to align the model toward desired objectives, including improved stability-related metrics and property-conditioned

generation.

At the same time, several challenges remain. First, stability and novelty scores remain limited across the models considered, suggesting that further study is needed to understand when and how symmetry-aware training can improve the generation of both stable and novel crystals. Second, a more systematic comparison with customized transformer architectures would help clarify the relative advantages of fine-tuning pretrained LLMs versus training domain-specific models from scratch. Finally, RL fine-tuning introduces an important trade-off: although it can shift the policy toward lower-energy and metastable configurations, it may also reduce novelty and diversity, indicating potential mode collapse. Overall, CrysTune highlights the potential of combining LLM fine-tuning, crystallographic representations, and reinforcement learning, and we hope it motivates further work on symmetry-aware language models for materials discovery.

References

Albergo, M. S. and Vanden-Eijnden, E. Building normalizing flows with stochastic interpolants. *ICLR*, 2023.

Antunes, L. M., Butler, K. T., and Grau-Crespo, R. Crystal structure generation with autoregressive large language modeling. *Nature Communications*, 15(1):10570, 2024.

Batatia, I., Benner, P., Chiang, Y., Elena, A. M., Kovács, D. P., Riebesell, J., Advincula, X. R., Asta, M., Avaylon, M., Baldwin, W. J., et al. A foundation model for atomistic materials chemistry. *The Journal of chemical physics*, 163(18), 2025.

Betala, S., Gleason, S. P., Ramlaoui, A., Xu, A., Channing, G., Levy, D., Fourrier, C., Kazeev, N., Joshi, C. K., Kaba, S.-O., et al. Lemat-genbench: A unified evaluation framework for crystal generative models. *arXiv preprint arXiv:2512.04562*, 2025.

Cao, Z., Luo, X., Lv, J., and Wang, L. Space group informed transformer for crystalline materials generation, 2024. *URL https://arxiv. org/abs/2403.15734*, 11.

Cao, Z., Luo, X., Lv, J., and Wang, L. Space group informed transformer for crystalline materials generation. *Science Bulletin*, 2025.

Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.*, 31(9):3564–3572, 2019.

Choudhary, K. and DeCost, B. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):1–8, 2021.

Das, K., Samanta, B., Goyal, P., et al. CrysXPP: An explainable property predictor for crystalline materials. *npj Computational Materials*, 8(43), 2022. doi: 10.1038/s41524-022-00716-8. URL https://doi.org/10.1038/s41524-022-00716-8.

Das, K., Goyal, P., Lee, S.-C., Bhattacharjee, S., and Ganguly, N. Crysmmnet: multimodal representation for crystal property prediction. In *Uncertainty in Artificial Intelligence*, pp. 507–517. PMLR, 2023a.

Das, K., Samanta, B., Goyal, P., Lee, S.-C., Bhattacharjee, S., and Ganguly, N. Crysgnn: Distilling pre-trained knowledge to enhance property prediction for crystalline materials. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 7323–7331, 2023b.

Das, K., Khastagir, S., Goyal, P., Lee, S.-C., Bhattacharjee, S., and Ganguly, N. Periodic materials generation using text-guided joint diffusion model. In *ICLR*, 2025.

Govindarajan, P., Miret, S., Rector-Brooks, J., Phielipp, M., Rajendran, J., and Chandar, S. Learning conditional policies for crystal design using offline reinforcement learning. *Digital Discovery*, 3:769–785, 2024. doi: 10.1039/D4DD00024B. URL http://dx.doi.org/10.1039/D4DD00024B.

Govindarajan, P., Reymond, M., Clavaud, A., Phielipp, M., Miret, S., and Chandar, S. Crystalgym: A new benchmark for materials discovery using reinforcement learning. *arXiv preprint arXiv:2509.23156*, 2025.

Gruver, N., Sriram, A., Madotto, A., Wilson, A. G., Zitnick, C. L., and Ulissi, Z. W. Fine-tuned language models generate stable inorganic materials as text. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=vN9fpfqoP1.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al. Lora: Low-rank adaptation of large language models. *Iclr*, 1(2):3, 2022.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.

Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., and Liu, Y. Crystal structure prediction by joint equivariant diffusion. *arXiv preprint arXiv:2309.04475*, 2023.

Jiao, R., Huang, W., Liu, Y., Zhao, D., and Liu, Y. Space group constrained crystal generation. *arXiv preprint arXiv:2402.03992*, 2024.

Kazeev, N., Nong, W., Romanov, I., Zhu, R., Ustyuzhanin, A., Yamazaki, S., and Hippalgaonkar, K. Wyckoff transformer: Generation of symmetric crystals. *arXiv preprint arXiv:2503.02407*, 2025.

Khastagir, S., Das, K., Goyal, P., Lee, S.-C., Bhattacharjee, S., and Ganguly, N. LLM Meets Diffusion: A Hybrid Framework for Crystal Material Generation. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026. URL https://openreview.net/forum?id=E6gwPtWjb1.

Kohn, W. and Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

Kwon, W. *vLLM: An Efficient Inference Engine for Large Language Models*. PhD thesis, UC Berkeley, 2025.

Levy, D., Panigrahi, S. S., Kaba, S.-O., Zhu, Q., Lee, K. L. K., Galkin, M., Miret, S., and Ravanbakhsh, S. SymmCD: Symmetry-preserving crystal generation with diffusion models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=xnssGv9rpW.

Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M. Flow matching for generative modeling. *ICLR*, 2023.

Liu, X., Gong, C., and Liu, Q. Flow straight and fast: Learning to generate and transfer data with rectified flow. *ICLR*, 2023.

Luo, Y., Liu, C., and Ji, S. Towards symmetry-aware generation of periodic materials. *Advances in Neural Information Processing Systems*, 2023.

Miller, B. K., Chen, R. T., Sriram, A., and Wood, B. M. Flowmm: Generating materials with riemannian flow matching. *arXiv preprint arXiv:2406.04713*, 2024.

Mohanty, T., Mehta, M., Sayeed, H. M., Oded, B.-E., Pitussi, I., Borenstein, A., Srikumar, V., and Sparks, T. D. Crystext: a generative ai approach for text-conditioned crystal structure generation using llm. *Integrating Materials and Manufacturing Innovation*, pp. 1–15, 2026.

Mukherjee, S., Das, K., Basuchowdhuri, P., Goyal, P., and Ganguly, N. Latent diffusion pretraining for crystal property prediction. *arXiv preprint arXiv:2606.00776*, 2026.

Park, H. and Walsh, A. Guiding generative models to uncover diverse and novel crystals via reinforcement learning. *arXiv preprint arXiv:2511.07158*, 2025.

Park, H., Onwuli, A., and Walsh, A. Exploration of crystal chemical space using text-guided generative artificial intelligence. *Nature Communications*, 16(1):4379, 2025.

Rhodes, B., Vandenhaute, S., Šimkus, V., Gin, J., Godwin, J., Duignan, T., and Neumann, M. Orb-v3: atomistic simulation at scale. *arXiv preprint arXiv:2504.06231*, 2025.

Shao, Z., Wang, P., Zhu, Q., Xu, R., Song, J., Bi, X., Zhang, H., Zhang, M., Li, Y., Wu, Y., et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

Sheng, G., Zhang, C., Ye, Z., Wu, X., Zhang, W., Zhang, R., Peng, Y., Lin, H., and Wu, C. Hybridflow: A flexible and efficient rlhf framework. *arXiv preprint arXiv: 2409.19256*, 2024.

Sriram, A., Miller, B., Chen, R. T., and Wood, B. Flowllm: Flow matching for material generation with large language models as base distributions. *NeurIPS*, 37:46025–46046, 2024.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.

Van, M.-H., Verma, P., Zhao, C., and Wu, X. A survey of ai for materials science: foundation models, llm agents, datasets, and tools. *arXiv preprint arXiv:2506.20743*, 2025.

Wondratschek, H., Müller, U., and internationale de cristallographie, U. *International tables for crystallography*, volume 1. Wiley Online Library, 2004.

Xie, T. and Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.*, 120(14):145301, 2018.

Xie, T., Fu, X., Ganea, O.-E., Barzilay, R., and Jaakkola, T. Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197*, 2021.

Xu, A., Desai, R., Wang, L., Hope, G., and Ritz, E. Plaid++: A preference aligned language model for targeted inorganic materials design. *arXiv preprint arXiv:2509.07150*, 2025.

Yan, K., Liu, Y., Lin, Y., and Ji, S. Periodic graph transformers for crystal material property prediction. *Advances in Neural Information Processing Systems*, 35:15066–15080, 2022.

Yang, M., Cho, K., Merchant, A., Abbeel, P., Schuurmans, D., Mordatch, I., and Cubuk, E. D. Scalable diffusion for materials generation. *arXiv preprint arXiv:2311.09235*, 2023.

Zeni, C., Pinsler, R., Zügner, D., Fowler, A., Horton, M., Fu, X., Shysheya, S., Crabbé, J., Sun, L., Smith, J., et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.

A. Hyperparameters

A.1. SFT Hyperparameters

We supervised fine-tune the LLaMA-2-7B-chat model on the MP-20 Wyckoff representation dataset using parameter-efficient LoRA adapters. Training is distributed over four H100 GPUs with one process per GPU. The SFT mixture contains the main property-conditioned Wyckoff generation objective together with auxiliary canonicalization and template prediction tasks. Symmetry-fixed Wyckoff coordinates are kept in the target sequence as context, but their loss contribution is masked during training.

Table 2. SFT Hyperparameters.

Hyperparameter	Value
Base model	`meta-llama/Llama-2-7b-chat-hf`
Training data	`data/mp_20/train.csv`
Validation data	`data/mp_20/val.csv`
Number of epochs	4
GPUs	4 H100 GPUs
Per-device batch size	1
Gradient accumulation steps	1
Effective global batch size	4
Learning rate	1×10^{-4}
Learning-rate scheduler	Cosine
Warmup steps	100
Weight decay	0.0
Maximum sequence length	2048 tokens
LoRA rank	8
LoRA alpha	32
LoRA dropout	0.05
LoRA bias	None
Generation task probability	0.65
Canonicalization task probability	0.25
Template prediction task probability	0.10
Property conditioning	Enabled
Mask fixed Wyckoff coordinate loss	Enabled
Optimizer	AdamW

A.2. GRPO Hyperparameters

We perform GRPO fine-tuning using VERL (Sheng et al., 2024) on $4\times$H100 GPUs with FSDP sharding and bf16 weights; vLLM (Kwon, 2025) handles on-policy rollout generation with tensor parallelism. The PPD, MACE-MP calculator, and VAE featurizer are instantiated once per worker and reused across steps; a batch-level timeout of $180\,\mathrm{s}$ on MACE and $120\,\mathrm{s}$ on the VAE forward falls back to NaN (mapped to $R_{\min}$). Table 3 provides the values of the important hyperparameters of the RL fine-tuning pipeline.

Table 3. Shared RL hyperparameters for the two GRPO runs (unconditional and crystal system conditional).

Setting	Value
Actor backbone	LLaMA-2-7B-chat-hf, SFT checkpoint
Algorithm	GRPO
Prompts per step (batch)	64
Rollouts per prompt K	5
Mini-batch size	16
Micro-batch per GPU	4
Max prompt / response tokens	512 / 500
Sampling temperature	1.0
Actor learning rate	1×10^{-6} (AdamW, FSDP)
KL coefficient β	10^{-3} (low-variance KL)
KL placement	actor loss only (not in reward)
Entropy bonus	0
Epochs	20
Energy MLIP	MACE-MP `medium-mpa-0` (**?**)
Hull reference	`PatchedPhaseDiagram` (MP, 2025-04-09)
Precision	bf16 (actor, reference)
Rollout engine	vLLM v0, TP=2

B. Chemeleon2 Reward Function

Table 4. Reward components active during GRPO fine-tuning. The † term is only used in the crystal system conditional run.

Component	Weight w_c	Normalization	Range (pre-norm)
r_E (energy above hull)	1.0	min–max	$[-1, 0]$
r_C (creativity, unique + novel)	1.0	–	$\{0\} \cup (0, 1]$
r_X (composition MMD)	1.0	min–max	$\mathbb{R}$
r_S (structure MMD)	0.1	min–max	$\mathbb{R}$
r_{CS} (crystal system) †	1.0	–	$\{0, 1\}$

The reward implementation is a port/adaptation of the Chemeleon2 reward stack. In the command used for the unconditional Wyckoff run, the active components are r_E, r_C, r_X, and r_S with no density, band-gap, neural reward-model, or global-normalization term. The crystal system term r_{CS} is enabled only for the crystal system conditional script.

Energy above hull. For each generated structure, the energy is predicted with MACE-MP and wrapped as a `PDEntry`; `PatchedPhaseDiagram.get_e_above_hull` is then used to obtain ΔE_{hull} in $\mathrm{eV/atom}$. The raw reward is

$$r_E(\mathcal{C}) = -\mathrm{clip}(\Delta E_{\text{hull}}(\mathcal{C}),\, 0,\, 1)\,. \tag{8}$$

Failed energy evaluations map to NaN and receive the maximum raw penalty $r_E = -1$.

Creativity. A structure is *unique* if it is the first member of its `StructureMatcher` group within the rollout batch, and *novel* if it matches no MP-20 reference structure with the same reduced formula. Writing $u(\mathcal{C}), v(\mathcal{C}) \in \{0, 1\}$ for these indicators,

$$r_C(\mathcal{C}) \;=\; \begin{cases} 1, & u(\mathcal{C}) \wedge v(\mathcal{C}), \\ 0, & \neg u(\mathcal{C}) \wedge \neg v(\mathcal{C}), \\ \min_{\mathcal{C}' \in \mathcal{N}(\mathcal{C})} d_{\mathrm{AMD}}(\mathcal{C}, \mathcal{C}'), & \text{otherwise}, \end{cases} \tag{9}$$

where d_{AMD} is the Average Minimum Distance descriptor metric (**?**) and $\mathcal{N}(\mathcal{C})$ contains reference and in-batch structures with the same reduced formula.

Distributional diversity. Let z denote either the structure-level or composition-level Chemeleon2 VAE embedding. With a cubic polynomial kernel $k(x, y) = (x^\top y/d + 1)^3$ we form a biased MMD^2 estimator and assign each rollout the change

in MMD induced by *leaving it out* of the batch:

$$\widehat{\mathrm{MMD}}^2(\mathbf{z}, \mathbf{z}^{\mathrm{ref}}) \;=\; \tfrac{1}{M(M-1)}\sum_{i\neq j} k(z_i, z_j) + \tfrac{1}{N(N-1)}\sum_{i\neq j} k(z_i^{\mathrm{ref}}, z_j^{\mathrm{ref}}) - \tfrac{2}{MN}\sum_{i,j} k(z_i, z_j^{\mathrm{ref}}), \tag{10}$$

$$r_{\{S,X\}}(\mathcal{C}^{(k)}) \;=\; \widehat{\mathrm{MMD}}^2_{\setminus k}(\mathbf{z}, \mathbf{z}^{\mathrm{ref}}) \;-\; \widehat{\mathrm{MMD}}^2(\mathbf{z}, \mathbf{z}^{\mathrm{ref}}). \tag{11}$$

Positive r_S, r_X indicate rollouts whose removal would increase the MMD to the reference distribution, so these samples receive higher reward under the implemented Chemeleon2 diversity term.

Crystal system target matching. For the crystal system conditional run, each prompt p carries a target crystal system $y_p \in \{\text{cubic}, \ldots, \text{triclinic}\}$. We compute the generated structure's crystal system with `SpacegroupAnalyzer` using `symprec`$= 0.1$ and `angle_tol`$= 5^\circ$, then assign

$$r_{CS}(\mathcal{C}; y_p) \;=\; \mathbf{1}[\,\mathrm{CS}(\mathcal{C}) = y_p\,]\,. \tag{12}$$

Table 5. Benchmark results for CrysTune$^{\mathrm{RL}}{}_{\mathrm{CS}}$ across all 7 crystal systems.

Metric	Cubic	Hexagonal	Monoclinic	Orthorhombic	Tetragonal	Triclinic	Trigonal
Validity (%)	**98.20**	58.00	27.20	31.96	74.72	31.64	36.92
Uniqueness (%)	81.51	83.66	93.68	91.11	56.10	**97.47**	91.87
Novelty (%)	6.92	84.62	96.62	95.74	66.33	**99.24**	88.52
SUN score	**0.1222**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
MSUN score	0.4481	0.5517	0.0000	0.0000	0.2141	0.3793	**0.5417**
Stability (%)	**0.6517**	0.1379	0.0000	0.0000	1.1242	0.0000	0.0000
Metastability (%)	**53.65**	4.41	0.00	0.13	11.83	0.38	1.63
E_{hull} (eV)	**0.205**	2.429	2.172	1.883	1.109	2.368	1.651
E_{form} (eV)	-0.332	2.068	1.294	1.011	0.693	1.324	0.993
Relaxation RMSE	**0.155**	0.978	1.491	1.357	1.485	1.670	1.258
Element diversity	**0.712**	0.703	0.636	0.678	0.695	0.610	0.669
Space group diversity	0.035	0.052	0.074	**0.078**	**0.078**	0.043	0.070
Site number diversity	1.890	5.173	4.721	3.212	3.504	**9.267**	6.186
Physical size diversity	0.213	0.239	**0.262**	0.256	0.218	0.234	0.260

C. Wyckoff Tables: 174 and 167

Table 6. Wyckoff positions for space group 174 ($P\bar{6}$).

Multiplicity	Wyckoff letter	Site symmetry	Coordinates
1	a	$\bar{6}$ ..	$(0,\ 0,\ 0)$
1	b	$\bar{6}$ ..	$(0,\ 0,\ 1/2)$
1	c	$\bar{6}$ ..	$(1/3,\ 2/3,\ 0)$
1	d	$\bar{6}$ ..	$(1/3,\ 2/3,\ 1/2)$
1	e	$\bar{6}$ ..	$(2/3,\ 1/3,\ 0)$
1	f	$\bar{6}$ ..	$(2/3,\ 1/3,\ 1/2)$
2	g	3 ..	$(0,\ 0,\ z),\ (0,\ 0,\ -z)$
2	h	3 ..	$(1/3,\ 2/3,\ z),\ (1/3,\ 2/3,\ -z)$
2	i	3 ..	$(2/3,\ 1/3,\ z),\ (2/3,\ 1/3,\ -z)$
3	j	m ..	$(x,\ y,\ 0),\ (-y,\ x-y,\ 0),$ $(-x+y,\ -x,\ 0)$
3	k	m ..	$(x,\ y,\ 1/2),\ (-y,\ x-y,\ 1/2),$ $(-x+y,\ -x,\ 1/2)$
6	l	1	$(x,\ y,\ z),\ (-y,\ x-y,\ z),$ $(-x+y,\ -x,\ z),\ (x,\ y,\ -z),$ $(-y,\ x-y,\ -z),\ (-x+y,\ -x,\ -z)$

Table 7. Wyckoff positions for space group 167 ($R\bar{3}c$, hexagonal axes). Unlike the primitive group in Table 6, the rhombohedral lattice carries three centering operations (top row): each listed coordinate is replicated by adding every centering vector, tripling the row count to give the full multiplicity. The $12c$ row therefore expands to $4 \times 3 = 12$ atoms in the unit cell.

Multiplicity	Wyckoff letter	Site symmetry	Coordinates (per centering block)
			Centering vectors: $(0,\ 0,\ 0)+\quad(\frac{2}{3},\ \frac{1}{3},\ \frac{1}{3})+\quad(\frac{1}{3},\ \frac{2}{3},\ \frac{2}{3})+$
6	a	32	$(0,\ 0,\ 1/4),\ (0,\ 0,\ 3/4)$
6	b	$\bar{3}$	$(0,\ 0,\ 0),\ (0,\ 0,\ 1/2)$
12	c	3 ..	$(0,\ 0,\ z),\ (0,\ 0,\ -z+1/2),$ $(0,\ 0,\ -z),\ (0,\ 0,\ z+1/2)$
18	d	$\bar{1}$	$(1/2,\ 0,\ 0),\ (0,\ 1/2,\ 0),$ $(1/2,\ 1/2,\ 0),\ (0,\ 1/2,\ 1/2),$ $(1/2,\ 0,\ 1/2),\ (1/2,\ 1/2,\ 1/2)$
18	e	.2	$(x,\ 0,\ 1/4),\ (0,\ x,\ 1/4),$ $(-x,\ -x,\ 1/4),\ (-x,\ 0,\ 3/4),$ $(0,\ -x,\ 3/4),\ (x,\ x,\ 3/4)$
36	f	1	$(x,\ y,\ z),\ (-y,\ x-y,\ z),$ $(-x+y,\ -x,\ z),\ (y,\ x,\ -z+1/2),$ $(x-y,\ -y,\ -z+1/2),$ $(-x,\ -x+y,\ -z+1/2),$ $(-x,\ -y,\ -z),\ (y,\ -x+y,\ -z),$ $(x-y,\ x,\ -z),\ (-y,\ -x,\ z+1/2),$ $(-x+y,\ y,\ z+1/2),$ $(x,\ x-y,\ z+1/2)$

D. Inference Projection

Tag	Failure mode
`parsing_error`	Syntax, space group, Wyckoff symbol, or numeric field (Sec. 3.3)
`comp_invalid`	Element symbol $\hat{e}_i$ unrecognized by pymatgen
`max_atoms`	Any $Z(\hat{e}_i) > 94$
`valid`	Parser returns $\mathcal{P}(\hat{s}) \in \mathcal{S}$

For ease of reimplementation, we restate the C2 and C3 operations of Section 3.3 in compact symbolic form.

Free-axis set. For each (n_G, ℓ), let $\mathcal{F}(n_G, \ell) \subseteq \{1, 2, 3\}$ be the indices of axes that appear as free parameters in $\tau(n_G, \ell)$; the complement $\{1, 2, 3\} \setminus \mathcal{F}(n_G, \ell)$ is the set of pinned axes. For example, $\mathcal{F}(174, 3j) = \{1, 2\}$ (axes x and y free) and $\mathcal{F}(167, 6a) = \varnothing$ (all axes pinned).

C2 (fixed-coordinate snap). For each site i,

$$\hat{r}_{i,a} \;\leftarrow\; \tau(\hat{n}_G, \hat{\ell}_i)_a \quad \text{for every } a \notin \mathcal{F}(\hat{n}_G, \hat{\ell}_i), \tag{13}$$

leaving free axes unchanged. After this update, the pinned components of $\hat{\mathbf{r}}_i$ exactly match the site template.

Site-specific Wyckoff operations. For Wyckoff letter ℓ in space group n_G, the International Tables list $m = \mathrm{Mult}(n_G, \ell)$ affine maps $W_1^{(\ell)}, \ldots, W_m^{(\ell)}$, each of the form $W_j\mathbf{r} \;=\; \mathbf{R}_j\mathbf{r} + \mathbf{t}_j \pmod{\mathbf{1}}$, following the same convention as the elements of G in Section 2; for centered lattices, the centering vectors are folded into the $\mathbf{t}_j$. The rows of Tables 6 and 7 display these maps as coordinate expressions for two example space groups.

C3 (orbit expansion). The m_i atomic positions of site i are

$$\mathcal{O}_i \;=\; \left\{ W_j^{(\hat{\ell}_i)}\, \hat{\mathbf{r}}_i \;:\; j = 1, \ldots, m_i \right\}. \tag{14}$$

The CIF is built from $\bigcup_i \mathcal{O}_i$, paired with the corresponding species $\hat{e}_i$.

E. Additional Results

E.1. SFT with more epochs

To test whether longer supervised fine-tuning improves the Wyckoff generator, we compare the 4-epoch checkpoint against the 10-epoch checkpoint. This is a comparison of the generation+canonicalization auxiliary setting.

The 4-epoch checkpoint appears better on most metrics that directly determine whether a generated crystal is usable: validity is higher by 9.44 percentage points, novelty is higher by 6.38 percentage points, stability remains nonzero, and the relaxed structures have lower energy above hull and lower relaxation RMSE. The 10-epoch checkpoint mainly improves uniqueness, metastability, site-number diversity, and physical-size diversity, but these gains come with a substantial drop in validity and worse physical-realism metrics.

A plausible explanation is mild over-training of the SFT objective. Additional epochs can make the model more confident in frequent textual/Wyckoff patterns without improving the geometric consistency required by the hard parser, Wyckoff-site constraints, and downstream relaxation. This can increase within-sample uniqueness by reducing exact repeats, while still producing more invalid or physically strained structures. The lower novelty and space group diversity after 10 epochs also suggest that longer SFT may pull generations closer to dominant training-distribution templates. In this setting, the 4-epoch checkpoint therefore provides a better balance between learning the representation and preserving enough generative diversity and physical plausibility for evaluation.

E.2. Density-conditioned generation

To evaluate conditional generation on a continuous property, we fine-tuned the model using a distance-based reward that penalizes deviations from a target density. As illustrated in Figure 4, the SFT baseline performs predictably poorly at

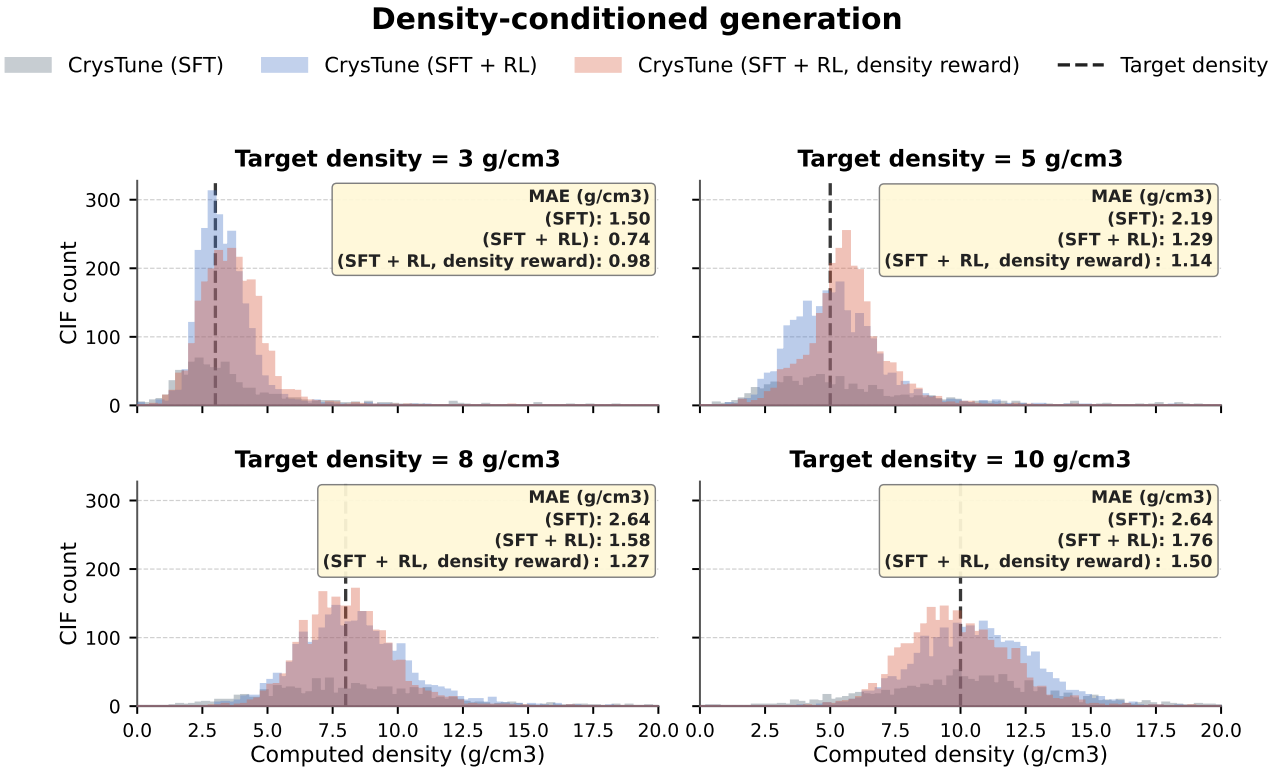

Figure 4. Distribution of computed densities for target density values of 3, 5, 8, and 10 $\mathrm{g/cm^3}$. The histograms compare the SFT baseline, standard RL (SFT + RL), and RL with an explicit density penalty (SFT + RL, density reward), illustrating the modest improvements and ongoing challenges in optimizing continuous material properties.

Table 8. Comparison of 4-epoch and 10-epoch SFT checkpoints for the Wyckoff auxiliary-task model. Best result per metric is **bolded**. ↑/↓ denotes higher/lower is better.

Metric	4 epochs	10 epochs
Generation Quality		
Validity (%) ↑	**46.33**	36.89
Uniqueness (%) ↑	90.83	**98.05**
Novelty (%) ↑	**61.46**	55.08
Stability		
SUN score ↑	**0.000**	**0.000**
MSUN score ↑	**0.833**	0.781
Stability (%) ↑	**0.208**	0.000
Metastability (%) ↑	8.75	**17.19**
Energy/Relaxation		
E_{hull} (eV/atom) ↓	**1.362**	1.508
E_{form} (eV/atom) ↓	**0.863**	0.868
Relax. RMSE ↓	**0.741**	0.870
Diversity		
Element diversity ↑	**0.678**	**0.678**
Space group diversity ↑	**0.161**	0.139
Site number diversity ↑	8.885	**11.084**
Physical size diversity ↑	0.306	**0.342**

matching the target density across all tested values (3, 5, 8, and 10 $\mathrm{g/cm^3}$), which is expected since standard supervised training lacks strong explicit optimization for precise continuous targets. While standard RL fine-tuning narrows the output distributions, applying the explicit density reward leads to modest improvements in Mean Absolute Error (MAE) for most targets compared to standard RL, and the model still struggles to tightly center the generations around the exact target values. Ultimately, these distributions highlight the inherent difficulty of optimizing continuous property values through current RL fine-tuning frameworks.

E.3. Alternative parsing strategy

We evaluate the effect of making the Wyckoff parser more stringent in terms of space group consistency and flexible in terms of fixed coordinate snapping. The default parser follows the projection pipeline described in Section 3.3: after reading the generated space group, Wyckoff symbols, coordinates, and lattice, it snaps pinned coordinates to the canonical Wyckoff template and then expands the corresponding orbit. The alternative mode adds two hard validity checks before this expansion: the generated Wyckoff letter must be a valid position for the generated space group, and the generated lattice metric must match the crystal system implied by that space group. For example, tetragonal and cubic groups must satisfy the expected length and angle equalities, while trigonal groups allow either the hexagonal or rhombohedral setting.

This mode also changes how fixed-coordinate snapping is applied. In the default parser, pinned axes are snapped before orbit expansion. In the alternative mode, the parser first walks the symmetry orbit to find the canonical generator, which can recover cases where the model placed a free coordinate in a symmetry-equivalent but nominally pinned slot; it then snaps only the pinned axes of this recovered generator and expands the orbit. Thus, strict parsing is not only more restrictive: it also makes the coordinate correction more symmetry-aware.

The two parsers produce similar validity rates: 52.21% for default parsing and 53.16% for alternate parsing. The stricter evaluation improves the metrics most closely tied to energy and relaxation. It also increases metastability from 30.29% to 32.39% and MSUN from 0.445 to 1.006. These improvements are consistent with the additional space-group/lattice consistency checks and symmetry-aware coordinate recovery filtering or correcting geometrically inconsistent outputs. The tradeoff is a small reduction in uniqueness, novelty, and diversity, suggesting that strict parsing preferentially retains or reconstructs more regular structures while discarding some atypical generations.

Table 9. Default and strict Wyckoff parsing for the `aux-form` SFT model. Best result per metric is **bolded**. ↑/↓ denotes higher/lower is better.

Metric	Default	Alt
Generation Quality		
Validity (%) ↑	52.21	**53.16**
Uniqueness (%) ↑	**94.65**	93.76
Novelty (%) ↑	**29.40**	28.77
Stability		
SUN score ↑	**0.223**	0.000
MSUN score ↑	0.445	**1.006**
Stability (%) ↑	1.336	**1.408**
Metastability (%) ↑	30.29	**32.39**
Energy/Relaxation		
E_{hull} (eV/atom) ↓	1.049	**0.850**
E_{form} (eV/atom) ↓	0.511	**0.330**
Relax. RMSE ↓	0.500	**0.470**
Diversity		
Element diversity ↑	**0.712**	0.695
Space group diversity ↑	**0.130**	0.122
Site number diversity ↑	**6.404**	5.911
Physical size diversity ↑	**0.314**	0.307

E.4. Auxiliary Task Ablations

Table 10. Ablation study for CrysTune SFT variants. Best result per metric is **bolded**. ↑/↓ denotes higher/lower is better.

Metric	CrysTune^{gen}	CrysTune$^{gen+can}$	CrysTune^{SFT}
Generation Quality			
Validity (%) ↑	34.14	46.33	**52.21**
Uniqueness (%) ↑	89.71	90.83	**94.65**
Novelty (%) ↑	54.66	**61.46**	29.40
Stability			
SUN (%) ↑	0.0	0.0	**0.223**
MSUN (%) ↑	0.322	**0.833**	0.445
Stability (%) ↑	0.322	0.208	**1.336**
Metastability (%) ↑	12.54	8.75	**30.29**
Energy/Relaxation			
E_{hull} (eV/atom) ↓	1.194	1.362	**1.049**
E_{form} (eV/atom) ↓	**0.511**	0.863	**0.511**
Relax. RMSE ↓	0.889	0.741	**0.500**
Diversity			
Element ↑	0.669	0.678	**0.712**
Space group ↑	0.139	**0.161**	0.130
Site number ↑	7.46	**8.89**	6.40
Physical size ↑	**0.339**	0.306	0.314

E.5. Auxiliary-task exact-match evaluation

We separately evaluate the two auxiliary skills introduced in the Wyckoff SFT mixture: canonicalization and Wyckoff coordinate-template prediction. For canonicalization, prompts are constructed from the MP-20 test split by first converting each crystal to the canonical Wyckoff string and then applying every non-identity pure translation returned by the symmetry analyzer. Crystals with no nontrivial pure translation are not included, so the resulting benchmark contains only genuine symmetry-equivalent canonicalization cases. For template prediction, we exhaustively enumerate all Wyckoff positions exposed by `pyxtal` for space groups 1–230 and ask the model to output the canonical generator coordinate form.

Table 11. Exact-match performance on the two Wyckoff auxiliary tasks. The base model is `meta-llama/Llama-2-7b-chat-hf`; the SFT model is CrysTune^{SFT}. Accuracy is computed after collapsing whitespace in the first generated non-empty line.

Task	Model	Examples	Correct	Accuracy (%)
Canonicalization	Base LLaMA-2-7B-chat	9,399	0	0.00
Canonicalization	CrysTune^{SFT}	9,399	3,671	39.06
Wyckoff template prediction	Base LLaMA-2-7B-chat	1,731	0	0.00
Wyckoff template prediction	CrysTune^{SFT}	1,731	696	40.21

The base LLaMA-2-chat model receives the same instructional prompts but obtains zero exact matches on both tasks. This is expected under a strict string-level metric: the model has not been trained to emit the repository's compact Wyckoff serialization, and even semantically related free-form answers are counted incorrect if they do not exactly match the target line. The SFT checkpoint learns both auxiliary behaviors substantially better, reaching 39.06% exact match on translated-to-canonical Wyckoff strings and 40.21% exact match on the full space-group/Wyckoff-position template inventory.

E.6. Canonicalization with mixed orbit perturbations

The main text describes the default canonicalization auxiliary task, where the non-canonical input is produced by applying one non-trivial pure translation $\mathbf{v} \in \mathcal{T}(\mathcal{C})$ to every Wyckoff site and training the model to recover $s(\mathcal{C})$ (Equation (6)). We compare this default setting with a broader `mixed` mode implemented in the fine-tuning code. In the default `global` mode, the perturbed input is $\hat{s}_{\mathbf{v}}$: all site coordinates are shifted by the same $\mathbf{v}$. In `mixed` mode, each canonicalization example either uses this same global shift or, with equal probability, replaces each site coordinate independently by a random non-identity element of that site's Wyckoff orbit. If no non-trivial pure translation is available, mixed mode falls back to the per-site orbit perturbation. The target remains the canonical string $s(\mathcal{C})$ in both cases.

This ablation therefore extends the canonicalization task from the origin convention alone to both conventions described in Section 3.1: the global origin choice and the per-site representative choice. The change makes the auxiliary task more comprehensive, but also less structured. A global translation can be inverted by identifying a single shift shared by every site, whereas the per-site orbit perturbation requires the model to recover the first listed representative separately for each (n_G, ℓ_i).

Mixed canonicalization degrades most downstream metrics. Validity decreases from 52.21% to 35.21%, uniqueness decreases from 94.65% to 72.61%, and the stability-related metrics deteriorate: the stable fraction decreases from 1.34% to 0.23%, metastability decreases from 30.29% to 16.48%, and the mean energy above hull increases from 1.049 to 1.439 eV/atom. Relaxation RMSE also increases from 0.500 to 0.724, indicating that the generated structures are less geometrically consistent before relaxation. The main improvements are novelty, which increases from 29.40% to 34.43%, and site-number diversity, which increases from 6.40 to 7.76.

These results suggest that per-site orbit perturbations make the canonicalization objective less aligned with the generation and inference pipeline used in the main model. The default task teaches a simple global invariance while keeping each site's canonical representative structure intact, so it reinforces the origin convention without exposing the model to many alternative coordinate layouts. Mixed mode presents a richer set of symmetry-equivalent strings, but the independently sampled site representatives make the input distribution more heterogeneous than the canonical generation target. This appears to loosen the model's attachment to common canonical templates, increasing novelty and site-number diversity, but it also weakens the Wyckoff grammar needed for valid, low-energy outputs. We therefore retain the default global canonicalization mode in the main experiments.

E.7. LeMat-GenBench Metrics

We evaluate each generated crystal after parsing it as a `pymatgen` structure. The reported JSON file is a compact post-processed summary of the benchmark output. Ratio-based quantities are reported as percentages, while energy and displacement quantities are reported in their native units. Although the benchmark configuration supports multiple machine-learned interatomic potentials (MLIPs), all stability-related values reported here use the ORB model only, specifically the ORB-specific convex hull results, rather than an ensemble average.

Table 12. Effect of changing the canonicalization auxiliary task from the default global pure-translation mode to mixed global/per-site orbit mode for the `aux-form` SFT model. Best result per metric is **bolded**. ↑/↓ denotes higher/lower is better.

Metric	Default	Mixed orbit
Generation Quality		
Validity (%) ↑	**52.21**	35.21
Uniqueness (%) ↑	**94.65**	72.61
Novelty (%) ↑	29.40	**34.43**
Stability		
SUN score ↑	**0.223**	0.000
MSUN score ↑	**0.445**	0.114
Stability (%) ↑	**1.336**	0.227
Metastability (%) ↑	**30.29**	16.48
Energy/Relaxation		
E_{hull} (eV/atom) ↓	**1.049**	1.439
E_{form} (eV/atom) ↓	**0.511**	1.162
Relax. RMSE ↓	**0.500**	0.724
Diversity		
Element diversity ↑	**0.712**	**0.712**
Space group diversity ↑	**0.130**	0.126
Site number diversity ↑	6.404	**7.761**
Physical size diversity ↑	**0.314**	0.278

Validity. `overall_validity` is the percentage of generated structures that pass all validity checks. A structure is counted as valid only if it is charge neutral within the configured tolerance, has no unrealistically short interatomic distances, and passes basic physical plausibility checks on density, lattice parameters, CIF serialization, and space-group assignment.

Uniqueness and novelty. `overall_uniqueness` is the percentage of non-duplicate structures in the generated set. Duplicate detection is performed by the configured structural fingerprinting or structure-matching method. `overall_novelty` is the percentage of generated structures that are not matched to the LeMat-Bulk reference set under the same comparison procedure. Thus, uniqueness is measured within the generated batch, whereas novelty is measured against known reference materials.

Stability and MetaSUN. For each structure, ORB is used to compute the formation energy and the energy above the ORB convex hull. `stability_score` is the percentage of structures with ORB energy above hull less than or equal to 0 eV/atom. `metastability_score` is the percentage with ORB energy above hull less than or equal to 0.1 eV/atom. The mean ORB energy above hull is reported as `e_above_hull` in eV/atom, and the mean ORB formation energy is reported as `formation_energy` in eV/atom. Lower values indicate structures that are closer to the predicted thermodynamic hull.

SUN and MSUN. `sun_score` is the percentage of generated structures that are simultaneously stable, unique, and novel. The computation first filters structures by stability, then removes duplicates within the stable subset, and finally tests novelty against LeMat-Bulk. `msun_score` applies the same unique-and-novel criterion to the metastable subset, using the 0.1 eV/atom threshold. In the SUN calculation, structures exactly stable with respect to the hull are handled in the stable subset, while metastable structures are those above 0 and up to the metastability threshold.

Relaxation and diversity. `relaxation_rmse` is the mean root-mean-square Cartesian displacement between the input structure and the ORB-relaxed structure, in Angstrom. It measures how much the generated geometry moves during local relaxation. The diversity metrics summarize the spread of the generated batch: element diversity measures coverage of chemical elements, space-group diversity measures coverage of crystallographic symmetries, site-number diversity measures the spread in numbers of atomic sites per cell, and physical-size diversity summarizes variation in density, lattice parameters, volume, and packing factor.

Table 13. Summary of fields in the parsed benchmark JSON.

Field	Meaning
`overall_validity`	Percent passing all validity checks
`overall_uniqueness`	Percent unique within the generated set
`overall_novelty`	Percent not matched to LeMat-Bulk
`sun_score`	Percent stable, unique, and novel
`msun_score`	Percent metastable, unique, and novel
`stability_score`	Percent with ORB $E_{\text{hull}} \leq 0$ eV/atom
`metastability_score`	Percent with ORB $E_{\text{hull}} \leq 0.1$ eV/atom
`e_above_hull`	Mean ORB energy above hull, eV/atom
`formation_energy`	Mean ORB formation energy, eV/atom
`relaxation_rmse`	Mean ORB relaxation displacement, Angstrom
`element_diversity`	Elemental coverage/diversity score
`space_group_diversity`	Space-group coverage/diversity score
`site_number_diversity`	Diversity of atom counts per cell
`physical_size_diversity`	Diversity of density, lattice, volume, and packing features

E.8. Compute

All experiments were conducted using a hardware setup equipped with four NVIDIA H100 80GB GPUs, 128GB of system RAM, and 24 to 48 CPU cores. In terms of computational time, the supervised fine-tuning (SFT) experiments required approximately 4 hours to complete, whereas the reinforcement learning (RL) fine-tuning stages took roughly 10 hours. During the inference phase, generating 2,500 valid CIF files took approximately 30 minutes. Finally, the comprehensive downstream evaluation of these generated structures required about 1 hour and 30 minutes.