CRYSTAL: Coordinated Multi-Objective Reinforcement Learning for Crystal Generation

Zhan'ao Yao [*1,2] **Jingyuan Shu** [*3,4] **Boxuan Zhang** [*5,6] **Xiaoyu Wu** [*4] **Rongyan Wang** [1,2] **Tingwei Chen** [7] **Linjing Li** [5,6] **Daniel Dajun Zeng** [5,6] **Yu-Dong Yao** [*3] **Xiaolin Zhao** [*1,2] **Jiahui Shi** [*5,6] **Jianjun Liu** [*1,2]

Abstract

In materials science, *de novo* generation of crystal structures that simultaneously satisfy stability, diversity, novelty, and validity is a central challenge for accelerating new materials discovery. Large language models, leveraging a two-stage paradigm of fine-tuning and reinforcement learning with verifiable rewards (RLVR), have demonstrated significant advantages over conventional diffusion-based approaches. However, existing RLVR methods typically constrain only one or two objectives, making it difficult to coordinate multi-objective optimization and prone to reward hacking—where one objective collapses sharply during training, leading to severe imbalance in overall performance. To address this, we propose CRYSTAL, a method based on Group Relative Policy Optimization that jointly optimizes four critical attributes: leveraging physics-grounded verifiable signals to ensure structural stability and physical correctness, regulating novelty through explicit comparison with established materials databases, and generating multiple candidate structures within a single inference step to explicitly promote diversity, ultimately achieving coordinated multi-objective control through multiplicative aggregation. The proposed method effectively mitigates reward hacking in multi-objective reinforcement learning, achieves state-of-the-art performance in comprehensive multi-objective evaluations, and attains an S.U.N. metric of 25.1 on 1,000 generated materials, demonstrating the potential to further extend large language models toward multi-objective on-demand materials design.

[1]State Key Laboratory of High Performance Ceramics, Shanghai Institute of Ceramics, Chinese Academy of Sciences, 1295 Dingxi Road, Shanghai 200050, China [2]Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China [3]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China [4]Beijing Wenge Technology Co., Ltd., Beijing, China [5]State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. [6]School of Artificial Intelligence, University of Chinese Academy of Science, Beijing 100049, China. [7]Faculty of Information, Liaoning University, Shenyang, China. Correspondence to: Jianjun Liu <jliu@mail.sic.ac.cn>, Jiahui Shi <jiahui.shi@ia.ac.cn>, Xiaolin Zhao <zhaoxiaolin@mail.sic.ac.cn>.

Accepted by ICML 2026 AI for Science Workshop

1. INTRODUCTION

Although materials databases such as Materials Project (Jain et al., 2013) and OQMD (Saal et al., 2013) have accumulated data on hundreds of thousands of known compounds, theoretical studies suggest that many additional stable compounds remain unexplored (Davies et al., 2016). Traditional high-throughput screening and element-substitution strategies are limited by combinatorial explosion and localized search, making it difficult to systematically explore this vast space. By learning the underlying distributions of material compositions and crystal structures, generative models enable the direct, from-scratch design of novel candidate materials in previously unsampled regions of chemical space (Xie et al., 2021; Guo et al., 2025), providing a powerful paradigm for accelerating materials discovery.

Recently, large language models (LLMs) fine-tuned (Gruver et al., 2024; Sriram et al., 2024) on crystallographic text representations and further enhanced by reinforcement learning (RL)(Cao & Wang, 2026) have demonstrated superior performance and faster inference in crystal structure generation compared to variational autoencoders (VAEs) (Xie et al., 2021) and diffusion models (Zeni et al., 2025; Jiao et al., 2023; Miller et al., 2024). However, current RL-based approaches still face critical challenges in coordinating multiple generation objectives. Most existing methods(Cao & Wang, 2026)(Mohanty et al., 2026) design reward functions around a single objective (*e.g.*, stability), where improving one performance metric often leads to the degradation of others. Offline RL approaches such as PLaID (Xu et al., 2025b), which adopt DPO-based preference optimization, improve upon this by incorporating stability and novelty as

pairwise preferences; however, uniqueness is inherently a batch-level metric that cannot be naturally formulated as pairwise comparisons, limiting joint optimization within the DPO paradigm. Moreover, LLMs' autoregressive generation lacks the inherent stochastic exploration mechanism of diffusion models, and the initial generation quality after SFT falls short of diffusion model pre-training levels, making diversity collapse particularly severe during RL training.

To address these challenges, we propose CRYSTAL (Coordinated Multi-Objective Reinforcement Learning for Crystal Generation), a multi-objective reward framework built upon Group Relative Policy Optimization (GRPO) (Shao et al., 2024). CRYSTAL leverages physics-grounded verifiable signals to ensure structural stability and physical correctness, regulates novelty through explicit comparison with established materials databases, and generates multiple candidate structures within a single inference step, enabling diversity-oriented policy updates that explicitly promote structural uniqueness. The multiplicative aggregation of sub-objectives enforces hard constraints on each dimension—since any factor approaching zero drives the product to zero—effectively mitigating reward hacking without the trade-off assumptions inherent in Pareto-based or additive aggregation approaches. Experimental results demonstrate that CRYSTAL achieves state-of-the-art performance, with the S.U.N. metric reaching 25.1 on 1000 generated materials, representing an approximately 70% improvement over prior methods, and attaining broader coverage of chemical space. These results demonstrate the potential to further extend LLMs toward multi-objective on-demand materials design.

2. Related Work

Recently, reinforcement learning has been increasingly introduced into crystal structure generation to improve generation quality. Within the diffusion model framework, MatInvent (Chen et al., 2025) formulates equivariant denoising as a multi-step decision-making process for target-property-driven conditional generation, applying S.U.N. as a pre-filtering condition and computing target property rewards only for structures that pass the filter, with min aggregation or weighted summation to reduce multiple property objectives into a single scalar reward for multi-objective design. Chemeleon2 (Park & Walsh, 2025) targets unconditional generation and directly incorporates generation quality metrics such as stability as reward signals, adopting additive aggregation of multiple rewards. Both works adopt evaluation metrics such as MSUN that differ from the standard S.U.N. framework. Within the LLM model framework, CrystalFormer-RL (Cao & Wang, 2026) employ PPO-based online training, while CrysTex-RL (Mohanty et al., 2026) adopts GRPO. However, these methods all rely on a single reward signal (*e.g.*, energy above hull) for optimization, where improving one performance metric often leads to the degradation of others—for example, CrystalFormer-RL achieves a stability rate of 73.4%, but its S.U.N. remains only 21.6%. Moreover, the autoregressive generation of LLMs lacks the inherent random exploration mechanism of diffusion models, and the initial generation quality after SFT falls short of the pre-training level of diffusion models, making diversity collapse more prone to occur during RL training. How to effectively maintain generation diversity in multi-objective optimization remains an open challenge.

3. Method

As shown in the Figure 1, we propose the CRYSTAL pipeline. We first perform supervised fine-tuning on the LLM to enable it to generate complete, crystallographically compliant structural descriptions from a given target instruction. The training data is adopted from the previous work (Xu et al., 2025b). Specifically, crystal structures are encoded using the Wyckoff notation—a compact textual representation that explicitly incorporates key physical knowledge such as space group symmetry, thereby significantly increasing information density while drastically reducing sequence length. Through this stage of learning, the general-purpose language model gradually develops an understanding of the relationship between a crystal's chemical composition and its atomic arrangement, establishing a reliable initial policy for the subsequent phase of physics-informed reinforcement learning.

3.1. Valid Constraint Reward

Since language models can generate atomic arrangements that violate fundamental physical rules under unconstrained conditions (such as atomic overlap or charged crystals), directly using them for materials discovery would lead to a large number of invalid candidates, severely reducing sampling efficiency. Therefore, we introduce three binary validity criteria to rapidly screen for physically reasonable crystal structures. $S = \{s_1, s_2, \ldots, s_n\}$ denote all structures generated in a single rollout, where each s_i represents an individual crystal structure.

We define the structural validity criterion $\mathcal{V}_{\text{struct}}(s) : s \to \{0, 1\}$ as:

$$\mathcal{V}_{\text{struct}}(s) = \begin{cases} 1, & \text{if } \min_{i \neq j} d_{ij} \geq 0.5 \text{ Å}, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where d_{ij} denotes the distance between atoms i and j, ensuring no non-physical atomic overlaps.

We define the charge neutrality criterion $\mathcal{V}_{\text{charge}}(s) : s \to$

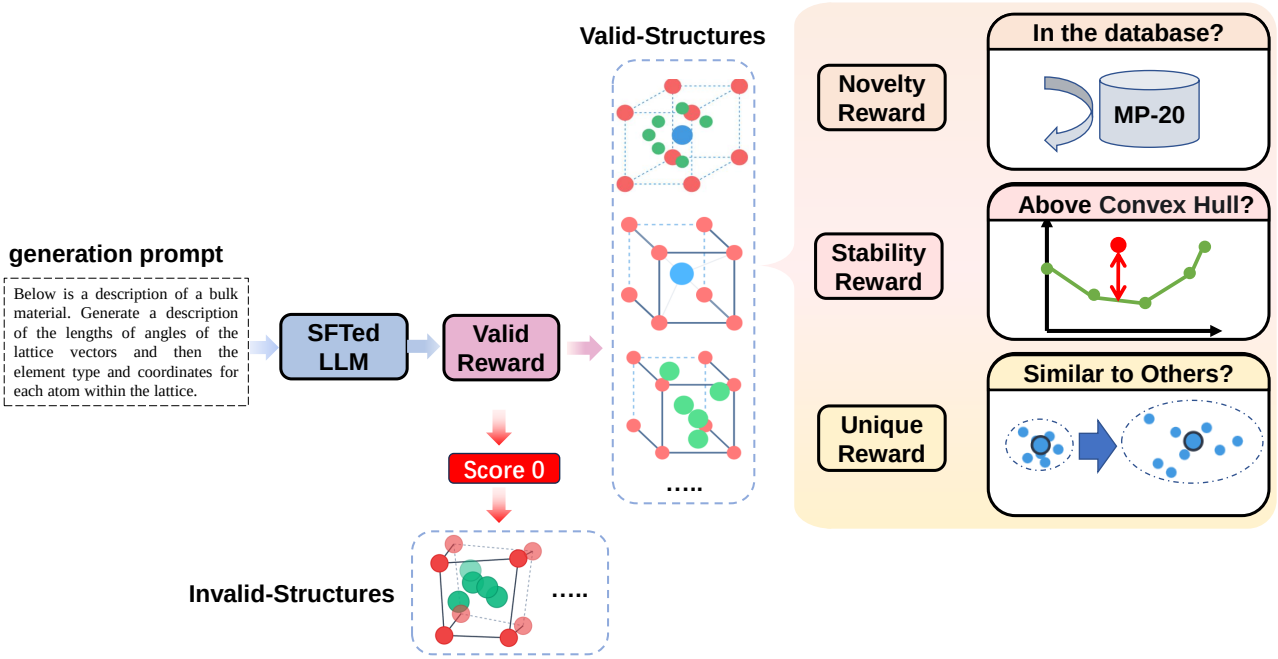

Figure 1. This is the pipeline of our CRYSTAL. We further fine-tune the model using reinforcement learning, with a reward function composed of four components: Valid Reward, Novelty Reward, Stability Reward, and Uniqueness Reward.

$\{0, 1\}$ as:

$$\mathcal{V}_{\text{charge}}(s) = \begin{cases} 1, & \text{if } \sum_{i=1}^{N} q_i = 0, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where q_i is the formal charge of atom i, ensuring the crystal conforms to basic chemical stability rules.

Only when the structure can stably converge to a local energy minimum is it considered mechanically feasible. If relaxation fails, even if the previous two criteria are satisfied, the structure is still judged as invalid. So we perform geometric relaxation via machine learning interatomic potential (MLIP) (Chen & Ong, 2022).We define the relaxation convergence criterion $\mathcal{V}_{\text{relax}}(S) : S \to \{0, 1\}$ as:

$$\mathcal{V}_{\text{relax}}(s) = \begin{cases} 1, & \text{if } s \text{ passes MLIP relaxation,} \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

where we perform geometric relaxation via MLIP,

The overall validity is determined by the conjunction of all three criteria:

$$\mathcal{V}(s) = \mathcal{V}_{\text{struct}}(s) \cdot \mathcal{V}_{\text{charge}}(s) \cdot \mathcal{V}_{\text{relax}}(s). \tag{4}$$

The reward value is then assigned as $r_{valid} = \mathcal{V}(S)$, which equals 0 if any criterion is not satisfied, and 1 otherwise. Finally, we define $S_v = \{s \in S \mid \mathcal{V}(s) = 1\}$ as the set of all valid structures that pass all three criteria and relaxed.

3.2. Stability Reward

In computational materials discovery, the convex hull energy E_{hull} (Jain et al., 2013) serves as a key indicator of thermodynamic stability. It quantifies the energy difference between a candidate structure and the most stable phases at the same composition. Structures with $E_{\text{hull}} \leq 0$ are thermodynamically stable and considered synthesizable, while those with small positive E_{hull} (e.g., ≤ 0.1 eV/atom) are metastable and may still be accessible under kinetic control. Thus, E_{hull} directly enables inference of both stability and synthetic feasibility.

To guide the model toward generating diverse stable crystal structures, we adopt a continuous reward function rather than a single stability threshold. This design provides graded incentives for materials with varying functional properties (such as thermoelectrics or photovoltaics), which often remain in metastable states while fully suppressing structures with $E_{\text{hull}} > 0$ that strictly violate chemical space exploration capabilities.

Specifically, for each candidate structure $s \in S_v$ that passes the validity criteria, we first perform geometric relaxation using machine learning interatomic potential (MLIP) to obtain the optimized structure s_{relaxed}, and then use the same MLIP model to calculate its convex hull energy with respect

to elemental composition:

$$E_{\text{hull}-s} = \text{MLIP}(s_v). \tag{5}$$

In materials science, $E_{\text{hull}} \leq 0.1$ eV/atom is widely recognized as a threshold for synthesizable stable structures. Therefore, we define the following energy-based reward function:

$$r_{stab}(s) = \begin{cases} 1, & E_{\text{hull}} - s \leq 0, \\ 1 - \frac{E_{\text{hull}-s} - 0.1}{0.1}, & 0 < E_{\text{hull}-s} \leq 0.1, \\ 0, & E_{\text{hull}-s} > 0.1, \end{cases} \tag{6}$$

where the reward linearly decays from 1 to 0 as $Ehull_s$ increases from 0 to 0.1 eV/atom. This linear decay provides smooth gradient signals for thermodynamic stability guidance, while simultaneously enabling prediction of metastable structures in the $(0, 0.1]$ region. High-energy structures ($E_{\text{hull}-s} > 0.1$) are completely suppressed with zero reward. Finally, we define $S_{stable} = \{s \in S_v \mid E_{\text{hull}-s} \leq 0.1\}$ as the set of all stable structures.

3.3. Novelty Reward

In materials discovery, generating novel structures is as crucial as achieving thermodynamic stability. Without novelty incentives, models converge to repeatedly generating known materials, limiting exploration of uncharted chemical space. We define a binary novelty criterion $\mathcal{N}(s) : s \to \{0, 1\}$ that evaluates whether structure s is sufficiently distinct from all known materials in reference database S_{db}, using two complementary descriptors(CrystalNNFingerprint and ElementProperty) from Matminer (Ward et al., 2018). Meanwhile, we introduce a novelty reward on the stable structure set S_{stable}.

The structural descriptor uses CrystalNNFingerprint:

$$\text{fp}_{\text{struc}}(s) = \text{CrystalNNFingerprint}(s), \tag{7}$$

encoding local coordination environments and topological connectivity, independent of chemical identity.

The compositional descriptor uses ElementProperty:

$$\text{fp}_{\text{comp}}(s) = \text{ElementProperty}(s), \tag{8}$$

quantifying elemental composition via atomic properties (electronegativity, radius, electron count), independent of geometric arrangement.

These orthogonal descriptors separately characterize *what* elements are present versus *how* they are arranged. We define:

$$\mathcal{N}(s) = \begin{cases} & \min\limits_{x \in S_{\text{db}}} \|\text{fp}_{\text{struc}}(s) - \text{fp}_{\text{struc}}(x)\|_2 > d_{\text{struc}} \\ 1, & \text{and } \min\limits_{x \ \in S_{\text{db}}} \|\text{fp}_{\text{comp}}(s) - \text{fp}_{\text{comp}}(x)\|_2 > d_{\text{comp}}, \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where thresholds d_{struc} and d_{comp} follow prior work (**?**) standards.

We set $r_{novel}(s) = \mathcal{N}(s)$. Specifically, the conjunction (AND) is critical: disjunction (OR) would allow satisfying novelty through only compositional or only structural innovation. This creates unbalanced exploration where models exploit shortcuts—discovering new compositions with known structures, or new structures with known compositions—while ignoring rich spaces like polymorphs. The AND constraint enforces simultaneous innovation across both dimensions, preventing such degeneracy.

3.4. Uniqueness Reward

While stability and novelty incentives guide generation toward thermodynamically stable structures, initial experiments reveal a critical issue: the model tends to converge toward a few high-reward structures, repeatedly generating near-identical or slightly perturbed variants. This model collapse behavior wastes computational resources on redundant candidates rather than exploring diverse chemical space. To address this, we introduce a multi-level diversity reward on the stable structure set S_{stable}.

Given the generated stable structure sequence $\{s_1, s_2, \ldots, s_n\}$, we first employ StructureMatcher (Ong et al., 2013) to evaluate structural similarity. This tool compares the geometric arrangements of structures *with identical elemental composition*, determining whether they are crystallographically equivalent up to symmetry operations and atomic position perturbations. We define the matching function:

$$M(s_i, s_j) = \mathbb{I}(\text{StructureMatcher}().\text{fit}(s_i, s_j)), \tag{10}$$

where $M(s_i, s_j) = 1$ indicates that structures s_i and s_j are structurally equivalent (same elements, same geometric arrangement), and $M(s_i, s_j) = 0$ otherwise. Note that StructureMatcher inherently requires matching elemental composition before comparing structures.

However, structural matching alone is insufficient as it cannot adequately account for compositional diversity. Therefore, we introduce penalties based on chemical element comparison $r(s)$ and stoichiometry comparison $k(s)$. Structures with minor compositional variations (such as ABO_3 perovskites differing only in A-site cation) are penalized to encourage broader exploration. We define separate counters for element set and stoichiometry repetition:

$$c_r^{(t)}(s) = \sum_{j=1}^{t} \mathbb{I}(r(s_j) = r(s)), \tag{11}$$

$$c_k^{(t)}(s) = \sum_{j=1}^{t} \mathbb{I}(k(s_j) = k(s)). \tag{12}$$

which count how many prior structures share the same element set or stoichiometry with the current structure, respectively.

The intra-sequence diversity reward is then:

$$r_{\text{intra}}(s) = \begin{cases} 0, & c_r^{(t)}(s) > K_{\text{intra}}^r \text{ or } c_k^{(t)}(s) > K_{\text{intra}}^k, \\ 1, & \text{otherwise}, \end{cases} \tag{13}$$

where K_{intra}^r and K_{intra}^k are hyperparameters controlling the tolerance for element set and stoichiometry repetition within a sequence, respectively. If either counter exceeds its threshold, the reward becomes 0.

Additionally, to prevent repeated generation of the same structure across multiple training steps (i.e., inter-sequence redundancy), we track the generation frequency over recent rollouts. We maintain separate counters for element sets and stoichiometry patterns:

$$f_r(s) = \sum_{\tau \in T} \sum_{s' \in G(\tau)} \mathbb{I}(r(s') = r(s)), \tag{14}$$

$$f_k(s) = \sum_{\tau \in T} \sum_{s' \in G(\tau)} \mathbb{I}(k(s') = k(s)), \tag{15}$$

where T represents recent rollouts and $G(\tau)$ denotes structures generated in rollout τ.

The inter-sequence diversity reward uses a three-tier system to balance exploration. We first determine the tier based on the maximum of the two counters:

$$f(s) = \max(f_r(s), f_k(s)), \tag{16}$$

then assign:

$$r_{\text{inter}}(s) = \begin{cases} 2, & f(s) \le K_{\text{inter}}^{\text{low}}, \\ 1, & K_{\text{inter}}^{\text{low}} < f(s) \le K_{\text{inter}}^{\text{high}}, \\ 0, & f(s) > K_{\text{inter}}^{\text{high}}, \end{cases} \tag{17}$$

where $K_{\text{inter}}^{\text{low}}$ and $K_{\text{inter}}^{\text{high}}$ define thresholds for rare and excessive generation. By using the maximum of both counters, we ensure that over-generation in either element set or stoichiometry dimension triggers the appropriate penalty. This design provides strong positive incentive (2) for discovering rarely-generated structure types, standard reward (1) for structures at normal frequency, and zero reward (0) for over-exploited structures.

Finally, we introduce a match-based reward that ensures structural uniqueness:

$$r_{\text{match}}(s_t) = \begin{cases} 0, & \exists j < t \text{ such that } M(s_t, s_j) = 1, \\ 1, & \text{otherwise}. \end{cases} \tag{18}$$

This reward immediately becomes 0 if any prior structure in the sequence matches the current one, preventing exact duplicates.

The overall diversity reward combines these three components:

$$r_{\text{div}}(s) = r_{\text{match}}(s) + r_{\text{intra}}(s) + r_{\text{inter}}(s), \tag{19}$$

where each component is evaluated sequentially. The matching reward $r_{\text{match}}(s) \in \{0, 1\}$ ensures structural uniqueness within the sequence, the intra-sequence reward $r_{\text{intra}}(s) \in \{0, 1\}$ prevents compositional redundancy, and the inter-sequence reward $r_{\text{inter}}(s) \in \{0, 1, 2\}$ modulates based on cross-sequence frequency. When a structure is duplicated within the sequence ($r_{\text{match}}(s) = 0$) or compositionally redundant ($r_{\text{intra}}(s) = 0$), the diversity reward is correspondingly reduced. When both local diversity criteria are satisfied, the final reward is determined by the inter-sequence frequency, ranging from 1 (normal frequency) to 2 (rare structures), yielding an overall reward range of 0 to 4 that effectively balances local uniqueness with global diversity exploration during training.

3.5. Multiple Objectives Based Material Generation Strategy optimization

The final reward function integrates validity, stability, novelty, and diversity components:

$$R(s) = \begin{cases} r_{\text{stab}}(s) \cdot r_{\text{nov}}(s) \cdot r_{\text{div}}(s), & r_{valid} = 1 \\ & \text{and } r_{\text{stab}}(s) > 0, \\ 0, & \text{otherwise}. \end{cases} \tag{20}$$

This composition enables exploration across the stable and metastable regions of chemical space, where $r_{\text{stab}}(s) \in [0, 1]$ captures thermodynamic stability, $r_{\text{nov}}(s) \in \{0, 1\}$ incentivizes database novelty, and $r_{\text{div}}(s) \in \{0, 1, 2, 3, 4\}$ promotes structural diversity.

We optimize the model using GRPO. For each prompt q, we sample a group of G outputs $\{o_i\}_{i=1}^G$ from the current policy π_θ. The advantage for each output is computed by normalizing its reward relative to the group distribution:

$$\hat{A}_i = \frac{R(o_i) - \mu_R}{\sigma_R + \epsilon}, \tag{21}$$

where μ_R and σ_R are the mean and standard deviation of rewards $\{R(o_j)\}_{j=1}^G$ within the group, and ϵ is a small constant for numerical stability. This formulation inherently centers the rewards around the group average while scaling by their spread, yielding stable and low-variance advantage estimates.

The GRPO objective is:

$$\mathcal{J}_{\text{GRPO}}(\theta) = \mathbb{E}\Big[q \sim P(Q),\ \{o_i\}_{i=1}^{G} \sim \pi_{\theta_{\text{old}}}(O|q)\Big] \frac{1}{G}\sum_{i=1}^{G}\frac{1}{|o_i|}\sum_{t=1}^{|o_i|}\Bigg\{\min\Bigg[\frac{\pi_\theta(o_{i,t}|q,o_{i,<t})}{\pi_{\theta_{\text{old}}}(o_{i,t}|q,o_{i,<t})}\hat{A}_{i,t}, \text{clip}\Bigg(\frac{\pi_\theta(o_{i,t}|q,o_{i,<t})}{\pi_{\theta_{\text{old}}}(o_{i,t}|q,o_{i,<t})}, 1-\varepsilon, 1+\varepsilon\Bigg)\hat{A}_{i,t}\Bigg] -\beta D_{\text{KL}}\big[\pi_\theta \,\|\, \pi_{\text{ref}}\big]\Bigg\}, \tag{22}$$

where ε is the clipping parameter, β controls KL regularization to the reference policy π_{ref}, and the expectation is over the prompt distribution $P(Q)$ and rollouts from the old policy. Through iterative updates, the model learns to generate structures that are physically valid, thermodynamically reasonable, novel, and diverse.

4. EXPERIMENTS

4.1. SETUP

We conduct all experiments on the widely used MP-20 dataset (Jain et al., 2013; Xie et al., 2021), which consists of 45,231 inorganic crystal structures curated from the Materials Project. All structures satisfy metastability criteria and contain no more than 20 atoms per unit cell. To efficiently adapt LLMs to the task of from-scratch crystal generation, we build our approach on the Qwen 2.5-7B Instruction (Qwen et al., 2025) and employ parameter-efficient fine-tuning via 4-bit quantization (Dettmers et al., 2023) together with low-rank adapters (Hu et al., 2022).

The entire training pipeline is implemented in PyTorch (Paszke et al., 2019), with model loading and training orchestration handled through the Hugging Face Transformers library (Wolf et al., 2019). Wyckoff positions and space group symmetry information for crystal structures are automatically generated using the PyXtal toolkit (Fredericks et al., 2021), following standard crystallographic conventions, to ensure physically consistent input representations. After supervised fine-tuning (SFT), we further perform online RL on the same instruction set, guiding the model toward generating energetically more stable crystal configurations. Detailed training hyperparameters, optimization settings, and sampling strategies are provided in Appendix A.

4.2. METRICS

To provide an initial assessment of the quality of the generated crystal structures, we follow (Xie et al., 2021) and report structural validity and compositional validity as auxiliary metrics, which reflect the model's ability to generate physically and chemically plausible crystals. As our primary evaluation criterion, we adopt the S.U.N. Rate introduced by (Zeni et al., 2024), which measures the fraction of generated structures that simultaneously satisfy stability, uniqueness, and novelty, and serves as the core indicator of overall generation performance.

Stability is determined by comparing the formation energy of each relaxed structure against the energy convex hull constructed by (Riebesell et al., 2025). Specifically, we follow the standard relaxation protocol of Matbench Discovery (Riebesell et al., 2025) and use `eSEN` to relax the structures generated during validation, with the Materials Project database serving as the reference convex hull for computing the energy above the hull (E_{hull}). All total energies are corrected using the MP2020 compatibility scheme to ensure consistency across different exchange-correlation functionals (e.g., DFT and DFT+U). Structures lying on or below the convex hull, i.e., with $E_{\text{hull}} \leq 0$ eV/atom, are considered stable. Uniqueness and novelty are both evaluated on relaxed structures using Pymatgen's `StructureMatcher` (Ong et al., 2013): uniqueness captures the model's ability to generate structurally diverse crystals by identifying distinct structures among all generated samples, while novelty quantifies the structural and chemical distinctiveness of generated samples relative to the training set (MP-20), thereby reflecting the model's capacity to explore previously unseen regions of crystal space. These three criteria are jointly computed on relaxed structures to obtain the S.U.N. metric, formally defined as:

$$\text{Stability Rate} = \frac{N_{\text{stable}}}{N_{\text{gen}}} \times 100\% \tag{23}$$

$$\text{SUN Rate} = \frac{N_{\text{SUN}}}{N_{\text{gen}}} \times 100\% \tag{24}$$

Here, N_{gen} denotes the total number of generated structures, N_{stable} is the count of thermodynamically stable ones ($E_{\text{hull}} \leq 0$ eV/atom), and N_{SUN} represents those that satisfy all three criteria: stability, uniqueness, and novelty.

During RL for unconditional from-scratch crystal generation, we primarily employ the eqV2 (Liao et al., 2023) machine-learned interatomic potential to evaluate structural stability. To ensure consistency with prior work(Xu et al., 2025a), different MLIPs are used at different stages of evaluation: specifically, `eSEN` (Fu et al., 2025) is adopted for S.U.N. computation during preference dataset construction and final result evaluation, following the standard practice of testing with the same potential as in previous studies.

4.3. RESULTS

As shown in Table 1, we systematically compare the proposed method with representative approaches for from-scratch crystal generation on the MP-20 dataset. Early

Table 1. Results of From-scratch Crystal Structures Generation on the MP-20 Dataset

Method	Validity (%)		Stability (%)	S.U.N. (%)
	Structural	Composition		
CDVAE (Xie et al., 2021)	99.9	86.9	3.6	3.5
DiffCSP (Jiao et al., 2023)	99.6	82.2	12.5	9.7
DiffCSP++ (Jiao et al., 2024)	99.9	85.8	13.2	9.1
FlowMM (Miller et al., 2024)	96.4	83.4	9.3	6.3
FlowLLM (Sriram et al., 2024)	99.9	90.8	13.9	4.7
SymmCD (Levy et al., 2025)	94.3	85.9	9.4	7.0
CrystalLLM-70B (Gruver et al., 2024)	99.6	95.4	5.3	-
Jointly-trained ADIT (Joshi et al., 2025)	99.7	92.1	15.4	5.3
PLaID (Xu et al., 2025b)	99.6	93.0	18.9	10.3
PLaID++ (Xu et al., 2025a)	99.7	95.4	22.5	12.5
CRYSTAL	99.9	98.4	58.4	25.1

Table 2. Ablation Study of Reward Configurations

Reward Configuration	S.U.N. (%)
Without Uniqueness Reward	12.1
Without Novelty Reward	7.3
Without Stability Rewards	9.8
All Rewards	25.1

generative models such as CDVAE and DiffCSP achieve relatively high validity but are severely limited in thermodynamic stability, leading to consistently low S.U.N. rates. Subsequent diffusion-based methods incorporating symmetry constraints achieve moderate stability improvements, yet their overall generation quality remains constrained by unresolved multi-objective trade-offs. LLM-based approaches demonstrate clear advantages in modeling discrete crystal representations. PLaID and PLAID++, which incorporate DPO-based and type-enhanced DPO-based preference optimization, achieves improved stability and S.U.N. performance. The reported results for PLaID and PLAID++ are obtained by repeatedly sampling from the publicly released generated outputs of the PLAID++ model and recomputing the metrics, with the highest values reported. The pairwise preference framework of DPO inherently struggles to coordinate all three objectives simultaneously—while stability and novelty can be formulated as pairwise preferences, uniqueness is a batch-level metric that cannot be naturally incorporated into pairwise comparisons, limiting joint optimization within the DPO paradigm.

In contrast, CRYSTAL consistently outperforms all baselines across all metrics, achieving substantially higher stability and S.U.N. By leveraging GRPO-based multi-objective reward with multiplicative aggregation, CRYSTAL enables coordinated optimization of stability, novelty, and uniqueness, effectively mitigating reward conflict in multi-objective RL and significantly enhancing the quality of from-scratch crystal generation. Furthermore, in the appendix G, we present three novel generated structures visualized using VESTA.

5. Analysis

5.1. Ablation Study

To further evaluate the role of the proposed GRPO-based Multi-Objective reward framework in addressing reward conflict and reward hacking in multi-objective RL, we perform a comprehensive ablation study across different reward configurations. Specifically, we systematically ablate the individual reward components—uniqueness, novelty, and stability—by removing each constraint from the full reward set (ALL), and compare their impact using the S.U.N. metric.

The results show that optimizing a single constraint in iso-

Table 3. Experiments with Numbers of Rollouts

Rollout	Stability (%)	Novelty (%)	Uniqueness (%)	S.U.N. (%)
16	70.2	95.0	18.3	12.2
32	67.9	96.8	20.3	13.0
64	62.5	93.5	27.1	16.4
128	58.4	96.3	46.7	25.1

lation is insufficient for coordinated multi-objective crystal generation. Models trained with individual rewards exhibit strong objective bias, achieving S.U.N. success rates of approximately 12.1%, 7.3%, and 9.8%, respectively, reflecting limited overall performance. In contrast, jointly optimizing all verifiable rewards yields a substantial improvement, increasing the S.U.N. success rate to 25.1%. Importantly, the full model maintains novelty above 96.3% and simultaneously achieves a balanced compromise, with stability reaching roughly 58.4% and uniqueness about 46.7%. These results clearly demonstrate that Multi-Objective rewards are crucial within the GRPO framework for mitigating reward bias in multi-objective RL. Without their joint contribution, the model tends to favor specific objectives; in contrast, coordinated updates across the reward set enable the simultaneous optimization of stability, novelty, and uniqueness, thereby substantially enhancing the overall quality and practical utility of from-scratch crystal generation.

5.2. Experiments with Numbers of Rollouts

We investigate the impact of rollout size on from-scratch crystal generation under the GRPO framework (Table 3). In this setting, crystal generation is formulated as a single-step decision process, and each generation step is treated as one rollout. As the rollout size decreases from 128 to 16, the stability metric consistently increases (from 58.4% to 70.2%), while novelty remains largely stable. In contrast, uniqueness drops sharply (from 46.7% to 18.3%), leading to a corresponding degradation in the S.U.N. success rate.

This behavior indicates that when the group size is insufficient, the hierarchical reward structure fails to impose effective joint constraints on policy updates. As a result, optimization becomes dominated by local objectives such as stability, giving rise to policy drift and reward hijacking. Although high stability is preserved, reduced uniqueness and weakened multi-objective coordination ultimately limit overall S.U.N. performance. By contrast, larger rollout sizes enable richer intra-group comparisons, maintaining a balanced interplay among stability, novelty, and uniqueness, and thereby substantially improving the quality and practical utility of from-scratch crystal generation.

Conclusion

We propose a Multi-Objective reward framework based on GRPO for from-scratch generation of crystalline materials that simultaneously satisfy stability, novelty, diversity, and correctness. Our approach decomposes multi-objective optimization into independently verifiable sub-rewards and leverages intra-group relative policy updates to mitigate reward hijacking and imbalance issues. Experimental results demonstrate that CRYSTAL significantly outperforms existing methods across S.U.N. success rate, stability, novelty, and uniqueness metrics, validating the efficacy of multi-objective RL for crystal generation and providing a robust pathway for on-demand materials design.

Acknowledgements

The AI-driven experiments, simulations, and model training were performed on the robotic AI-Scientist platform of the Chinese Academy of Sciences. The authors gratefully acknowledge financial support from the National Key R&D Program of China (2025YFF0516300), the National Natural Science Foundation of China (NSFC) (52541102, 22133005, U25A20229, 22403102), the Science and Technology Commission of Shanghai Municipality (LJ2024049, 23ZR1472600, 25CL2902100), the Shanghai Municipal Commission of Economy and Informatization (2024-GZL-RGZN-01026), the Youth Innovation Promotion Association of CAS (2022251), the Shanghai Sailing Program (23YF1454900, 24YF2753300), and the Open Research Fund of Suzhou Laboratory (No. SZLAB-1508-2024-ZD018).

References

Betala, S., Gleason, S. P., Ramlaoui, A., Xu, A., Channing, G., Levy, D., Fourrier, C., Kazeev, N., Joshi, C. K., Kaba, S.-O., Therrien, F., Hernandez-Garcia, A., Mercado, R., Krishnan, N. M. A., and Duval, A. Lemat-genbench: A unified evaluation framework for crystal generative mod-

els, 2026. URL https://arxiv.org/abs/2512.04562.

Cao, Z. and Wang, L. Reinforcement fine-tuning for materials design. *Physical Review B*, 113(2), January 2026. ISSN 2469-9969. doi: 10.1103/45zh-44bg. URL http://dx.doi.org/10.1103/45zh-44bg.

Chen, C. and Ong, S. P. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

Chen, J., Guo, J., Fako, E., and Schwaller, P. Accelerating inverse materials design using generative diffusion models with reinforcement learning, 2025. URL https://arxiv.org/abs/2511.03112.

Davies, D. W., Butler, K. T., Jackson, A. J., Morris, A., Frost, J. M., Skelton, J. M., and Walsh, A. Computational screening of all stoichiometric inorganic materials. *Chem*, 1(4):617–627, 2016.

Dettmers, T., Pagnoni, A., Holtzman, A., and Zettlemoyer, L. Qlora: Efficient finetuning of quantized llms. *Advances in neural information processing systems*, 36:10088–10115, 2023.

Fredericks, S., Parrish, K., Sayre, D., and Zhu, Q. Pyxtal: A python library for crystal structure generation and symmetry analysis. *Computer Physics Communications*, 261: 107810, 2021.

Fu, X., Wood, B. M., Barroso-Luque, L., Levine, D. S., Gao, M., Dzamba, M., and Zitnick, C. L. Learning smooth and expressive interatomic potentials for physical property prediction. *arXiv preprint arXiv:2502.12147*, 2025.

Gruver, N., Sriram, A., Madotto, A., Wilson, A. G., Zitnick, C. L., and Ulissi, Z. Fine-tuned language models generate stable inorganic materials as text. *arXiv preprint arXiv:2402.04379*, 2024.

Guo, G., Saidi, T. L., Terban, M. W., Valsecchi, M., Billinge, S. J., and Lipson, H. Ab initio structure solutions from nanocrystalline powder diffraction data via diffusion models. *Nature Materials*, pp. 1–9, 2025.

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., and Persson, K. A. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 07 2013. ISSN 2166-532X. doi: 10.1063/1.4812323. URL https://doi.org/10.1063/1.4812323.

Jiao, R., Huang, W., Lin, P., Han, J., Chen, P., Lu, Y., and Liu, Y. Crystal structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing Systems*, 36:17464–17497, 2023.

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

Joshi, C. K., Fu, X., Liao, Y.-L., Gharakhanyan, V., Miller, B. K., Sriram, A., and Ulissi, Z. W. All-atom diffusion transformers: Unified generative modelling of molecules and materials. *arXiv preprint arXiv:2503.03965*, 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. *arXiv preprint arXiv:2502.03638*, 2025.

Liao, Y.-L., Wood, B., Das, A., and Smidt, T. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. *arXiv preprint arXiv:2306.12059*, 2023.

Meta Fundamental AI Research. fairchem-data-omat, 2024. URL https://pypi.org/project/fairchem-data-omat/.

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., and Sparks, T. D. Crystext: A generative ai approach for text-conditioned crystal structure generation using llm. *Integrating Materials and Manufacturing Innovation*, 2026. doi: 10.1007/s40192-026-00451-8.

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

Park, H. and Walsh, A. Guiding generative models to uncover diverse and novel crystals via reinforcement learning, 2025. URL https://arxiv.org/abs/2511.07158.

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

Qwen, :, Yang, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Li, C., Liu, D., Huang, F., Wei, H., Lin, H.,

Yang, J., Tu, J., Zhang, J., Yang, J., Yang, J., Zhou, J., Lin, J., Dang, K., Lu, K., Bao, K., Yang, K., Yu, L., Li, M., Xue, M., Zhang, P., Zhu, Q., Men, R., Lin, R., Li, T., Tang, T., Xia, T., Ren, X., Ren, X., Fan, Y., Su, Y., Zhang, Y., Wan, Y., Liu, Y., Cui, Z., Zhang, Z., and Qiu, Z. Qwen2.5 technical report, 2025. URL https://arxiv.org/abs/2412.15115.

Riebesell, J., Goodall, R. E., Benner, P., Chiang, Y., Deng, B., Ceder, G., Asta, M., Lee, A. A., Jain, A., and Persson, K. A. A framework to evaluate machine learning crystal stability predictions. *Nature Machine Intelligence*, 7(6): 836–847, 2025.

Saal, J. E., Kirklin, S., Aykol, M., Meredig, B., and Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *Jom*, 65(11):1501–1509, 2013.

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.

Sriram, A., Miller, B. K., Chen, R. T., and Wood, B. M. Flowllm: Flow matching for material generation with large language models as base distributions. *Advances in Neural Information Processing Systems*, 37:46025–46046, 2024.

von Werra, L., Belkada, Y., Tunstall, L., Beeching, E., Thrush, T., Lambert, N., Huang, S., Rasul, K., and Gallouédec, Q. TRL: Transformers Reinforcement Learning, 2020. URL https://github.com/huggingface/trl.

Ward, L., Dunn, A., Faghaninia, A., Zimmermann, N. E., Bajaj, S., Wang, Q., Montoya, J., Chen, J., Bystrom, K., Dylla, M., et al. Matminer: An open source toolkit for materials data mining. *Computational Materials Science*, 152:60–69, 2018.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*, 2019.

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*, 2025a.

Xu, A., Desai, R., Wang, L., Hope, G., and Ritz, E. T. PLaID: Preference aligned language model for targeted inorganic materials design. In *AI for Accelerated Materials Design - ICLR 2025*, 2025b. URL https://openreview.net/forum?id=7aoP3ZeBfy.

Zeni, C., Pinsler, R., Zügner, D., Fowler, A., Horton, M., Fu, X., Shysheya, S., Crabbé, J., Sun, L., Smith, J., Nguyen, B., Schulz, H., Lewis, S., Huang, C.-W., Lu, Z., Zhou, Y., Yang, H., Hao, H., Li, J., Tomioka, R., and Xie, T. Mattergen: a generative model for inorganic materials design, 2024. URL https://arxiv.org/abs/2312.03687.

Zeni, C., Pinsler, R., Zügner, D., Fowler, A., Horton, M., Fu, X., Wang, Z., Shysheya, A., Crabbé, J., Ueda, S., et al. A generative model for inorganic materials design. *Nature*, 639(8055):624–632, 2025.

Table 4. Comparison of novelty evaluation strategies (OR vs. AND) for CRYSTAL on the MP-20 dataset. Reported values are percentages (%).

STRAT.	COM. NOV FIG	STR. NOV FIG	ALL NOV FIG
OR	50.2	10.3	53.2
AND	89.2	52.1	92.6

Table 5. Impact of structure comparison methods on evaluation metrics for CRYSTAL on the MP-20 dataset. Reported values are percentages (%).

METHOD	NOVELTY	UNIQUENESS
FINGERPRINT	92.6	32.3
STRUCTUREMATCHER	96.3	46.7

A. Hyperparameter Setting

We fine-tune the model using the AdamW optimizer with a batch size of 16 and a learning rate of 10^{-5}, employing fp4 mixed-precision training. Parameter-efficient fine-tuning is performed via LoRA adapters, configured with a rank of 8, an alpha scaling factor of 32, and a dropout rate of 0.05. The training runs for 10 epochs on the MP-20 dataset.

Following supervised fine-tuning, we apply Group Relative Policy Optimization (GRPO) to align the model's policy with desired structural and compositional preferences. We implement GRPO using the TRL library (von Werra et al., 2020), with an Adam optimizer, a batch size of 16, and a learning rate of 5×10^{-6}. Training is performed in 4-bit precision (fp4/bfloat16) for one epoch on our dataset. The GRPO configuration includes a rollout size of 128, preference loss coefficients $\beta_{\text{struc}} = 0.1$ and $\beta_{\text{comp}} = 2.0$ for structural and compositional rewards, All other hyperparameters are set to their default values, with a decoding temperature of 1.0. Meanwhile, $K^{r}_{\text{intra}} = K^{k}_{\text{intra}} = 3$, and $K^{\text{high}}_{\text{inter}} = 5$, $K^{\text{low}}_{\text{inter}} = \lceil K^{\text{high}}_{\text{inter}}/2 \rceil = 3$.

B. Comparison of Reward Function Based on AND OR Judgments

In this section, we first evaluate the impact of two different novelty reward strategies (OR and AND) on model training effectiveness, and then discuss the influence of different structure comparison methods on evaluation results.

Novelty is assessed using the fingerprint-based method proposed by CrystaLLM (**?**), evaluated along two dimensions: compositional novelty and structural novelty. Under the OR strategy, a generated structure receives the novelty reward if it is novel in either composition or structure; the AND strategy requires novelty in both dimensions simultaneously. As shown in Table 4, under the OR strategy, the low threshold for obtaining rewards results in insufficient exploration incentive, with both compositional novelty and structural novelty remaining at low levels (50.2% and 10.3%, respectively). The AND strategy significantly raises the exploration threshold by requiring simultaneous novelty in both dimensions, driving the model to explore more deeply in both composition and structure: compositional novelty increases substantially from 50.2% to 89.2%, structural novelty from 10.3% to 52.1%, overall novelty from 53.2% to 92.6%. This demonstrates that stricter joint constraints not only improve structural novelty but also simultaneously promote compositional novelty, with the two reinforcing rather than competing with each other.

C. Impact of Structure Comparison Methods (Fingerprint vs. StructureMatcher)

The choice of structure comparison method directly affects the evaluation of uniqueness and novelty as shown in Table 5. The fingerprint-based method adopted in this work determines structural equivalence by computing distances between structure fingerprints (CrystalNNFingerprint) and compositional fingerprints (ElementProperty), exhibiting higher sensitivity to subtle structural differences and thus applying a relatively stricter criterion. In comparison, Pymatgen's StructureMatcher performs atom-level structure matching within specified tolerances for lattice parameters, atomic positions, and angles, applying a relatively more lenient criterion. Under StructureMatcher evaluation, novelty increases from 50.0% to 70.2% and uniqueness from 92.4% to 96.3%, reflecting the inherent difference in judgment granularity between the two methods. Since all baseline methods in the main table are evaluated using the fingerprint-based method, we adopt it as the primary reporting standard to ensure fair comparability. Notably, CRYSTAL significantly outperforms all baseline methods under both comparison methods, validating the robustness of the results.

Table 6. Impact of intra-sequence diversity threshold K_{intra} on S.U.N. rate (%). $K_{\text{inter}}^{\text{high}}$ is fixed at 5.

$K_{\text{INTRA}}^{r} = K_{\text{INTRA}}^{k}$	1	3	5
S.U.N.	24.8	25.1	24.6

Table 7. Impact of inter-sequence diversity threshold K_{inter} on S.U.N. rate (%). $K_{\text{intra}}^{r} = K_{\text{intra}}^{k}$ is fixed at 3.

$(K_{\text{INTER}}^{\text{LOW}}, K_{\text{INTER}}^{\text{HIGH}})$	(1, 1)	(2, 3)	(3, 5)
S.U.N.	22.3	23.7	25.1

D. Impact of Intra-sequence Diversity Threshold

To examine the effect of intra-sequence diversity constraint strength on generation quality, we fix $K_{\text{inter}}^{\text{high}} = 5$ and vary $K_{\text{intra}}^{r} = K_{\text{intra}}^{k}$ across 1, 3, and 5. As shown in Table 6, the results indicate that overly strict intra-sequence constraints excessively limit generation freedom—within a single rollout of 128 structures, the same composition is only allowed to appear once, causing many reasonable structural variants to receive zero reward, which in turn disrupts learning. Moderately relaxing the threshold allows the model to sample multiple times within the same compositional space, facilitating the discovery of different stable structural configurations under the same elemental system.

E. Impact of Inter-sequence Diversity Threshold

To examine the effect of cross-rollout diversity constraints on long-term exploration capability, we fix $K_{\text{intra}}^{r} = K_{\text{intra}}^{k} = 3$ and vary $K_{\text{inter}}^{\text{high}}$ across 1, 3, and 5, with $K_{\text{inter}}^{\text{low}} = \lceil K_{\text{inter}}^{\text{high}}/2 \rceil$ correspondingly set to 1, 2, and 3. As shown in Table 7, the results indicate that smaller thresholds cover a limited historical range, where even a small number of structural repetitions triggers penalties, causing the model to prematurely abandon exploration of promising regions of chemical space. Larger thresholds provide more sufficient tolerance, enabling the model to perform multiple samplings and in-depth exploration of valuable structure types, while still effectively suppressing genuine over-generation.

F. Evaluation under Corrected Phase Diagram

F.0.1. Evaluation under Corrected Phase Diagram

To further validate the robustness of our results, we evaluate all three LLM-based RL methods under the stricter OMAT24-corrected convex hull using eSEN (Fu et al., 2025), with energy corrections obtained from fairchem-data-omat(Meta Fundamental AI Research, 2024). This setup adopts a self-consistent hull rather than mixing different reference systems, consistent with the design philosophy of LeMat-GenBench (Betala et al., 2026). We verified the reliability of this pipeline on DFT-labeled data, achieving an F1 score of 0.9033 (Precision = 0.8958, Recall = 0.9110), confirming DFT-level stability classification accuracy. As shown in Table 8, under this more stringent evaluation standard, the stability rates of all methods decrease compared to the uncorrected setting, confirming that the corrected phase diagram applies stricter criteria. Nevertheless, CRYSTAL maintains a substantial advantage over both PLaID and PLaID++ in terms of stability and S.U.N., confirming that the performance improvement is robust across different evaluation standards and is not an artifact of a particular choice of reference phase diagram.

G. The presentation of the new crystal structure

Table 8. Comparison under OMAT24-corrected convex hull (%).

METHOD	STABILITY	S.U.N.
PLAID	21.8	12.7
PLAID++	20.2	13.4
CRYSTAL	52.8	23.6

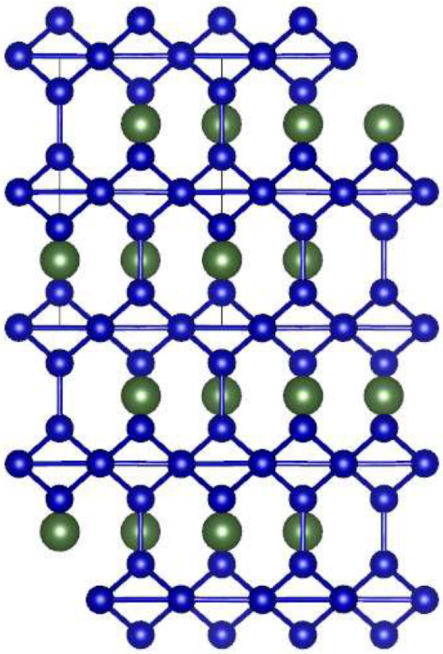

Figure 2. The crystal structure of $ErCo_4$

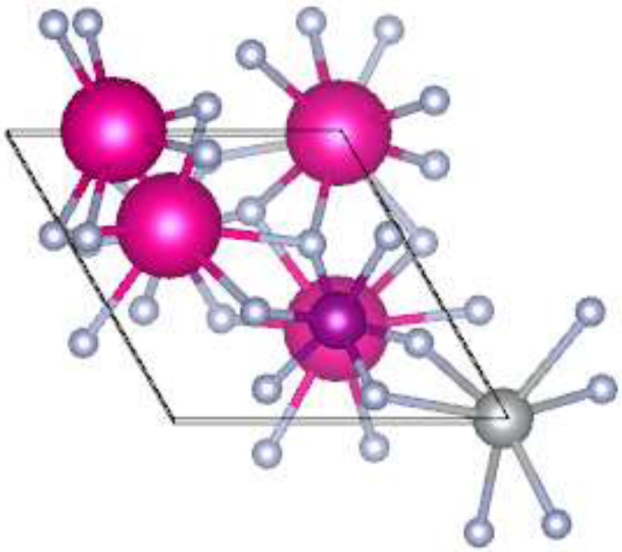

Figure 3. The crystal structure of Rb_2MnAgF_6

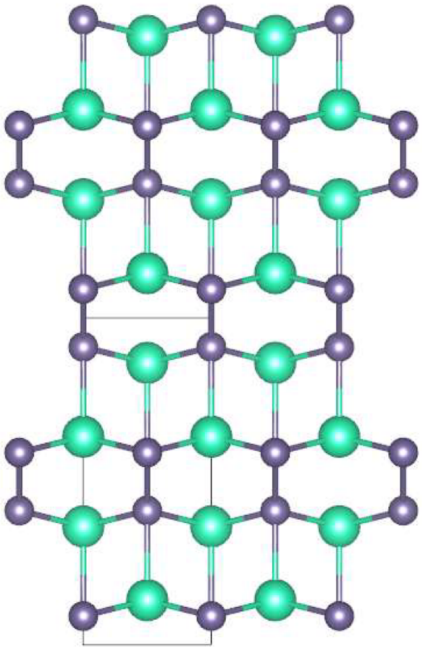

Figure 4. The crystal structure of LuGe