Chinese Physics Letters

LETTER

NextCrystal: A Symmetry-Driven Generative Framework for Crystal Structure Prediction

Cite as: Jinming Mu *et al*. 2026 *Chin. Phys. Lett.* **43** 100802

Recommended Articles

Predicting Many Crystal Properties via an Adaptive Transformer-Based Framework
Haosheng Xu, Dongheng Qian, and Jing Wang
Chin. Phys. Lett. **42**, 070706 (2025)

AI-Driven Inverse Design of Materials: Past, Present, and Future
Xiao-Qi Han, Xin-De Wang, Meng-Yuan Xu, Zhen Feng, Bo-Wen Yao, Peng-Jie Guo, Ze-Feng Gao, and Zhong-Yi Lu
Chin. Phys. Lett. **42**, 027403 (2025)

Machine Learning Prediction of Crystal Structure Stability toward the Design of High-Entropy Oxides
Qiancheng Zhou, Chunchun Wang, Liyuan Wang, Zhouzhou Wang, Ming Qiu, Mingdong Dong, and Ying Yu
Chin. Phys. Lett. **43**, 050801 (2026)

Learning-at-Criticality in Large Language Models for Quantum Field Theory and Beyond
Xiansheng Cai, Sihan Hu, Tao Wang, Yuan Huang, Pan Zhang, Youjin Deng, and Kun Chen
Chin. Phys. Lett. **42**, 120002 (2025)

High-Throughput Screening $MgFe_6Ge_6$-Prototype Bilayer Kagome Family
Lan-Ting Shi, Jian-Guo Si, Hui-Bin Zhou, Xian-Biao Shi, Peng-Fei Liu, Le-De Xian, Ping Zhang, and Bao-Tian Wang
Chin. Phys. Lett. **43**, 030711 (2026)

Ambient-Stable Polymeric Nitrogen Achieved through Multi-Stage Computational Design
Jiani Lin, Jianfu Li, Yong Liu, Jianan Yuan, and Xiaoli Wang
Chin. Phys. Lett. **42**, 070708 (2025)

A series journal of Chinese Physical Society distributed by IOP Publishing

Chinese Physics Letters

Chin. Phys. Lett. **43** 100802 (2026)

LETTER

DOI: 10.1088/0256-307X/43/10/100802
CSTR: 32039.14.0256-307X.43.10.100802

NextCrystal: A Symmetry-Driven Generative Framework for Crystal Structure Prediction

Jinming Mu(穆锦铭)[1], Lixin He(何力新)[1,2,*], Xudong Zhu(朱旭东)[2,*], and Shi Yin(殷实)[2,*]

[1]Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China
[2]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China

[*]Corresponding authors. Email: helx@ustc.edu.cn; zhu@iai.ustc.edu.cn; shiyin@iai.ustc.edu.cn

Crystal structure prediction (CSP), which aims to predict the 3D atomic arrangement of a crystal from its composition, is central to materials discovery and mechanistic understanding. Crystal symmetry plays a crucial role in CSP, but given the composition in a unit cell, existing methods either struggle with the NP-hard combinatorial challenge of enforcing symmetry rigorously or rely on retrieving known templates, inherently limiting both physical fidelity and the discovery of genuinely new materials. To address this challenge, we introduce NextCrystal, a symmetry-driven generative framework that employs language models to encode chemical semantics and directly generate fine-grained Wyckoff site patterns from atomic stoichiometry, without retrieving pre-existing structural templates at inference. To overcome the combinatorial complexity of site assignments, we incorporate domain knowledge via an efficient, linear-complexity heuristic beam search, rigorously enforcing algebraic consistency between site multiplicities and atomic stoichiometry. By integrating this symmetry-consistent template into a diffusion backbone, the framework constrains the stochastic generative trajectory to a physically plausible geometric manifold. NextCrystal achieves state-of-the-art performance on stability, uniqueness, and novelty benchmarks, as well as superior structural matching, establishing a paradigm for exploring previously unexplored crystallographic space without relying on prior structural templates. As a blind-rediscovery case study, NextCrystal recovers the known low-energy antipolar oV *Pnma* polymorph of HfO_2, which is absent from the training dataset, without retrieving a structural prototype during inference. Standardized structure matching after a common first-principles relaxation confirms its equivalence to the published oV/a62 structure and its distinction from the conventional high-pressure cotunnite/oII *Pnma* phase.

Crystalline materials are fundamental to modern technologies spanning energy, electronics, medicine, and aerospace, and advancing these sectors relies on identifying novel crystal structures with tailored properties [1−5]. The vastness of the compositional and structural parameter space presents a fundamental bottleneck for materials discovery. Conventional pipelines, guided by chemical intuition and constrained by costly computations and experiments [6,7], are intrinsically inefficient at exploring this space at scale.

In recent years, artificial intelligence (AI), and deep learning-based generative models in particular, have emerged as a promising route to proposing candidate materials at scale. They demonstrate superior generalization capabilities and significant advantages in computational efficiency [8−14]. Within this landscape, the crystal structure prediction (CSP) task [8,10,13], which seeks the three-dimensional arrangement of atoms in the unit cell from a given composition, plays a pivotal role in AI-driven materials design workflows [11]. Diffusion models have become the dominant architecture for 3D structure generation [15,16], and, for the CSP task, diffusion-based approaches also deliver state-of-the-art (SOTA) performance [10,14].

However, the CSP task is fundamentally governed by stringent crystallographic rules, imposing rigid constraints on the design methodology. Physically plausible crystal structures typically obey the symmetry constraints imposed by specific space groups, which strictly constrain the allowed Wyckoff positions, determining atomic multiplicities, site symmetries, and the relative arrangement of atoms within the unit cell. Only when the given chemical composition is mathematically consistent with the corresponding Wyckoff positions can physically reasonable crystal structures be obtained.

However, rigorously enforcing these physical constraints imposed by space groups remains a universal bottleneck across the entire CSP domain. Determining an exact combination of Wyckoff positions whose multiplicities precisely match the atom counts of a given unit-cell composition constitutes a combinatorial optimization problem, and its generalized multiplicity-constrained formulation is NP-hard, as formally established in Section A.3 of the Supplementary Materials (SM). Existing methods struggle to efficiently resolve this task, often requiring partial compromises. For instance, Xin *et al.* [17] proposed a pruning-based search for Wyckoff positions to address site assignments. Nevertheless, the approach remains fundamentally a brute-force enumeration. As the number of atoms in a unit cell increases, the possible partitions of atoms into available Wyckoff sites grow exponentially. For large systems, exhaustive enumeration becomes computationally prohibitive. Moreover, lacking a likelihood-based ranking mechanism, such methods cannot efficiently identify the most physically plausible candidates from the massive, unguided pool of combinations.

Within the landscape of generative AI, existing methods only partially address the discrete combinatorial challenge of site assignments. Early 3D generative models, such as CDVAE [18] and DiffCSP [19], bypass this challenge by operating directly on continuous coordinates. These models rely on neural networks to implicitly learn crystallographic symmetries, leaving the generative trajectory unconstrained in a high-dimensional search space. CrystaLLM [9] adopts a text-generation approach, autoregressively predicting Crystallographic Information File (CIF) formats; however, any generated symmetry or Wyckoff symbols are purely probabilistic textual outputs, lacking explicit algebraic or physical constraints. Some recent methods condition the generator on a space group label [11,14], but they do not explicitly represent or enforce Wyckoff positions, capturing symmetry only in a coarse, global sense. More recent explicit-symmetry generators include SymmCD [20], which jointly generates asymmetric units and crystallographic symmetry transformations, and WyckoffDiff [21],

Received 6 July 2026; accepted 11 September 2026; published 20 October 2026
© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Check for updates

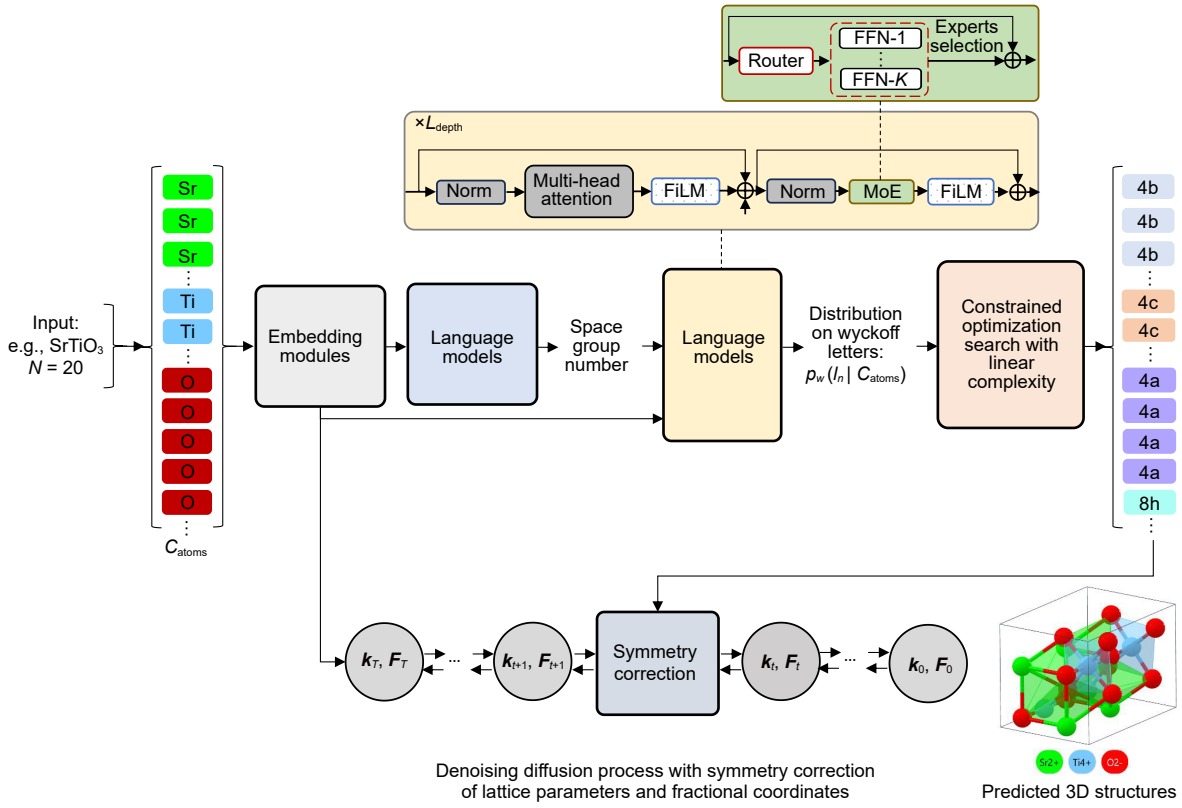

Fig. 1. Schematic of the proposed symmetry-driven generative framework. The Symmetry Correction block rectifies the lattice and fractional coordinates during denoising to enforce the crystal-family and Wyckoff-orbit constraints.

which generates symmetry-based crystal representations through discrete diffusion. Their native formulations target de-novo or unconditional crystal generation rather than exact-composition CSP, whereas NextCrystal predicts structures for a prescribed composition and can additionally operate under a specified space-group constraint.

Moving beyond coarse global symmetry, several recent AI methods attempt to explicitly utilize and strictly enforce Wyckoff positions, yet they remain hindered by the intractable complexity of exact site assignments. For composition-conditioned CSP, these approaches can be broadly organized into soft-penalty/post-filtering, Wyckoff-sequence generation without stoichiometric hard constraints, and retrieval-based template methods.

Soft-Penalty Methods

While WyCryst [12] guides generation via Wyckoff-based latent representations, its soft-penalty sampling cannot guarantee strict stoichiometric alignment. Consequently, extracting 3D structures consistent with Wyckoff positions relies on post-hoc filtering, a process that can become computationally prohibitive given the NP-hard nature of combinatorial site assignments.

Wyckoff-Sequence Generation Without Stoichiometric Hard Constraints

CrystalFormer [22] presents a compelling framework representing crystals as Wyckoff sequences with logit masking for symmetry enforcement. While effective for generating structures within a given space group number, the final stoichiometry is an emergent outcome of the sampling process rather than a hard constraint. Since the model focuses on generative modeling, it does not employ an explicit constrained optimization mechanism to align site multiplicities with a given composition and number of atoms in a cell.

Retrieval-Based Template Methods

DiffCSP++ [10] goes further by enforcing strict space group symmetry defined by Wyckoff templates during the 3D material generation process. It employs two strategies for obtaining these templates: the first directly utilizes the ground-truth (GT) Wyckoff templates of the target material to guide structure generation. However, in practical CSP scenarios, these GT templates are inherently unknown a priori. As an alternative, it relies on a retrieval-based approach, which assumes that a suitable space group and Wyckoff template are already known in the given lookup database. In practice, these templates are retrieved from existing structures via metric learning methods, such as CSPML [23]. However, restricting generation to retrieved templates confines the model to known symmetry patterns, thereby limiting its ability to discover genuinely new space-group and Wyckoff configurations for a given composition.

To address these fundamental limitations, we propose an efficient symmetry-driven generative framework, named NextCrystal, which enables data-driven prediction of fine-grained Wyckoff site assignments directly from composition and atom counts, without retrieving a pre-existing structural template at inference, as illustrated in Fig. 1. Building upon the probabilistic distributions inferred by domain-specific language models (LMs), we explicitly overcome the intractable exponential complexity of combinatorial site assignments discussed above by designing an efficient heuristic constrained-optimization search. This algorithm drastically compresses the time complexity from an exponential scale down to a linear scale. This renders the search computationally tractable while rigorously enforcing algebraic consistency between site multiplicities and atomic stoichiometry without relying on computationally prohibitive brute-force enumeration. Crucially, the predicted symmetries are then imposed as hard geometric constraints to guide and modulate the diffusion process for three-dimensional crystal geometries. This rectifies the denoising trajectory, confines the generative process to a physically valid manifold, and preserves both structural plausibility and computational precision.

Below, we outline the core methodology of NextCrystal in this section. For full implementation details, as well as training and testing configurations, please refer to Sections A and B of the SM, respectively.

First, we use two LMs to directly infer fine-grained crystallographic symmetry from compositional inputs. These LMs are domain-specific token-sequence models operating on composition-derived element tokens, rather than general-purpose natural-language foundation models. Both predictors are trained on existing crystal-structure data and retrieval-free refers only to inference, during which no database structure or pre-existing Wyckoff template is queried. Both models are built upon the Transformer architecture [24], and inspired by Soft MoE [25], we replace the standard feed-forward networks (FFNs) within each Transformer block with sparse Top-k Mixture-of-Experts (MoE) layers to enhance parameter capacity without a proportional increase in active matrix computation. This design enables multiple experts to specialize in distinct compositional patterns, with only the Top-2 experts requested for each token. The routing implementation, including capacity handling, router noise, and load balancing, is detailed in Section A.2 of the SM, while standard-, requested-active-compute-, and parameter-matched FFN comparisons, the FiLM/concatenation ablation, and the associated compute analyses are provided in Section G of the SM.

Under matched parameter capacity, the sparse Top-2 MoE yields statistically comparable prediction quality to the dense FFN while substantially reducing active matrix computation and peak memory.

The first LM, denoted as LM_g, takes as input an atomic sequence representation of the composition, expanded to explicitly reflect the atom count N in the unit cell, and outputs a probability distribution over the 230 crystallographic space groups:

$$S_g = \arg\max_i p_g(i|C_{\text{atoms}}),\ \ p_g(i|C_{\text{atoms}}) = LM_g(C_{\text{atoms}}),\ \ 1 \leqslant i \leqslant 230, \quad (1)$$

where $C_{\text{atoms}} = [A, A, A, \ldots, B, B, B, B, B, \ldots, C, C, \ldots]$ is the expanded atomic sequence for elements $A, B, C, \ldots$ derived from the given composition, proportionally expanded to match the number of atoms in the unit cell; $p_g(i|C_{\text{atoms}})$ is the predicted probability that the crystal belongs to the i-th space group; S_g represents the predicted space group.

The second LM, denoted as LM_w, takes the same expanded atomic sequence together with the predicted space group S_g and outputs a probability distribution over Wyckoff letters compatible with that space group for each atom. Rather than concatenating S_g naively to the input, we encode the predicted space group into a continuous embedding and inject it into each transformer block via the feature-wise linear modulation (FiLM) [26] module. This FiLM-based conditioning ensures that Wyckoff letter predictions are explicitly modulated by the inferred space group, rather than being learned in an implicitly entangled fashion. Concretely, FiLM(S_g) produces per-channel scaling and shifting coefficients that modulate the intermediate representations of LM_w, allowing the model to adapt its representation to the symmetry class:

$$p_w(l_n = l \mid C_{\text{atoms}}) = LM_w\left(C_{\text{atoms}}, \text{FiLM}(S_g)\right),\ l_n \in L(S_g),\ 1 \leqslant n \leqslant N, \quad (2)$$

where $p_w(l_n = l|C_{\text{atoms}})$ denotes the probability that the n-th atom is assigned to letter l, and $L(S_g)$ is the set of admissible Wyckoff letters for space group S_g. Specific details regarding the network architectures and hyperparameter configurations of these two LMs are provided in Sections A.1, A.2, and B of the SM.

Second, we formulate the Wyckoff letter assignment for each atom as a constrained optimization problem based on the probability distribution $p_w(l_n|C_{\text{atoms}})$ predicted by the LM, and propose an efficient algorithm to solve it. This formulation strictly adheres to the combinatorial constraints imposed by the Wyckoff site multiplicity rules under the inferred space group symmetry. Specifically, the constrained maximization objective is defined as follows:

$$\begin{aligned}
\{l_n^*\}_{n=1}^N = \arg\max_{\{l_n\}_{n=1}^N} & \sum_{n=1}^N \log p_w(l_n|C_{\text{atoms}}), \\
\text{s.t. } & l_n \in L(S_g), \\
& \text{Count}_e(l) \bmod \text{mult}(l) = 0, \quad \forall l \in L(S_g), \\
& \sum_{l \in L(S_g)} \text{Count}_e(l) = N_e, \\
& \sum_{e \in \mathcal{E}} \frac{\text{Count}_e(l)}{\text{mult}(l)} \leqslant 1, \quad \forall l \in L_{\text{fix}}(S_g), \\
& \forall e \in \mathcal{E} = \{A, B, C, \ldots\},
\end{aligned} \quad (3)$$

where $\text{mult}(l)$ refers to the multiplicity of Wyckoff letter l, and $\text{Count}_e(l)$ is the number of atoms of element e assigned to l, N_e is the number of atoms of element e in a cell, and $L_{\text{fix}}(S_g) = \{l \in L(S_g) \mid \text{DoF}(l) = 0\}$ denotes the set of zero-degree-of-freedom Wyckoff positions. Such a position corresponds to a unique fixed crystallographic orbit and therefore can be occupied at most once in a complete assignment. The set $\{l_n^*\}_{n=1}^N$ denotes the exact optimizer of the formal objective in Eq. (3) and serves as the reference target for the finite-beam decoder. The key point here is that $\text{Count}_e(l)$ must be an integer multiple of $\text{mult}(l)$, a hard constraint that cannot be satisfied by simply using the maximum probability outputs of each token from the LM. The generalized multiplicity-constrained form of Eq. (3) is NP-hard, and Section A.3 of the SM formally establishes this result by reduction from three-dimensional matching. Consequently, finding the solution for Eq. (3) via brute-force enumeration is computationally intractable, as the search space grows exponentially with the number of atoms, scaling as $O(|L(S_g)|^N)$. To overcome this bottleneck, we propose a constraint satisfaction beam search with an element-decomposed, atom-resolved search and multiplicity-block expansion to obtain high-scoring valid assignments as a finite-width approximation to the constrained optimization problem. Let K denote the beam width and let $|L(S_g)|$ be the number of admissible Wyckoff letters for a candidate space group. The atom-wise assignment is decomposed by element, and each search expansion assigns a complete multiplicity-sized group of atoms to one Wyckoff letter. For an element containing N_e atoms, at most N_e such expansion rounds are required, since each expansion assigns at least one previously unassigned atom. Each round retains at most K states and considers at most $|L(S_g)|$ candidate letters per state. Summing over all elements, with $\sum_e N_e = N$, gives $O\left(NK|L(S_g)|\right)$ state–letter expansion calls for one space-group query.

For each state–letter expansion, at most C multiplicity-sized atom combinations are considered. The corresponding candidate-generation complexity is therefore $O\left(NK|L(S_g)|C\right)$. Since C is fixed independently of N, this additional factor does not change the linear asymptotic dependence on N. The optional fallback search introduces at most one additional search for an unsuccessful element, while the per-element Top-M truncation and the subsequent beam-truncated combination of element-level assignments introduce only bounded additional factors. The exclusion of repeatedly occupied zero-degree-of-freedom Wyckoff positions likewise constitutes a bounded feasibility check and does not alter the asymptotic scaling. Across the Top-G candidate space groups, the corresponding state–letter expansion bound is $O\left(NK\sum_{g=1}^G |L(S_g)|\right)$. We use $G = 5$, $K = 100$, and $M = 4$ for inference. The search is therefore linear in N when the beam width, candidate-space-group count, combination cap, and crystallographic Wyckoff-letter sets are fixed. Further implementation details and measured scaling are provided in Sections A.3 and F of the SM.

Because finite-beam truncation can discard feasible or higher-scoring states, the decoder is a finite-width approximation and does not provide a global optimality guarantee. We benchmarked it against an MILP reference over all 45230 MP-20 queries and cross-validated the exact reference using dynamic programming and brute-force

enumeration on their applicable subsets. Detailed failure rates, objective gaps, runtime scaling, and beam-width sensitivity are reported in Section F of the SM.

To generate up to 20 candidates for each test composition, we employ a hierarchical sampling strategy to ensure both macroscopic and microscopic structural diversity. First, we select the Top-5 most probable space groups predicted by LM_g. Subsequently, for each selected space group, we extract Top-4 valid Wyckoff assignments using our constrained beam search algorithm. This combination yields up to 20 distinct symmetry templates per composition, which are then used to guide the diffusion generation process. If fewer than four distinct valid assignments survive for a space group, all available assignments are used without duplication or padding. The canonical deduplication, tie-handling, and final-selection rules are detailed in Section A.3 of the SM.

Before diffusion, repeated occurrences of the same element–Wyckoff-letter pair are deterministically expanded into distinct orbit instances, each assigned a separate anchor index and, where applicable, an independent orbit-generator coordinate. The complete construction is provided in Sections A.3 and A.4 of the SM.

Third, we guide and constrain the diffusion-based generation of 3D atomic structures using the predicted space group and Wyckoff site template. To instantiate this approach, we adopt the graph neural network (GNN) applied by DiffCSP++[10] as the backbone network of the diffusion model. The predicted symmetry constraints then guide its sampling trajectory. Specifically, we employ a joint-rectification mechanism on both lattice and fractional coordinates. For the lattice, we constrain its $O(3)$-invariant logarithmic parameterization by applying a binary mask dictated by the inferred crystal family, ensuring the unit cell geometry strictly adheres to the specific constraints on lattice lengths and inter-axial angles defined by the crystal family. Simultaneously, for the fractional coordinates $\boldsymbol{F}_t$, we enforce the predicted Wyckoff constraints through anchor-based orbit reconstruction. Before each coordinate update, the network-predicted coordinate vector field is mapped to a common orbit frame, averaged over symmetry-equivalent sites, and expanded back to the orbit. At initialization and after each coordinate update, all sites in each orbit are reconstructed from its anchor coordinate using the corresponding affine symmetry operations and wrapped into the unit cell. Further details regarding the lattice masking and coordinate reconstruction are available in Section A.4 of the SM. The same symmetry-conditioned generation path is applicable across all 230 international space groups without crystallographic-category-specific branches and its space-group coverage and symmetry-identification limitations are detailed in Section H of the SM.

To evaluate our approach, we benchmark it against a comprehensive suite of generative models. We adopt DiffCSP++ [10] as the primary baseline, as it represents the current SOTA for standard CSP and serves as the most stringent benchmark for composition-conditioned structural generation. To provide a broader context for structural reconstruction fidelity, we also encompass CDVAE [18], DiffCSP [19], and CrystaLLM [9] in our comparative analysis, specifically regarding the Matching Rate metric. While MatterGen [14] and CrystalFormer [22] are leading methods in materials generation, they are excluded from direct quantitative comparison in this work. These models primarily operate on a joint generation paradigm that co-samples chemical compositions and structural coordinates. Adapting them into dedicated CSP models and training them from scratch without the original authors' optimization protocol risks an unfair and unrepresentative comparison. Thus, a direct evaluation is omitted.

To provide a comprehensive evaluation of our model, we adopt the Top-20 Stability, Uniqueness, and Novelty (SUN) metrics based on the core principles of Ref. [14], alongside the Top-20 Matching Rate [19]. Within each dataset, NextCrystal and DiffCSP++ are evaluated using the same MatterSim relaxation, energy-reference, phase-diagram, and failure-handling protocol. The detailed definitions and adapted calculation protocols for these metrics, designed to accommodate our extensive Top-20 evaluation setting, are provided in Sections C.1 and C.2 of the SM. Uniqueness is evaluated using a sequential representative convention: candidates are processed in rank order, and the first member of each local duplicate class is retained as its representative. For joint SUN, the same candidate must satisfy the Stability and Novelty criteria and survive the uniqueness filter defined in Section C.1 of the SM. Thus, joint SUN is a same-candidate criterion rather than the product of the aggregate component rates. Stability requires $E_{\text{hull}} \leqslant 0.1$ eV/atom, and Table 1 reports composition-level success under the Top-20 evaluation protocol defined in Section C.1 of the SM. The SUN metrics emphasize physical plausibility, structural diversity, and originality, aspects that are more directly aligned with the objectives of open-ended discovery in materials science. In contrast, the Matching Rate prioritizes structural fidelity by quantifying the precision with which the model reconstructs the experimental GT geometry within defined coordinate and symmetry tolerances. It is worth noting that these two metrics can be inherently contradictory in a Top-1 evaluation because a model must choose between replicating a known structure to achieve a high Matching Rate or exploring a previously unknown one to achieve high Novelty. This tension is effectively resolved in our Top-20 setting. By generating multiple candidates for each composition, the model is afforded the capacity to simultaneously demonstrate structural fidelity by recovering the experimental GT and chemical creativity by proposing novel near-hull candidates. Thus, reporting both metrics at Top-20 provides a holistic assessment of the model's dual role as both a precise structural reconstructor and a robust discovery engine for uncharted materials space.

Table 1. Comparison of Top-20 SUN metrics (%) across three benchmarks. Stability, Uniqueness, Novelty, and joint SUN are reported as composition-level Top-20 at-least-one success rates over the complete official test splits, with the number of test compositions in each dataset used as the denominator. Results for DiffCSP++ [10] are reproduced using the official source code implementation [27], utilizing the CSPML [23] retrieval method.

Dataset	Method	Stability	Uniqueness	Novelty	**SUN**
MP-20	DiffCSP++	36.51	**100.0**	54.82	14.48
	NextCrystal	**81.40**	**100.0**	**94.33**	**35.37**
Perov-5	DiffCSP++	75.77	**100.0**	99.52	72.97
	NextCrystal	**99.08**	**100.0**	**99.97**	**92.42**
MPTS-52	DiffCSP++	11.95	**100.0**	45.97	5.44
	NextCrystal	**47.63**	**100.0**	**78.80**	**27.89**

Our experiments are conducted on three standard CSP benchmark datasets: MP-20 [28], Perov-5 [29], and MPTS-52 [10]. These datasets were previously used in DiffCSP++ and are employed here with identical training, validation and testing splits to ensure a fair comparison. Detailed descriptions and statistical information regarding these three datasets are available in Section C.2 of the SM. The experimental results for the DiffCSP++ baseline are obtained using its official open-source scripts for standard training and inference [27]. Table 1 presents a side-by-side comparison between the performance of this original method and that of our symmetry-enhanced method.

As shown in Table 1, our method yields notable and consistent improvements over DiffCSP++ across all three benchmarks on the comprehensive SUN metrics. On the MP-20 dataset, our method improves the Stability and Novelty components by approximately 123% and 72%, respectively, relative to the baseline. The advantages are even more pronounced on the challenging MPTS-52 dataset, where our method achieves relative gains of about 299% in Stability and 71% in Novelty. Crucially, these component-level improvements compound to yield an approximately 412% increase in the overall SUN metric. Even on Perov-5, which features relatively simple structural motifs, our method maintains a clear advantage in stability while matching the saturated uniqueness and Novelty scores of the baseline. The 100% composition-level Uniqueness values should not be interpreted as candidate-level diversity; the corresponding candidate-level

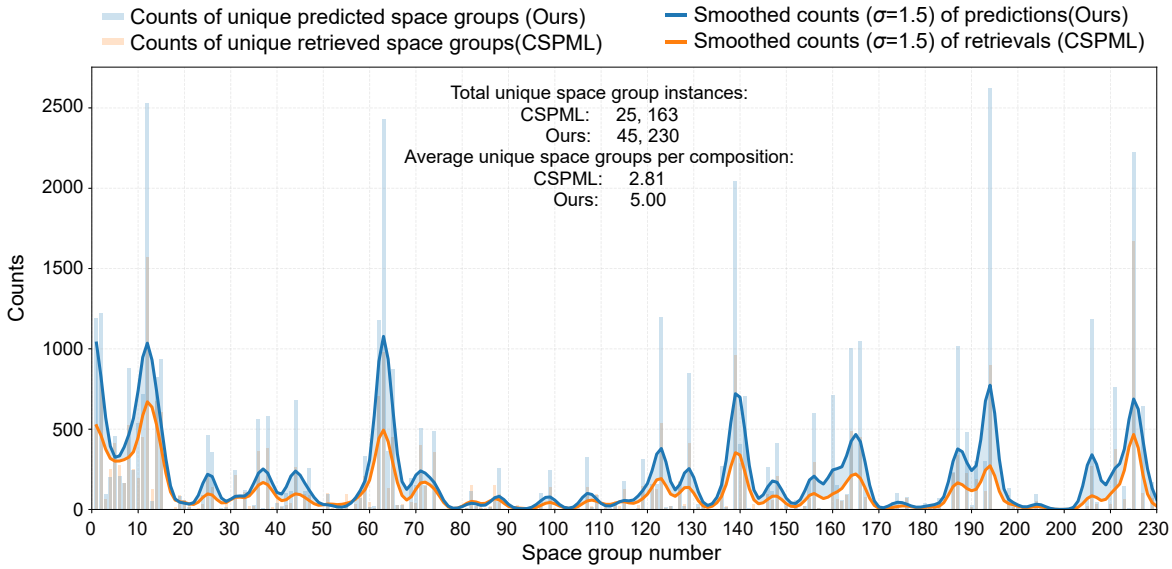

Fig. 2. Absolute frequency distribution (raw counts) of space groups among the candidates across the test set of MP-20. Counts are calculated using a unique-counting principle per composition, where any specific space group present in the candidate structures for a single testing composition is counted exactly once. Our generative approach yields a higher total unique count, demonstrating a consistently higher richness in symmetry exploration per composition compared to the localized redundancy of the retrieval baseline, i.e., CSPML [23].

yields are reported in Section C.1 of the SM. The significant rise in the overall SUN metric across all benchmark datasets confirms that our method generates candidates that more frequently satisfy the adopted near-hull, Uniqueness, and Novelty criteria jointly. Paired symmetry-rectification controls further show that enforcing the lattice and Wyckoff manifolds reduces initial geometric and mechanical residuals and substantially improves fixed-budget relaxability, providing a mechanistic pathway to the higher near-hull success (see Section H of the SM).

To identify the sources of these gains, we further performed controlled ablations and symmetry-mechanism analyses on the complete MP-20 test split. Sparse-MoE/FFN and FiLM/concatenation ablations are reported in Section G of the SM; GT versus predicted space groups, GT versus generated Wyckoff templates, and per-step rectification are analyzed in Section H of the SM. The complementary greedy/beam/exact-search analysis is provided in Section F of the SM, and metric-matched comparisons with explicit-symmetry generators are reported in Section I of the SM.

To gain deeper insights into the distinct generative behaviors of retrieval-based and symmetry-driven methods, we conducted detailed case studies on two representative samples, Zr_3Fe_8Mo and $NaCr_4O_8$, as provided in Section D of the SM. An analysis of these examples reveals that the baseline systematically struggles to satisfy the critical criteria for successful discovery. It often either relies on an exact structural match of a documented phase, which satisfies the adopted near-hull criterion but fails the Novelty criterion, or forces the atoms into a retrieved but topologically mismatched template. In the latter case, atoms are coerced into energetically unfavorable coordination environments to satisfy incorrect symmetry constraints, limiting internal structural relaxation and, in the $NaCr_4O_8$ case, producing larger relaxation-induced Cr-centered local strain. In contrast, our symmetry-driven approach generates novel candidates that satisfy the adopted near-hull criterion by adaptively allocating atoms to higher-multiplicity orbits with critical geometric flexibility. This use of predicted rather than retrieved Wyckoff templates allows NextCrystal to yield candidates that successfully satisfy all three SUN criteria. For $NaCr_4O_8$, NextCrystal reduces the MatterSim/MP2020 hull distance from 113.9 to 98.0 meV/atom and the Cr-centered relaxation-induced local strain from 4.80% to 2.90%. For Zr_3Fe_8Mo, the retrieval candidate matches `mp-1215698`, whereas the NextCrystal candidate remains near-hull without a match (see Section D of the SM).

The performance disparity observed in Table 1 and Section D of the SM is consistent with differences between retrieval-based and symmetry-driven generation. The baseline relies on CSPML, which essentially operates as a static lookup system. It rests on the isomorphism assumption, presupposing that for any new material, a structurally identical twin already exists in the training database. Its performance can therefore depend on whether a suitable prototype is present in the retrieval database. When the model forces a new query composition into a retrieved but topologically mismatched template, it introduces geometric frustration. The atoms are coerced into energetically unfavorable coordination environments to satisfy incorrect symmetry constraints, which can hinder relaxation toward low-energy configurations. This reliance on data lookup inherently confines the model to the known structural motifs, rendering it incapable of discovering new polymorphs that lie outside the existing library.

In contrast, our method performs symmetry prediction without template retrieval at inference and couples it to constrained optimization. Instead of retrieving existing solutions, the model learns a unified representation space from diverse structural data, enabling it to generalize and derive valid Wyckoff patterns for unseen compositions based on chemical rules. Crucially, a heuristic combinatorial optimization algorithm enforces algebraic consistency, ensuring the sum of site multiplicities matches the atomic stoichiometry. This mechanism allows the model to explore alternative symmetry-compatible candidates while enforcing the stoichiometric–multiplicity constraints before diffusion. The selected templates are subsequently imposed during diffusion through the symmetry-rectification procedure described above.

To further quantify generative symmetry diversity, we evaluate the distribution of unique space group instances generated across the testing set, as shown in Fig. 2. To faithfully capture the true exploratory richness for each individual composition, we apply a strict unique-counting principle: a unique space group instance is defined such that if a specific space group appears within the candidate structures generated for a single testing composition, it is counted exactly once for that composition, regardless of its redundancy. The results demonstrate that the CSPML retrieval baseline [23] often collapses into retrieving multiple structural templates that share the identical space group for a given composition. Because of this localized redundancy, multiple retrieved candidates frequently contribute only a single unique space group count, which drastically shrinks its total aggregate score. In contrast, our generative approach explicitly infers multiple mutually distinct space groups directly from the probability

Table 2. Comparison of Top-20 Matching Rates (%) across three benchmarks. Experimental results for CDVAE [18], DiffCSP [19], and CrystaLLM [9] are taken from Antunes *et al.* [9]. Results for DiffCSP++ [10] are reproduced from the official source code [27] using the CSPML [23] retrieval-based template method as described in the original study.

Dataset	Method	Matching Rate
MP-20	CDVAE	66.95
	DiffCSP	77.93
	CrystaLLM	73.97
	DiffCSP++	74.65
	NextCrystal	**81.70**
Perov-5	CDVAE	88.51
	DiffCSP	98.60
	CrystaLLM	97.60
	DiffCSP++	99.84
	NextCrystal	**99.95**
MPTS-52	CDVAE	20.79
	DiffCSP	34.02
	CrystaLLM	33.75
	DiffCSP++	40.53
	NextCrystal	**43.26**

distribution predicted by the language models. This mechanism yields a higher total count of unique instances and demonstrates a consistently richer and more abundant exploration across different space groups for each composition, as evidenced by the higher frequency observed across nearly all space group indices in Fig. 2.

Complementing the discovery-oriented SUN metrics, Table 2 presents the Top-20 Matching Rate, which quantifies the model's ability to reconstruct experimentally known GT structures. As shown, our method consistently achieves SOTA performance across all three benchmarks, outperforming all of the baseline methods. This result is scientifically significant when interpreted in the context of generative design. In composition-conditioned materials discovery, the Matching Rate confirms that the model's generative distribution covers the manifold of physically realized structures. A high Matching Rate indicates that the symmetry constraints imposed by our framework do not overly restrict the search space or exclude valid GT configurations. Instead, they effectively guide the diffusion process within a chemically reasonable manifold. Consequently, the combination of superior SUN scores (Table 1) and leading Matching Rates (Table 2) demonstrates that our method resolves the trade-off between exploration and exploitation: it is robust enough to recover known stable phases with high fidelity while possessing the generative flexibility to propose novel, metastable polymorphs beyond the existing samples.

To demonstrate the practical utility of our model, we applied the MP-20 pre-trained NextCrystal model to the generation of HfO_2 structures. HfO_2 exhibits a rich polymorphic landscape, with closely com-

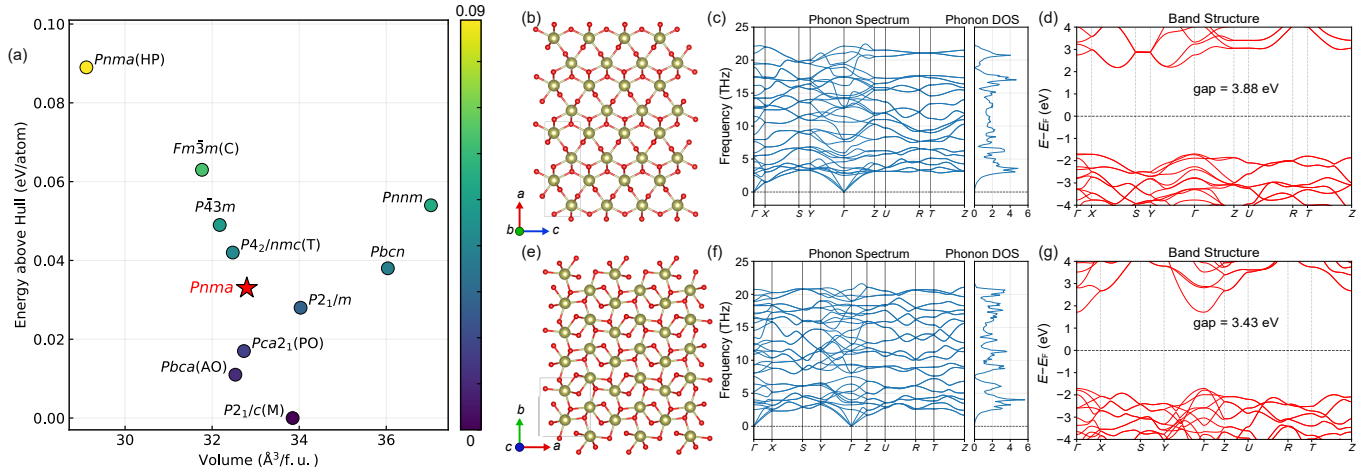

Fig. 3. Comparison between the independently recovered *Pnma* phase of HfO_2 and the conventional HP-*Pnma* phase. (a) Distribution of the 11 retained HfO_2 candidates in the volume–energy space, where the star marks the recovered *Pnma* phase. (b)–(d) Crystal structure, phonon dispersion, and electronic band structure of the recovered *Pnma* phase. (e)–(g) Crystal structure, phonon dispersion, and electronic band structure of the HP-*Pnma* phase.

peting nonpolar, ferroelectric, and antiferroelectric-related phases, making the identification of energetically accessible crystal structures central to understanding and tuning its functional properties [30]. We use $(HfO_2)_n$ ($n = 1$–12) as the initial inputs to NextCrystal. For each input, NextCrystal generates up to 100 structures by sampling 20 space groups and 5 sets of Wyckoff position assignments, resulting in a total of 887 HfO_2 candidate structures. These generated structures are subsequently validated using first-principles calculations within a workflow integrating ABACUS [31,32], PYATB [33], ZStar, and PHONOPY [34] packages. The exchange–correlation energy is treated within the PBEsol functional [35], and numerical atomic-orbital basis sets of TZDP and DZP were employed for Hf and O, respectively [36].

After structural relaxation, duplicate removal, and stability assessment, 11 structurally distinct candidates are retained. Among them, NextCrystal independently recovers the low-energy antipolar oV polymorph of HfO_2 (*Pnma*, No. 62) [Fig. 3(b)], previously reported by Antunes *et al.* [37,38]. Structure standardization and atom mapping confirm that the generated structure is equivalent to the published oV/a62 structure [39] and distinct from the conventional high-pressure cotunnite/oII *Pnma* phase [Fig. 3(e)]. The blind recovery therefore does not represent discovery of a new HfO_2 polymorph; instead, it provides an out-of-sample validation that NextCrystal can locate a known low-energy minimum without retrieving its structural prototype. Within our unified first-principles workflow, the recovered oV structure is 0.056 eV/atom lower in energy than the conventional high-pressure (HP)-*Pnma* phase at the PBEsol level. A PBE sensitivity calculation at 0 GPa retains $E_{oV} < E_{HP}$ while leaving oV 0.04073 eV/atom above monoclinic $P2_1/c$, consistent with its description as a metastable polymorph (see Section E of the SM).

Relative to the representative HfO_2 phases, namely monoclinic $P2_1/c$ (M), antiferroelectric orthorhombic Pbca (AO), polar orthorhombic $Pca2_1$ (PO), and tetragonal $P4_2/nmc$ (T), the recovered oV *Pnma* phase occupies an intermediate position in the low-energy metastable manifold [Fig. 3(a)]. Its energy above hull is 0.033 eV/atom, which is higher than those of M, AO, and PO by 0.033, 0.022, and 0.016 eV/atom, respectively, but lower than that of T by 0.009 eV/atom. Its intermediate energy and volume are consistent with the published low-energy oV phase.

Table 3. Relaxed lattice parameters, symmetry-inequivalent atomic sites, fractional coordinates, and Born effective charge tensors Z^* of the independently recovered *Pnma* phase of HfO_2.

Wyckoff site	Fractional coordinates	Born effective charge tensor Z^* (e)
Lattice: a = 10.1717 Å, b = 3.3978 Å, c = 3.7953 Å; $\alpha=\beta=\gamma=90°$		
4c (Hf)	(0.3817, 0.7500, 0.2524)	$\begin{bmatrix} 5.008 & 0.000 & 0.314 \\ 0.000 & 5.238 & 0.000 \\ 0.130 & 0.000 & 4.999 \end{bmatrix}$
4c (O1)	(0.7152, 0.7500, 0.5778)	$\begin{bmatrix} -2.661 & 0.000 & 0.870 \\ 0.000 & -3.016 & 0.000 \\ 0.812 & 0.000 & -2.046 \end{bmatrix}$
4c (O2)	(0.9746, 0.7500, 0.7391)	$\begin{bmatrix} -2.346 & 0.000 & -0.068 \\ 0.000 & -2.223 & 0.000 \\ 0.000 & 0.000 & -2.954 \end{bmatrix}$

The recovered oV phase has no imaginary phonon frequencies, confirming its dynamical stability [Fig. 3(c)]. It exhibits a semilocal PBEsol Kohn–Sham band gap of 3.88 eV [Fig. 3(d)], which is 0.45 eV larger than the corresponding value of 3.43 eV for HP-*Pnma* phase [Fig. 3(g)]. These absolute values should not be interpreted as quantitative quasiparticle or optical gaps because semilocal density functional theory generally underestimates the HfO_2 band gap; they are used here as a same-level comparison between the two *Pnma* structures. Spin–orbit coupling may further modify the electronic band structure of Hf-containing compounds and is not explicitly included in this comparison. Although this *Pnma* phase is centrosymmetric and therefore nonpolar, its Born effective charge tensors still exhibit pronounced anomalous dynamical charge response, indicating substantial Hf–O hybridization. The symmetry-inequivalent atomic positions and corresponding Born effective charge tensors are summarized in Table 3. Comparison with the conventional HP-*Pnma* phase shows distinct dynamical-charge responses despite the common space-group number; for example, O1 Z^*_{yy} changes from −3.016 in oV to −2.294 in HP-*Pnma*, while O2 Z^*_{zz} changes from −2.954 to −2.293 (see Section E of the SM). The calculated static dielectric constants, $\varepsilon_{xx}=17.85$, $\varepsilon_{yy}=19.93$, and $\varepsilon_{zz}=18.05$, reveal moderate dielectric anisotropy and substantial lattice polarizability.

These results show that the present model can independently recover physically meaningful low-energy HfO_2 polymorphs without retrieving their structural prototypes. Notably, the model is not trained specifically for the HfO_2 system, providing an out-of-sample test beyond system-specific training.

Conclusion

In conclusion, this work introduces NextCrystal, a symmetry-driven CSP framework that combines data-driven symmetry prediction, constrained Wyckoff assignment, and symmetry-conditioned diffusion without template retrieval at inference. By marrying language models with a linear-complexity heuristic search, NextCrystal successfully renders the inherently NP-hard combinatorial challenge of symmetry enforcement computationally tractable. This approach improves Top-20 SUN over DiffCSP++ across MP-20, Perov-5, and MPTS-52 and achieves higher Matching Rates on the three benchmarks. As a blind-rediscovery case study, NextCrystal recovers the known low-energy antipolar oV *Pnma* polymorph of HfO_2 without retrieving a structural prototype during inference; standardized structure matching after a common first-principles relaxation confirms its equivalence to the published oV/a62 structure.

Building upon this foundation, future extensions will integrate high-fidelity AI electronic structure prediction models, such as TraceGrad [40] and NextHAM [41] developed by our group, to enable rapid, high-throughput screening of the generated candidates. Furthermore, subsequent iterations of this framework will focus on integration with high-throughput computational screening and targeted experimental validation. In later stages, experimentally synthesized and characterized candidates could provide feedback for subsequent model refinement. This combination of computational design, AI-driven screening, and empirical validation provides a practical direction for improving iterative materials discovery.

Author Contributions

Jinming Mu: Software, Validation, Formal analysis, Writing-Original Draft, Visualization. Lixin He: Supervision, Project Administration, Formal analysis, Writing-Original Draft, Funding Acquisition. Xudong Zhu: Software, Validation, Formal analysis, Writing-Original Draft, Visualization. Shi Yin: Conceptualization, Methodology, Software, Formal analysis, Writing-Original Draft, Visualization, Funding Acquisition.

Code Availability

The source code of our method is available from the public GitHub repository at https://github.com/Mu-Jinming/NextCrystal. The code version accompanying this work is identified by commit `773549e0fc64c2b7ff161e4b38de82cbcc5b9c06`. The repository includes a top-level MIT license.

Data Availability

The crystal structures generated by our method for the test set compositions of the MP-20 [28], Perov-5 [29], and MPTS-52 [10] benchmarks, along with our predicted HfO_2 structures, are publicly available at: https://drive.google.com/drive/folders/1J78WYmN5YY-kwFP1z4XIRQ3VjAKRuPWyQ?usp=drive_link. The relaxed HfO_2 CIF used for the reported structural, phonon, electronic, and Born-effective-charge analyses is included in the public HfO_2 data folder at the same link.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the Advanced Materials—National Science and Technology Major Project (Grant No. 2025ZD0618401), the National Natural Science Foundation of China (Grant Nos. 62506112 and 12134012), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB0500201), and the Anhui Provincial Science and Technology Breakthrough Project (Grant No. 202523o09050015). The numerical calculations were performed on the USTC High-Performance Computing facilities and Hefei advanced computing center.

References

[1] Green M A, Ho-Baillie A, and Snaith H J 2014 *Nat. Photonics* **8** 506
[2] Berger L I 2020 *Semiconductor Materials* (CRC Press)
[3] Jiang Q and Zhu K 2024 *Nat. Rev. Mater.* **9** 399
[4] Ding J, Ji D, Yue Y, and Smedskjaer M M 2024 *Small* **20** 2304270
[5] Langer R and Tirrell D A 2004 *Nature* **428** 487
[6] Seh Z W, Kibsgaard J, Dickens C F, Chorkendorff I B, Nørskov J K, and Jaramillo T F 2017 *Science* **355** eaad4998
[7] Marzari N, Ferretti A, and Wolverton C 2021 *Nat. Mater.* **20** 736
[8] Jiao R, Huang W, Lin P, Han J, Chen P, Lu Y, and Liu Y 2023 *Proc. Int. Conf. Learn. Represent. (ICLR 2023)*
[9] Antunes L M, Butler K T, and Grau-Crespo R 2024 *Nat. Commun.* **15** 10570
[10] Jiao R, Huang W, Liu Y, Zhao D, and Liu Y 2024 *Proc. Int. Conf.*

Learn. Represent. (ICLR 2024)
[11] Yang S, Batzner S, Gao R, Aykol M, Gaunt A, McMorrow B, Rezende D, Schuurmans D, Mordatch I, and Cubuk E D 2024 *Adv. Neural Inf. Process. Syst.* **37** 38799
[12] Zhu R, Nong W, Yamazaki S, and Hippalgaonkar K 2024 *Matter* **7** 3469
[13] Han Y, Ding C, Wang J, Gao H, Shi J, Yu S, Jia Q, Pan S, and Sun J 2025 *Nat. Comput. Sci.* **5** 255
[14] Zeni C, Pinsler R, Zügner D *et al*. 2025 *Nature* **639** 624
[15] Xu J, Wang X, Cheng W, Cao Y P, Shan Y, Qie X, and Gao S 2023 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, BC, Canada, 2023, p. 20908
[16] Wang C, Peng H Y, Liu Y T, Gu J, and Hu S M 2025 *Comput. Vis. Media* **11** 1
[17] Xin Y, Xie Z, Chen Y, Gao P, Zhao H J, Lv J, Wang Y, and Ma Y 2025 *Phys. Rev. B* **111** 134112
[18] Xie T, Fu X, Ganea O E, Barzilay R, and Jaakkola T S 2022 *Proc. Int. Conf. Learn. Represent. (ICLR 2022)*
[19] Jiao R, Huang W, Lin P, Han J, Chen P, Lu Y, and Liu Y 2023 *Adv. Neural Inf. Process. Syst.* **36** 17464
[20] Levy D, Panigrahi S S, Kaba S O, Zhu Q, Lee K L K, Galkin M, Miret S, and Ravanbakhsh S 2025 *Proc. Int. Conf. Learn. Represent. (ICLR 2025)*
[21] Ekström Kelvinius F, Andersson O B, Parackal A S, Qian D, Armiento R, and Lindsten F 2025 *Proc. Mach. Learn. Res.* **267** 15130
[22] Cao Z, Luo X, Lv J, and Wang L 2025 *Sci. Bull.* **70** 3522
[23] Kusaba M, Liu C, and Yoshida R 2022 *Comput. Mater. Sci.* **211** 111496
[24] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L, and Polosukhin I 2017 *Conference on Neural Information Processing Systems* p. 5998
[25] Puigcerver J, Ruiz C R, Mustafa B, and Houlsby N 2024 *Proc. Int. Conf. Learn. Represent. (ICLR 2024)*
[26] Perez E, Strub F, de Vries H, Dumoulin V, and Courville A 2018 *Proc. AAAI Conf. Artif. Intell.* **32** 3942
[27] Jiao R, Huang W, Liu Y, Zhao D, and Liu Y 2024 *DiffCSP++: official implementation of Space Group Constrained Crystal Generation*. GitHub. https://github.com/jiaor17/DiffCSP-PP
[28] Jain A, Ong S P, Hautier G *et al*. 2013 *APL Mater.* **1** 011002
[29] Castelli I E, Landis D D, Thygesen K S, Dahl S, Chorkendorff I, Jaramillo T F, and Jacobsen K W 2012 *Energy Environ. Sci.* **5** 9034
[30] Zhu T, Ma L, Deng S, and Liu S 2024 *npj Comput. Mater.* **10** 188
[31] Li P, Liu X, Chen M, Lin P, Ren X, Lin L, Yang C, and He L 2016 *Comput. Mater. Sci.* **112** 503
[32] Chen M, Guo G C, and He L 2010 *J. Phys.: Condens. Matter* **22** 445501
[33] Jin G, Pang H, Ji Y, Dai Z, and He L 2023 *Comput. Phys. Commun.* **291** 108844
[34] Togo A, Chaput L, Tadano T, and Tanaka I 2023 *J. Phys.: Condens. Matter* **35** 353001
[35] Perdew J P, Ruzsinszky A, Csonka G I, Vydrov O A, Scuseria G E, Constantin L A, Zhou X, and Burke K 2008 *Phys. Rev. Lett.* **100** 136406
[36] Lin P, Ren X, and He L 2021 *Phys. Rev. B* **103** 235131
[37] Azevedo Antunes L, Ganser R, Alcala R, Mikolajick T, Schroeder U, and Kersch A 2021 *Appl. Phys. Lett.* **119** 082903
[38] Azevedo Antunes L, Ganser R, Kuenneth C, and Kersch A 2022 *Phys. Status Solidi RRL* **16** 2100636
[39] Kersch A, Ganser R, and Trien M 2022 *Front. Nanotechnol.* **4** 1026286
[40] Yin S, Pan X, Wang F, and He L 2025 *Proceedings of the 42nd International Conference on Machine Learning* (PMLR) 267 72364
[41] Yin S, Dai Z, Pan X, and He L 2026 *Proc. Int. Conf. Learn. Represent. (ICLR 2026)*