ADVANCED MATERIALS

REVIEW `OPEN ACCESS`

Generative Models for Crystalline Materials

Houssam Metni[1,2] | Laura Ruple[2] | Lauren N. Walters[3,4] | Luca Torresi[1,2] | Jonas Teufel[1,2] |
Henrik Schopmans[2] | Jona Östreicher[1,2] | Yumeng Zhang[1,2] | Marlen Neubert[1,2] | Yuri Koide[1,2] |
Kevin Steiner[2] | Paul Link[2] | Lukas Bär[2] | Mariana Petrova[2] | Gerbrand Ceder[3,4] | Pascal Friederich[1,2]

[1]Institute of Nanotechnology, Karlsruhe Institute of Technology, Karlsruhe, Germany | [2]Institute of Anthropomatics and Robotics, Karlsruhe Institute of Technology, Karlsruhe, Germany | [3]Department of Materials Science and Engineering, University of California, Berkeley, California, USA | [4]Materials Sciences Division, Lawrence Berkeley National Laboratory, California, USA

Correspondence: Pascal Friederich (pascal.friederich@kit.edu)

ABSTRACT

Understanding structure-property relationships in materials is fundamental in condensed matter physics and materials science. Over the past few years, machine learning (ML) has emerged as a powerful tool for advancing this understanding and accelerating materials discovery. Early ML approaches primarily focused on constructing and screening large material spaces to identify promising candidates for various applications. More recently, research efforts have increasingly shifted toward generating crystal structures using end-to-end generative models. This review analyzes the current state of generative modeling for crystal structure prediction and *de novo* generation. It examines crystal representations, outlines the generative models used to design crystal structures, and evaluates their respective strengths and limitations. Furthermore, the review highlights experimental considerations for evaluating generated structures and provides recommendations for suitable existing software tools. Emerging topics, such as modeling disorder and defects, integration in advanced characterization, incorporating synthetic feasibility constraints, and model explainability are explored. Ultimately, this work aims to inform both experimental scientists looking to adapt suitable ML models to their specific circumstances and ML specialists seeking to understand the unique challenges related to inverse materials design and discovery.

1 | Introduction

The discovery of new materials has been one of the most important drivers of technological development. Over time, efforts to address the materials discovery challenge have progressed through experiment-driven advances, the development of theoretical frameworks, and large-scale simulations. Today, these approaches are complemented by a fourth paradigm [1]: data-driven science, where machine learning (ML) approaches, powered by extensive databases of experimental and simulated data [2–8] as well as modern computing

infrastructure, open unprecedented opportunities for materials design [9–12].

Within the fourth paradigm, a variety of computational methods have been developed to explore and design crystal structures with application-specific properties. Conventional high-throughput virtual screening approaches were enhanced with active learning approaches to accelerate the generation of large informative sets of hypothetical materials using combinations of density functional theory (DFT), predictive ML methods, and uncertainty estimation [13]. Complementing this, machine-learned

interatomic potentials (MLIPs) accelerate optimization and sampling workflows and thus enable fast property estimation and efficient down-selection from massive candidate spaces [14]. Evolutionary and stochastic search algorithms [14], as well as rule-based structure generators, provide additional strategies to systematically explore structural possibilities, often guided by physical constraints or heuristic rules. Furthermore, the development of automated high-throughput labs [15–17] as well as increasingly autonomous materials acceleration platforms and self-driving labs [18–23] are starting to close the gap between virtual and experimental materials design and discovery. Together, these developments illustrate a clear methodological progression: from physics-based simulations to ML-augmented searches, towards autonomous virtual and experimental discovery, generating increasing volumes of ML-ready data, which in turn accelerates the design and discovery process.

One of the key components in this evolution in methodology is end-to-end generative models, which aim to directly propose material structures without relying on exhaustive searches or handcrafted rules [9]. By learning structural patterns and symmetries from data, generative approaches invert the traditional workflow: instead of generating candidates and evaluating them sequentially, they seek to directly suggest new structures that satisfy desired constraints from the outset. Over the past ten years, new developments in generative models from the core machine learning community were quickly adopted for inverse design of molecules and crystal structures. Compared to generative models for molecules [24], crystal generation presents additional challenges: crystals are periodic, often involve complex unit cells, and must respect strict symmetry constraints, making representation and model design substantially more intricate. For this reason, transfer of new approaches to crystal generation was usually delayed by 1-2 years. For instance, the first successful attempts to use variational autoencoders (VAEs) for images were published in 2013 [25], then adapted to generate molecular structures in 2016 [26], and to materials in 2019 [10]. Generative adversarial networks (GANs), introduced in 2014 [27], were first used for molecules in 2018 [28], for materials compositions in 2018 [29], and for crystal structures in 2020 [30]. Normalizing flows, introduced in 2014 [31], were applied to molecules in 2019 [32] and to crystal structures in 2022 and 2023 [33, 34]. The same is true for diffusion models (introduced in 2020 [35], applied to molecules in 2021 and 2022 [36–38], and to crystal structures in 2022 [39]), and transformers (introduced in 2017 [40], first applied to molecules in 2019-2021 [41, 42], and to crystal structures in 2023 and 2024 [43, 44]).

Recent reviews have broadly discussed and offered perspectives on ML-assisted inverse design in materials science [45–49], some of which have focused specifically on crystal structure prediction tasks [50–52]. Breuck et al. provided a comprehensive and technical overview of the methods and models used in generative approaches for crystal structures, with discussions on methodologies and future directions, but with a primary focus on the ML models themselves [53]. Li et al. similarly emphasized the models, while also discussing datasets and evaluation metrics, and offering a brief perspective on experimental implementation, highlighting challenges such as synthesizability [54]. More recently, Park et al. provided a perspective on the need to integrate synthesizability considerations into computational materials design [55]. This perspective article emphasized that effective inverse design must account not only for predicted stability but also for practical pathways to experimental realization, a topic of increasing importance that will also be reviewed here.

In this work, we extend these efforts in multiple aspects by focusing on generative models for crystalline materials with (*i*) an emphasis on recent innovations in ML approaches in the last years, i.e., models based on diffusion, flows, and transformers; (*ii*) a specific focus on introducing conceptual differences of representations, datasets, and evaluation metrics, also making these distinctions accessible to non-experts and new researchers in this field; and (*iii*) to enhance practical relevance beyond previous review papers, a discussion of the practical adaptation of generative models for experimental discovery, as well as the current limitations and emerging topics, e.g., defects and disorder, characterization workflow integration, synthetic feasibility, and model interpretability.

Combining these three aspects, this review should serve as a useful guide for an audience of experimental and computational materials scientists looking to find an entry point to existing generative models and how to use them in their specific applications and related materials design challenges. At the same time, this review should also inspire the materials informatics and computer science communities, looking for an experimental perspective on currently open challenges and considerations to take into account in the next generation of generative models for materials.

2 | Representations

Crystals are solids composed of atoms, ions, or molecules arranged in a periodic, three-dimensional pattern. This periodicity is described using a unit cell, the smallest repeating unit that fully captures the symmetry and structure of the crystal. Unit cells are defined by lattice parameters—lengths a, b, c and angles α, β, γ—and contain a basis of atoms. Translating the unit cell through space along lattice vectors generates the infinite crystal lattice.

Symmetry plays a fundamental role in crystallography (Figure 1). Crystals are categorized into space groups, which define all symmetry operations (including rotations, reflections, inversions, and fractional translations) that leave the crystal structure invariant [56]. Each space group also defines an asymmetric unit, the smallest portion of the structure that can be used to reconstruct the entire unit cell through the application of the group's symmetry operations. Atomic positions in the unit cell are described using Wyckoff positions, which specify the symmetry-allowed coordinates of atoms. Each Wyckoff position is associated with a site multiplicity (the number of equivalent positions generated by symmetry) and a site symmetry (the local point group symmetry of that position). This formalism allows crystal structures to be described compactly, while strictly adhering to symmetry constraints. Generative models for crystal structures employ a variety of machine-readable representations, such as the Crystallographic Information File (CIF) [57], graph-, and
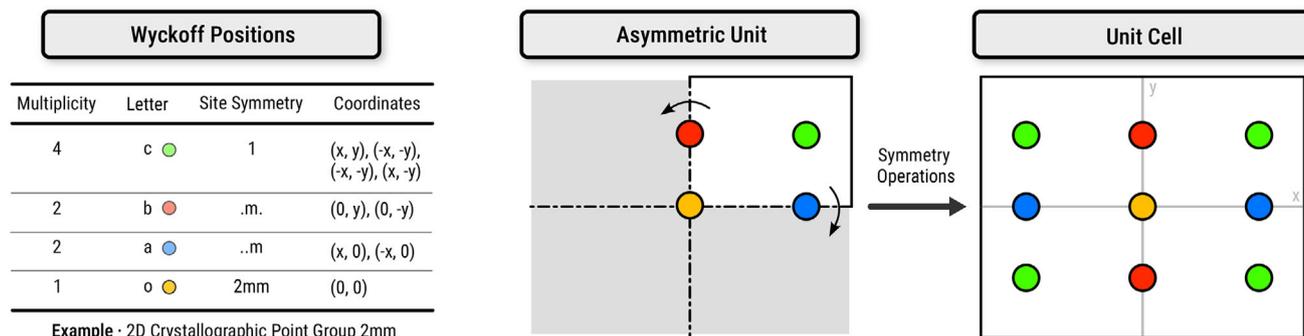
Multiplicity	Letter	Site Symmetry	Coordinates
4	c	1	(x, y), (-x, -y), (-x, -y), (x, -y)
2	b	.m.	(0, y), (0, -y)
2	a	..m	(x, 0), (-x, 0)
1	o	2mm	(0, 0)

Example · 2D Crystallographic Point Group 2mm

FIGURE 1 | Overview of the Wyckoff positions, asymmetric units, and symmetry operations in a unit cell.
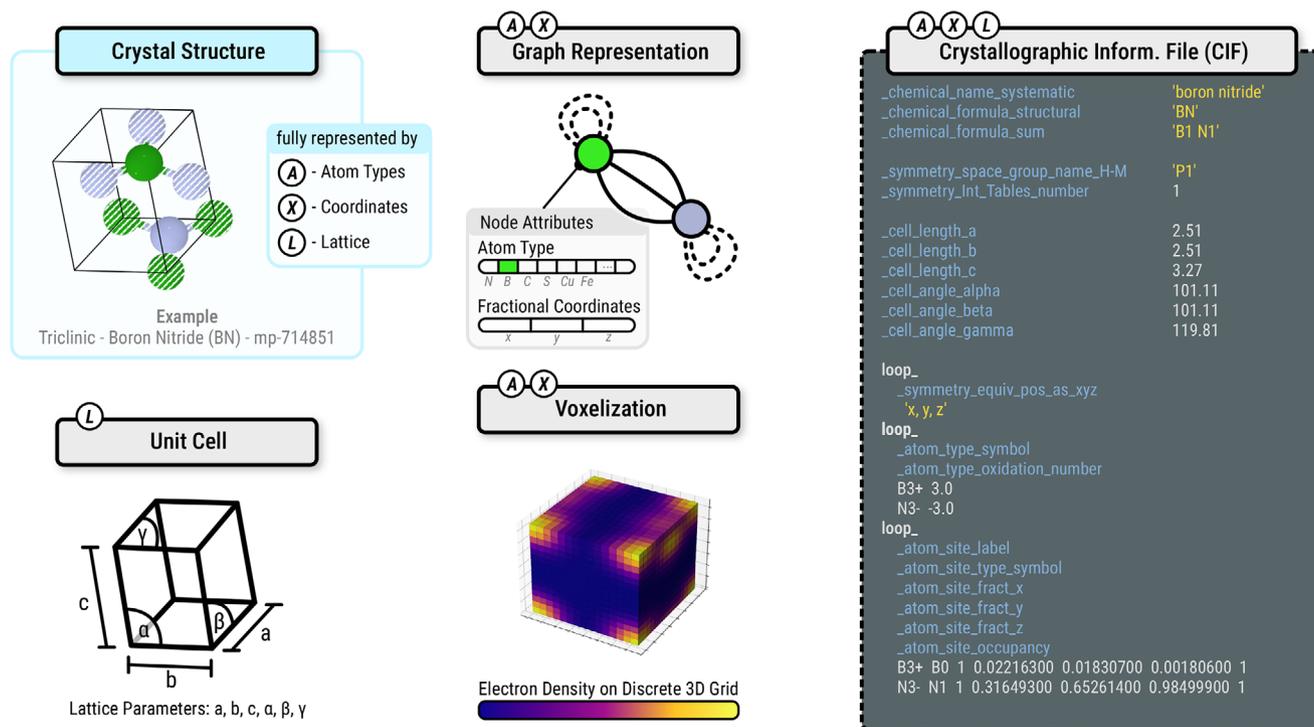

FIGURE 2 | Overview of the unit cell and machine-readable representations of crystal structures.

voxel-based formats (Figure 2). Additionally, some models also enforce space group symmetry constraints during training or generation, ensuring that generated structures adhere to crystallographic symmetry rules.

CIF is the standard format for storing and exchanging crystal structure data. CIF files contain comprehensive details about unit cell parameters, symmetry information (space group and symmetry operations), atomic positions (fractional coordinates, occupancy), and often Wyckoff positions, either explicitly or implicitly. Additional metadata from experimental or computational sources is typically included as well. CIF files are widely used in crystallographic databases and serve as a standard input/output format for many tools in generative modeling.

A more flexible and machine-learning-friendly representation of crystal structures is through graphs [58]. In graph-based models, nodes correspond to atoms and are represented with

necessary features such as element type (e.g., atomic number) and additional information such as electronegativity and valence. Edges represent interatomic interactions or spatial proximity, typically determined using distance-based cutoffs or k-nearest neighbors [59]. Periodic boundary conditions are incorporated by including edges that span across unit cell boundaries, effectively capturing the infinite periodic nature of crystals. These representations preserve the spatial and chemical relationships within the structure and are well-suited for use with graph neural networks (GNNs). Furthermore, symmetry can be encoded directly into the graph model: translational symmetry is handled through periodic boundary conditions, while rotational and reflectional symmetries can be incorporated via equivariant neural networks (e.g., E(3)-equivariant GNNs or SE(3)-Transformers), which ensure that the model's outputs transform consistently under these operations [60, 61].

Crystal structures can also be represented using voxelization, a technique in which the 3D space of the unit cell is

discretized into a regular grid of volumetric pixels (voxels). Each voxel encodes physical information, such as atomic occupancy, electron density, or atomic species, within that spatial region. The resulting 3D tensor representation is analogous to a 2D image and can be processed by 3D convolutional neural networks (CNNs) or volumetric generative models [62]. Voxel-based representations, when processed with fully-convolutional networks, are translation-equivariant. A shift in the input produces a corresponding shift in the feature maps. True translation invariance is only obtained if the architecture removes spatial information (e.g., via pooling). Rotational invariance can be approximated with data augmentation. However, due to the sparseness of the unit cell, voxel methods are computationally expensive and resolution-limited, making them less common than graph-based approaches for generative modeling tasks.

It is worth noting that alternative crystal structure representation methods have recently been proposed, although they have not yet been applied in the context of discriminative or generative ML models. For instance, Mrdjenovich and Persson introduced an algorithm representing each crystal as a unique set of integers derived from its lattice and atomic basis [63]. The lattice is encoded through an obtuse superbasis whose geometric features are discretized, while atomic positions are binned and ordered to remove symmetry and permutation redundancies. Together, these yield a canonical "crystal normal form", enabling efficient comparison and exploration of relationships between different crystal structures. Exploring new representations in generative models might increase performance without requiring more data or innovations in the model architectures.

3 | Databases

The effectiveness of ML methods strongly depends on the quality of the underlying data, as high-quality data is crucial for ensuring model reliability and generalization. The development of comprehensive and accurate crystal structure databases is thus of paramount importance in computational materials science, motivating the community efforts to invest in curated databases. On the one hand, the experimental materials science community has been gathering measurements and characterization results into open-access databases. On the other hand, the scarcity of experimental data has motivated the generation of large-scale simulated datasets using DFT, which complement and expand existing resources. While these databases are often categorized as experimental or DFT-calculated, many are hybrid in practice. Additional efforts have also led to the creation of application-specific databases tailored to particular materials. It is important to note that existing crystal structure databases are inherently imbalanced. Certain chemical elements, space groups, and structure types are strongly overrepresented due to historical research focus, experimental accessibility, and computational screening strategies. These representation biases can influence generative models in the absence of additional constraints. However, the quantitative impact of these biases remains difficult to assess given the lack of alternative, unbiased reference datasets. In this section, we present existing publicly available crystal structure databases, also summarized in Table 1.

3.1 | Experimental Crystal Structure Databases

The Inorganic Crystal Structure Database (ICSD), managed by the Leibniz Institute for Information Infrastructure (FIZ) Karlsruhe, is one of the largest databases for over 320 000 completely identified inorganic crystal structures, all of which have undergone quality checks to be verified [65]. The Cambridge Structural Database (CSD), managed by the Cambridge Crystallographic Data Centre (CCDC), consists of 1.25 million entries of small-molecule organic and metal-organic crystal structures, mainly data from X-ray and neutron diffraction analyses, spanning various experimental applications [66]. The Crystallographic Open Database (COD) is a collection of 520 000 small to medium-sized unit cell crystal structures of organic, inorganic, metal-organic compounds, and minerals [67]. Due to the open-access contribution model of this database, there are fewer filters to verify the validity and stability of the added structures. The Predicted Crystallographic Open Database (PCOD) is a COD subset of crystal structures predicted by focusing on predicted structures generated through computational methods such as GRINSP [84], while COD encompasses both experimentally determined and computationally predicted structures.

3.2 | Databases with DFT-Calculated Properties

Beyond the structure and potential accompanying experimental analyses, some databases also calculate further properties using DFT. The Materials Project (MP) is an open database containing structural, electronic, and energetic properties of known and predicted materials, including over 500 000 molecules and 200 000 inorganic materials [3]. A subset of the MP consisting of inorganic materials with less than 20 atoms in the unit cell, labeled as MP-20, contains about 45,000 metastable crystal structures spanning 89 different element types [39]. The Materials Project time split data (MPTS-52) is another subset of the MP comprising experimentally verified compounds with less than 52 atoms per unit cell [64].

The AFLOW (Automatic FLOW for Materials Discovery) repository encompasses data on approximately 3.9 million material compounds, along with a vast array of computed electronic, thermal, mechanical and thermodynamic properties [68]. Similarly, the Alexandria library is a quantum-chemical database of inorganic crystals, molecules, and low-dimensional materials whose geometries were optimized using over 5 million DFT calculations [69]. Other notable databases and benchmark tasks include the JARVIS-DFT [70], OQMD [71], Meta's OMAT24 [6], NOMAD [4, 82, 83], as well as the C2DB [72] and 2DMatPedia [73], both with a focus on two-dimensional (2D) materials.

3.3 | Application-Specific Databases

The Carbon-24 dataset focuses on 10 153 carbon structures selected from an original set of 101 529 structures [74]. These share the same elemental composition but vary in structure, spanning 6 to 24 atoms per unit cell. The Perov-5 dataset comprises properties of about 19 000 perovskite materials, curated from a dataset originally intended for water-splitting [75, 76]. In both of the

TABLE 1 | Summary of major publicly available materials databases from the literature, highlighting key characteristics including database size and material types covered. Experimental validation is indicated where applicable — while some databases are purely computational, others combine theoretical and experimentally verified structures to varying degrees.

Name and reference	Experimental validation	Database size	Material types
Materials Project (Full) [3]	Mixed (some verified)	>700 000 total, 200 000 crystals	Molecules and inorganic crystals
Materials Project (MP-20) [39]	Yes	~45 000	Inorganic crystals
Materials Project (MPTS-52) [64]	Yes	~40 000	Inorganic crystals
ICSD [65]	Mixed (mostly verified)	>320 000	Inorganic crystals
CSD [66]	Yes	1.25 million	Organic and metal-organic crystals
COD [67]	Mixed (less filters)	>520 000	Organic, inorganic, metal-organic crystals
AFLOW [68]	Computational	>3.9 million	Inorganic crystals
Alexandria [69]	Computational	>5 million	Molecules, inorganic, and low-dimensional materials
JARVIS-DFT [70]	Computational	~41 000	0D to 3D materials
OQMD [71]	Mixed (some verified)	>1.3 million	Inorganic crystals
C2DB [72]	Mixed (some verified)	>4000	2D materials
2DMatPedia [73]	Computational	>6000	2D materials
Carbon-24 [74]	Computational	~10 000	Carbon structures
Perov-5 [75, 76]	Computational	~19 000	Perovskite structures
OC20 [77]	Computational	>1.2 million	Catalysts
NEMAD [78, 79]	Yes	~67 000	Magnetic materials
SuperCon [80]	Yes	~33 000	Superconductors
3DSC [81]	Yes	~9150	Superconductors
NOMAD [4, 82, 83]	Mixed (some verified)	>19 million	A wide range of materials

previous datasets, the structures considered are not necessarily thermodynamically stable, so they might require further pre-selection protocols to filter the experimentally synthesizable and stable structures.

For catalysts, the OC20 dataset consists of 1 281 040 DFT relaxations (264 890 000 single-point evaluations) across a wide swath of materials, surfaces, and adsorbates (nitrogen, carbon, and oxygen chemistries) [77]. The Northeast Materials Database (NEMAD) is a database that consists of 67 573 magnetic materials collected via LLM-based automated scientific data extraction from experimental literature [78, 79]. Supercon focuses on high-T_c oxide superconductors, metal-alloy superconductors, and organic superconductors, with data extracted from academic papers and other sources [80], complemented by 3DSC, a dataset of superconductors based on Supercon but including matched crystal structures from the ICSD and Materials Project [81].

4 | Crystal Structure Related Tasks, Generation Approaches, and Machine Learning Models

Crystal structure–related tasks can be broadly categorized into two main types: property prediction and structure generation. Property prediction is a "forward task" which involves esti-

mating characteristics such as formation energy, band gap, or mechanical stability for a given crystal structure, using methods ranging from classical first-principles approaches like DFT to modern predictive ML models such as GNNs and large language models (LLMs). When combined with large-scale computational screening, predictive approaches enable the rapid exploration of vast materials libraries. In contrast, structure generation is an "inverse task", focusing on designing or generating new crystal structures that satisfy specific criteria. This includes tasks such as predicting stable crystal structures from chemical composition and generating entirely new crystal structures from scratch. Crystal structure prediction (CSP) has long been of paramount importance for understanding materials' properties and designing them. It consists of determining the most stable atomic arrangement for a given chemical composition. *De novo* generation aims to go a step further by creating entirely new crystal structures, potentially with novel compositions, from scratch. While CSP is constrained to exploring structural possibilities within a fixed chemical formula, typically relying on search algorithms and energy evaluations to identify low-energy configurations, *de novo* generation is inherently open-ended: it seeks to design both composition and structure simultaneously, often guided by desired property targets rather than stability alone. This makes *de novo* approaches particularly powerful for inverse design, where the goal is to discover materials with tailored functionalities beyond those found in existing databases.

While property prediction of crystal structures is mainly achieved through predictive machine learning, structure discovery generation tasks such as CSP and *de novo* generation are conventionally tackled using search and optimization methods, or, more recently, generative models.

4.1 | Design by Combining Property Prediction with High-Throughput Screening and Search

4.1.1 | Forward Prediction Methods

Property prediction for crystal structures has been traditionally achieved through quantum mechanical simulation methods such as DFT, which provides energy evaluations and property predictions from first principles. However, DFT is computationally expensive and scales poorly with system size and time scale, making exhaustive searches impractical for large and complex material spaces.

As computational power grew, predictive ML models began to complement DFT-based methods. Two (interconnected) main directions developed over the last decade, namely general-purpose property prediction models which map crystal structures to materials properties, and more specialized MLIPs which map specific geometric configurations of atoms to energies and forces. Both directions are highly related in terms of underlying ML methods, but differ in the exact task definition and in the datasets and applications. GNNs such as CGCNN [58], MEGNet [85], ALIGNN [86], and co-NGN [59] have also shown strong performance in predicting crystal properties directly from structural inputs. Furthermore, LLMs such as MatBERT [87] and LLM-Prop [88] have also been explored to predict properties from text-based descriptions of crystal structures, capturing symmetries and site information that conventional models may miss, but not (yet) outperforming GNNs in prediction accuracy. MLIPs [89] are trained to predict the total energy and atomic forces of three-dimensional atom configurations of both periodic and non-periodic structures. Notable architectures include MACE [90, 91], M3GNet [92], and CHGNet [93], which have enabled fast and accurate evaluation of large systems.

Both types of predictive models substantially accelerate property estimation compared to quantum chemical calculations and experiments, and thus enable fast screening, geometry relaxation, and property prediction across large datasets without relying solely on expensive DFT calculations or even experiments.

4.1.2 | High-Throughput Virtual Screening and Search-Based Generation

Bridging property prediction and structure discovery, high-throughput virtual screening serves as a strategy that leverages predictive models to explore large materials libraries, typically defined in advance through combinatorial enumeration of possible compositions and structures.

Computational materials design dates back to seminal work in the 1990s on limited topology methods, such as the cluster expansion [94–96], and subsequently, with increasing computational power in the 2000s and 2010s, to screening approaches that broadened the candidate list in intelligent ways [97, 98], ultimately leading, to the best of our knowledge, to the first truly novel materials discovered using computational approaches [99, 100].

The introduction of ML methods allowed a further speedup: by coupling fast property predictors with systematic or active-learning-based sampling of candidate materials, high-throughput virtual screening enables the efficient identification of promising materials and establishes a first step in a materials design funnel process, prior to more costly computational evaluation, multi-scale simulations, or experimental synthesis and characterization [15, 101].

High-throughput virtual screening has successfully led to the discovery of promising candidates across diverse material classes. For instance, computational high-throughput screening based on DFT identified a new electrocatalyst for the hydrogen evolution reaction (HER) [102]. Similarly, screening of perovskite metal oxides for solar light absorption reduced a vast space of 5,400 compounds to just 15 promising candidates using electronic structure calculations [76].

In some cases, computational predictions have been experimentally validated. Notable success stories include the identification of IRMOF-20, which demonstrated exceptional hydrogen storage capacities out of 5,309 metal-organic frameworks (MOFs) [103], and the discovery of high-performance crystalline Cu–S based thermoelectric materials from the ICSD database [104].

Once the materials libraries become more open and thus larger, an exhaustive enumeration and high-throughput virtual screening quickly becomes unfeasible. As a result, materials design must incorporate adaptive search and optimization approaches such as evolutionary algorithms.

CSP traditionally has been achieved experimentally by synthesizing the crystals and (single crystal) diffraction pattern characterization, which can be directly inverted to determine the structure. However, experimentation is time-consuming and challenging, which has led to a switch of focus towards computational approaches. Before the rise of end-to-end generative ML models, CSP relied on high-throughput screening, heuristic search strategies, and combinatorial enumeration. These methods often incorporate "semi-generative" principles by building structures based on evolutionary concepts or physical design rules.

Among the earliest and most widely used methods is USPEX [105], an evolutionary genetic algorithm that applies operators like heredity and mutation to evolve populations of structures toward low-energy minima, guided by DFT fitness evaluations. CALYPSO also takes an evolutionary approach founded on particle swarm optimization, using symmetry-constrained random initialization, bond characterization matrices to diversify the search, and refining candidates via DFT [106, 107]. AIRSS employs a conceptually simple random search approach, generating random unit cells that respect minimal physical constraints and rely on brute-force DFT

relaxations for selection [108]. Although these methods have been highly successful in identifying both stable and metastable crystal structures across diverse material classes, they remain computationally expensive, as they rely heavily on DFT for structural relaxation.

Alongside evolutionary and stochastic methods, rule-based structure generators have been developed for the rapid creation of specific crystal frameworks like metal-organic frameworks (MOFs) or covalent organic frameworks (COFs). ToBaCCo assembles crystalline frameworks from modular building blocks and topological templates, scaling the unit cell to fit metal clusters and organic linkers [109]. AuToGraFS follows a similar principle but includes a force field for in-situ optimization and allows for controlled introduction of defects, functional groups, or supercells [110]. These tools enable high-throughput generation of diverse hypothetical structures, although they typically require post-processing to assess synthesizability and stability.

Recent approaches have begun integrating predictive ML with traditional search strategies to improve efficiency. GNOME is one such hybrid system: it combines AIRSS-style random sampling with property prediction using GNNs [14]. GNOME applies active learning to iteratively improve the predictive model as more DFT labels are acquired, enabling efficient down-selection from millions of generated candidates. Its workflow has led to the identification of over 380 000 novel stable crystals and tens of thousands of materials with promising energy-related applications, now cataloged in the Energy-GNoME database [111].

Taken together, these DFT-based, ML-assisted, and heuristic search methods form the foundation of modern CSP and *de novo* generation workflows. Figure 3 provides an overview of this landscape, summarizing how property prediction, search-based generation, and end-to-end generative models are connected, along with literature spotlights. These approaches illustrate a transition from purely physics-based prediction to data-augmented search and screening. These frameworks generate large volumes of plausible structures, evaluate them with ML and DFT, and refine candidates toward realistic, stable materials. This layered approach—generate, predict, validate—sets the stage for the emergence of end-to-end generative models (Figure 4), which now aim to invert the process entirely: to learn how to directly propose structures with desired properties, a topic explored in the following sections. In Table 2, we provide a first, concise overview of major end-to-end generative model families for crystal structure design, highlighting key trade-offs in generation quality, inference speed, training stability, and likelihood estimation.

4.2 | Early End-to-End Generative Models

4.2.1 | Variational Autoencoders

Variational autoencoders (VAEs), introduced by Kingma and Welling, are generative models that integrate deep neural networks with principles of variational inference [25]. Their generative capacity has led to successful applications in domains like image synthesis [112, 113]. VAEs have been adapted for complex scientific tasks, including *de novo* molecular design and the generation of new crystal structures.

VAEs are composed of a probabilistic encoder, which maps inputs to a distribution in a latent space, and a decoder, which reconstructs data from its latent representation. The training objective minimizes both a reconstruction error and the Kullback–Leibler divergence, which promotes a smooth, continuous, and well-structured latent space. This probabilistic framework enables the resulting model to be highly effective at both reducing dimensionality and generating new data samples.

Early approaches often used direct, image-like representations. For example, iMatGen represents the unit cell with voxels on a fine grid [10]. The FTCP (Fourier-transformed crystal properties) framework refines this concept by using a Fourier-transformed representation, which introduces physical constraints to ensure the grid is mathematically invertible [114].

Instead of relying on direct spatial grids, other VAE models operate on a more abstract level, using descriptive parameters to define a structure. As an example, PCVAE represents a crystal's geometry using chemistry- and physics-informed descriptors such as the Bravais lattice and space group, and then decodes them to classify the space group and determine the lattice constants [115].

Building on the idea of a structured latent space, MagGen uses graph-theoretic analysis to quantify structural similarity, streamlining generation and enforcing design constraints [116]. Other methods focus on improving the physical realism of the generated crystals. WyCryst, for instance, produces more plausible structures by adding penalty terms to its model, ensuring the crystals adhere to both standard Euclidean symmetries and specific constraints within their allowed space groups [117].

Beyond these architectural innovations, new training strategies have also emerged to guide the discovery process. The evolutionary VAE for perovskite discovery (EVAPD) uses genetic algorithms to identify and rank the most promising candidates [118], while TL-VAE improves outcomes through target learning [119].

Finally, a more fundamental solution to handling complex crystal properties has been the development of graph-based VAEs. These models are naturally suited to representing structural symmetries, moving beyond both grid representations and abstract descriptors. A key example is the model developed by Winter, Noe, and Clevert, which is permutation-invariant by design, ensuring that symmetry is inherently respected throughout the generation process [120].

4.2.2 | Generative Adversarial Networks

Generative adversarial networks (GANs), first proposed by Goodfellow et al., are generative models that consist of two neural networks—a generator and a discriminator—engaged in a minimax game [27]. The generator aims to produce synthetic data that mimics real samples, while the discriminator learns to distinguish between real and generated data. Through this competitive interplay, both networks iteratively improve until the

a Overview of Machine Learning Models for Crystal Structures
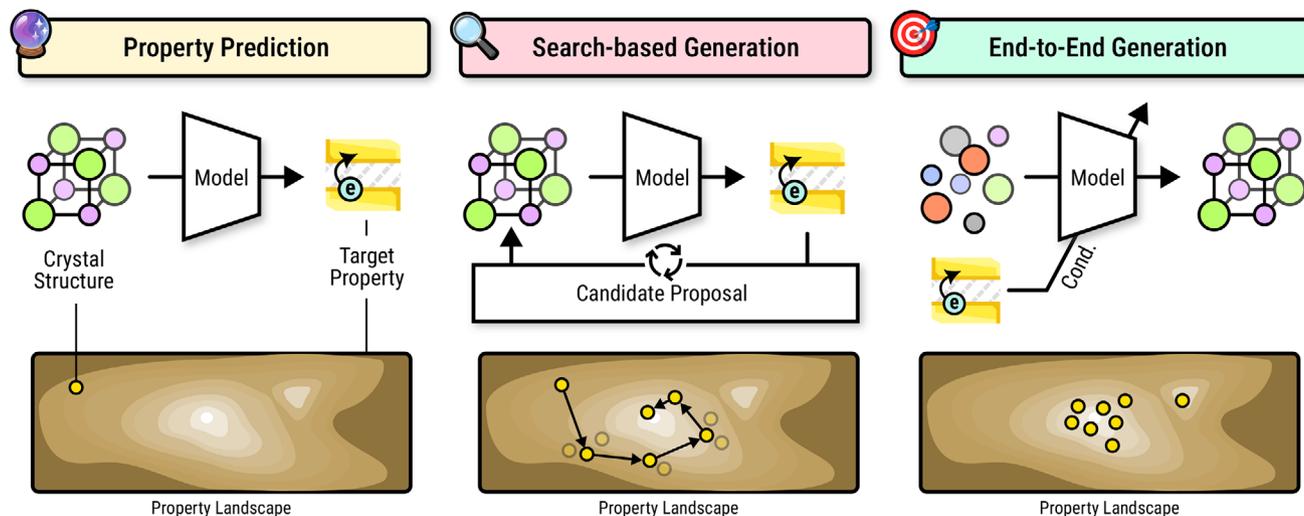
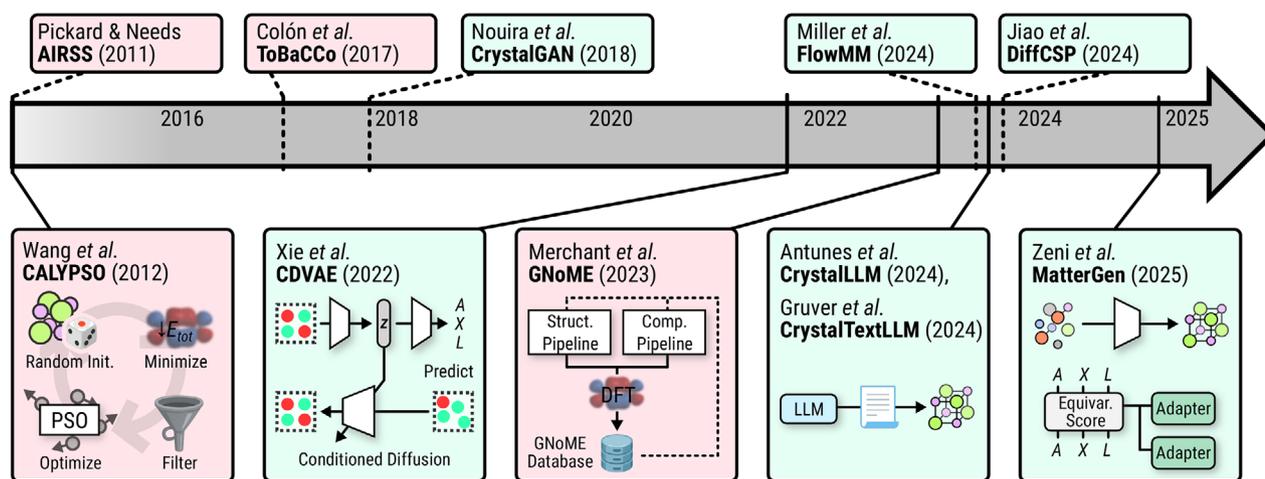**b** Literature Spotlights for Crystal Structure Generation

FIGURE 3 | (a) Overview of machine learning models in crystal structure generation. Left to right: Using machine learning models for simple property prediction, predictive models in conjunction with candidate proposal strategies for search-based generation, and end-to-end generation from a property conditional distribution. (b) Literature spotlights for crystal structure generation.

generator produces data that the discriminator can no longer reliably differentiate from real examples. This adversarial training allows GANs to model complex data distributions without relying on explicit probability functions. Due to their ability to produce high-quality synthetic data, GANs have achieved notable success in tasks such as image generation, style conversion, and data enhancement [121]. In addition, GANs have also been adapted for advanced scientific challenges, such as the design of novel molecules.

CrystalGAN is the first model to adapt GAN architectures to the synthesis of inorganic crystal structures [29], demonstrating that generative models could learn structural chemistry rules sufficiently well to produce stable crystals beyond the prototypes included in the training data.

Building upon this foundation, later research focused on guiding the generative process toward specific material properties. ZeoGAN introduces the concept of inverse design by conditioning

generation on adsorption energy grids [122]. In parallel, some models explore the advantages of a two-stage strategy: first generating plausible compositions, and then predicting their corresponding crystal structures. For example, the composition-conditioned crystal GAN (CCCGAN) generates crystals consistent with prescribed chemical formulas and validates their stability using DFT calculations [30], while MatGAN focuses exclusively on producing chemically reasonable and charge-balanced compositions [123].

Efforts to extend generative capabilities further have led to CubicGAN [124], which scales the approach to a dataset of over 370 000 ternary materials and successfully produces hundreds of new phonon-stable cubic structures. Meanwhile, the constrained crystals deep convolutional GAN (CCDCGAN) incorporates physical constraints directly into its training process by penalizing high-energy configurations [62]. This integration of energy-informed objectives marked a broader trend toward property-constrained generation, ensuring
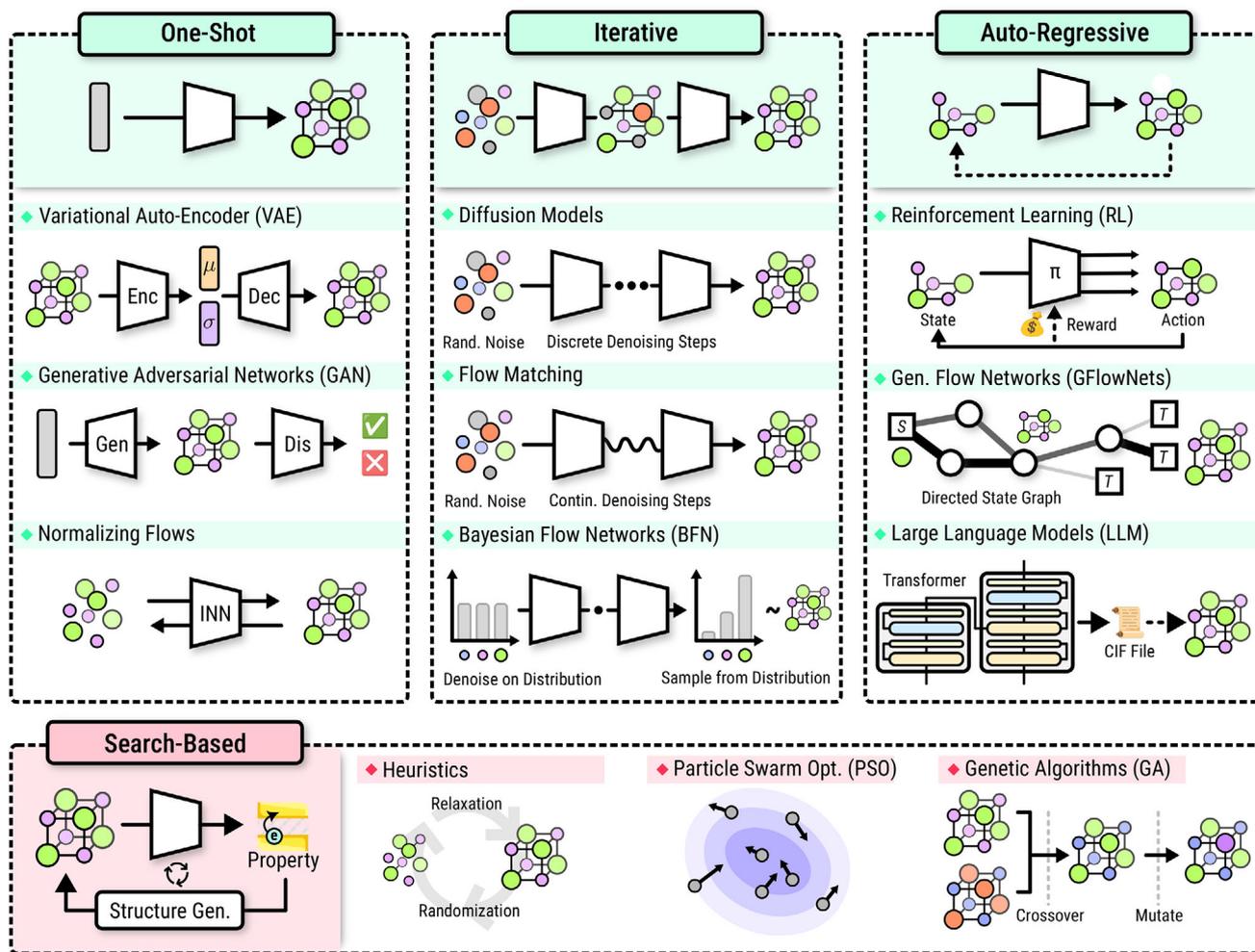
FIGURE 4 | Overview of different generative modeling approaches, divided into four conceptual categories. One-shot variational approaches directly generate crystal structures based on a latent space representation. Iterative methods repetitively refine the generation outcome. Auto-regressive methods iteratively construct samples from individual components of the crystal - or from tokens in the case of large language models. Search-based approaches employ property prediction models in conjunction with candidate proposal strategies, based on heuristics and/or physical intuition.

that chemical realism is maintained alongside structural novelty.

More recent advances have increasingly integrated crystallographic principles into GAN architectures to enhance the physical validity of generated structures. The physics-guided crystal generative model PGCGM explicitly encodes space-group symmetries, addressing a common limitation of unconstrained models that tend to produce low-symmetry or distorted structures [125]. Extending this paradigm, the crystal generative Wyckoff GAN (CGWGAN) [126] incorporates asymmetric units and Wyckoff positions to guarantee that generated templates satisfy both symmetry and atom-count requirements from the outset.

4.2.3 | Reinforcement Learning

Reinforcement learning (RL) is a branch of ML where an agent learns by interacting with its environment and receiving rewards or penalties. In the context of crystal generation, RL can help navigate the vast design space by learning strategies that lead to stable structures or materials with desired properties. Over the past few years, RL applications in materials science have evolved from accelerating structure searches to directly designing crystals with specific functionalities.

Early applications of RL to crystal problems focused on accelerating CSP. Zamaraeva et al. developed RL-CSP to dynamically choose moves such as atom swaps or lattice perturbations during CSP [127]. Compared to fixed, hand-tuned policies, this adaptive approach reduced the number of calculations required by up to 46 %, showing that agents could "learn" how to reach stable structures more efficiently. Similarly, Meldgaard et al. applied deep RL to surface reconstruction, where an agent proposed atomic placements for TiO_2 and SnO_2 surfaces, guided by DFT energies [128]. This demonstrated that RL can handle not only bulk structures but also the complex problem of predicting surface configurations.

RL has also been applied in inverse design to generate stable structures with target properties. For example, using offline RL as a way to guide crystal design, where agents trained on past data could generate new compositions whose band gaps matched desired targets [129]. By learning from past data instead of online

TABLE 2 | Qualitative comparison of model families for crystal structure generation. The entries provide approximate assessments intended to guide intuition rather than quantitative ranking. Actual performance depends strongly on representation, sampling steps, model and batch size, hardware, implementation details, and related experimental choices.

Model	Generation quality	Inference speed	Training stability	Likelihood estimation
Variational autoencoders	Medium	Fast: One-shot generation	Often hard to tune	Only ELBO
Generative adversarial networks	Medium	Fast: One-shot generation	Often prone to minimax oscillations	Not available
Diffusion models	Competing for SOTA	Typically slow: Many iterative steps	Usually stable: Score matching objectives; can be sensitive to noise scheduling	Possible but very expensive
Flow matching		Medium: Typically fewer steps than diffusion	Usually stable: Flow-matching objectives	Possible but very expensive
Bayesian flow networks		Medium to high: Typically fewer steps than diffusion	Usually stable: Bayesian update framework	Not reported in literature
Large language models		Typically slow to medium: Token-by-token generation	Usually stable: Standard cross-entropy objective	Only auto-regressive likelihoods

calculations, the model suggests chemically realistic structures while moving toward property goals. This showed that RL can shift from exploration to purposeful design.

Soon after, RL was extended to multi-objective optimization. RL agents were combined with predictive models to propose new inorganic materials that satisfied multiple criteria at once, such as band gap, mechanical strength, and even synthetic accessibility. By balancing these objectives, RL-based methods began to resemble practical design tools, addressing the reality that materials must perform across several metrics, not just one [130].

Most recently, Chen, Guo, and Schwaller combined RL with diffusion-based generative models for crystal structures [131]. Here, crystal generation was treated as a multi-step decision process within a diffusion model, and RL rewards guided the model toward crystal structures with target properties while maintaining diversity. This hybrid approach achieved higher success rates in goal-directed design than either RL or diffusion models alone, showing the potential of combining generative modeling with adaptive learning, as demonstrated recently in other works [132, 133].

While early end-to-end generative approaches, such as VAEs, GANs, and RL, proved useful in early CSP and *de novo* crystal design, newer architectures, presented in the following sections, surpass them by better capturing physical symmetries and material constraints, among other reasons.

4.3 | Recent End-to-End Generative Models

4.3.1 | Diffusion Models

First introduced by Sohl-Dickstein et al., diffusion-based models have rapidly become the dominant approach for image, video, and audio generation [134, 135]. More recently, diffusion models have gained increasing attention in graph-based data generation and are now considered the state-of-the-art method for generating molecular structures [136, 137].

The essential idea behind diffusion-based models is to systematically and gradually destroy structure in a data distribution, and then to learn to reverse this process to restore the data. This framework, while allowing several formulations such as Denoising Diffusion Probabilistic Models [35], Score Matching with Langevin Dynamics [138], and the generalized Stochastic Differential Equations [139], invariantly consists of two parts: a forward diffusion process, where noise is progressively introduced into real data until it approximates a chosen prior distribution, and a reverse generation process, where a model (typically parametrized with a neural network) is trained to recover the data distribution from noise.

The application of diffusion models to crystal structure generation was first introduced by Xie et al. with their crystal diffusion variational autoencoder (CDVAE) [39]. This model combines a VAE with a denoising diffusion model. Specifically, CDVAE uses a VAE to encode the joint distribution of composition, lattice, and the number of atoms. A diffusion-based model then learns the

joint distribution of atomic coordinates and types, conditioned on the latent code from the VAE. CDVAE explicitly encodes the interactions across periodic boundaries, ensuring it respects permutation, translation, rotation, and periodic invariances inherent in crystal structures. Subsequent works [140, 141] have extended the model by enabling the conditional generation of crystal structures with target properties.

Jiao et al. proposed DiffCSP, which introduces several novelties into diffusion-based crystal generation [142]. DiffCSP is the first generative model to use diffusion for jointly modeling atom type, lattice parameters, and atom coordinates, resulting in a more accurate representation of crystal geometry. The model operates on fractional coordinates, and it inherently encodes periodicity by employing the wrapped normal (WN) distribution [143]. Additionally, DiffCSP uses the Equivariant Graph Neural Network (EGNN) as a denoising network [60], thereby ensuring E(3) invariances for both lattice parameters and fractional coordinates. Many other frameworks in the literature have later adopted EGNN.

A significant advancement in the capabilities of diffusion-based generative models for crystals is the introduction of symmetry-aware models. Several recent models in the literature explore various approaches to ensure that the generated structures belong to specific space groups. Built on their previous work, Jiao et al. introduced DiffCSP++, which conditions its diffusion process on space groups and leverages predefined structural templates extracted from the training data [144]. This method ensures high-fidelity symmetry but potentially limits the discovery of novel structures beyond its training data. Levy et al. offered an alternative approach [64]. SymmCD generates crystal structures by jointly denoising a compact representation of the asymmetric unit—comprising lattice parameters, atom types, fractional coordinates, and site symmetries—while explicitly preserving space group constraints throughout the diffusion process. This allows for realistic and diverse space group distributions while avoiding reliance on templates. In contrast, WyckoffDiff inherently preserves space-group symmetry by directly diffusing over Wyckoff positions and site occupancies, thus encoding symmetry by construction rather than learning it [145].

Given the flexibility and sample quality demonstrated by diffusion-based approaches to generating crystalline structures, several specialized models have been developed for specific classes of materials. Diffusion models have been successfully applied for generating, among others, coarse-grained MOFs [146], new families of hypothetical superconductors [147], and zeolites with targeted properties [148].

Introduced by Zeni et al., MatterGen exemplifies the practical capabilities of diffusion models for crystal generation [149]. It applies diffusion jointly to the entire crystal representation, including lattice parameters, fractional coordinates, and atom types. MatterGen also incorporates adapter modules for fine-tuning the base model to achieve desired chemical compositions, symmetries, and scalar properties, demonstrating its effectiveness across magnetic, electronic, and mechanical domains. The model was trained on the large-scale Alexandria dataset [69], and its generated structures have undergone partial experimental validation, with usable code and pre-trained weights made publicly available.

Finally, Chemeleon represents a hybrid diffusion framework that combines text-conditioned generative modeling with contrastive representation learning, mapping textual and crystal graph embeddings into a shared latent space [150]. Building on this, a more recent hybrid approach extends Chemeleon's text-conditioned diffusion backbone by incorporating a reinforcement learning (RL) module that guides latent diffusion sampling through multi-objective rewards, thereby coupling diffusion-based generation with RL-driven exploration of the latent space [133].

4.3.2 | Normalizing Flows

Normalizing flows relate a complex target distribution to a simpler base distribution (often Gaussian or uniform) via an invertible function.

Discrete normalizing flows parametrize an invertible neural network, balancing expressiveness with the need for tractable Jacobian determinants. The Jacobian determinant is used to calculate the likelihood of each sample using the change-of-variables formula, which allows maximum-likelihood training [31, 151–154].

In the realm of crystal structure modeling, discrete normalizing flows are often employed as Boltzmann generators that model the full Boltzmann distribution for a given composition [155]. This is in line with the CSP task, though instead of only generating stable structures for a given composition, the full thermodynamic Boltzmann distribution is modeled. This can be used to determine (formation) free energies of crystal structures, which typically requires access to exact likelihoods. Since most other types of generative models lack straightforward access to likelihoods, discrete normalizing flows are a natural choice.

Wirnsberger et al. used normalizing flows to calculate free energy differences between two thermodynamic states [156]. The flow acts as an invertible mapping between the states to increase overlap compared to traditional free energy perturbation methods, yielding improved free energy estimates. Furthermore, Wirnsberger et al. trained a normalizing flow to model the Boltzmann distribution of Lennard-Jones crystals and cubic ice I to obtain absolute Helmholtz free energy estimates [33]. Similarly, Ahmad and Cai applied normalizing flows to compute Gibbs free energies of perfect diamond-cubic Si crystals and monovacancy defect formation energies [157]. Köhler et al. introduced rigid-body flows to model positions and orientations of molecules in molecular crystals [34]. They used this approach to calculate the free energy difference between two thermodynamic states of the hydrogen-ordered crystal phase of water. Schebek et al. developed conditional normalizing flows to map a single reference equilibrium distribution across a range of temperatures and pressures for accurate phase diagram prediction [158]. Applied to a Lennard-Jones face-centered cubic crystal, their model accurately predicted the solid–liquid coexistence line while reducing the required sampling compared to traditional methods.

While discrete normalizing flows transform a prior distribution to a target using a finite number of invertible neural network layers,

continuous normalizing flows are the continuous generalization thereof [159, 160]. Here, samples from a prior distribution are transformed by integrating an ordinary differential equation, moving the samples along a time-dependent vector field to match the target distribution. Analogous to discrete normalizing flows, the likelihood of a sample can be obtained by a continuous generalization of the change-of-variables formula. However, this requires the evaluation of the vector field divergence in each integration step, which is often prohibitively expensive. This makes them less suitable for calculating free energies compared to discrete normalizing flows.

However, continuous normalizing flows are typically more expressive compared to discrete normalizing flows, as they do not suffer from the same architectural constraints. The conditional flow matching training objective allows efficient simulation-free training of continuous normalizing flows [161, 162], similar to score matching in diffusion models. This makes continuous normalizing flows a promising competitor to diffusion models, which is attributed to the fewer integration steps they often require during sampling due to more efficient trajectories [161]. Recently, Riemannian flow matching was introduced, which extends flow matching to general manifolds [163]. Based on this approach, Miller et al. introduced FlowMM, a crystal generation framework that uses Riemannian Flow Matching to handle the periodic boundaries inherent in crystal modeling [164]. Similarly, Kim et al. used Riemannian flow matching to model MOFs, treating the metal nodes and organic linkers as rigid bodies to reduce the dimensionality of the problem [165].

4.3.3 | Generative Flow Networks

Generative flow networks (GFlowNets) represent a class of generative models that learn a stochastic policy to sequentially construct compositional objects, such as crystal structures, by traversing a directed acyclic graph from an initial state to a terminal state [166, 167]. Unlike traditional generative models that aim to maximize likelihood or minimize reconstruction error, GFlowNets are designed to sample diverse objects from a set of terminal states with probabilities proportional to a user-defined positive reward function. This proportionality-to-reward sampling enables efficient exploration of vast search spaces, facilitating the discovery of multiple high-reward candidates rather than converging on a single optimum. Training objectives such as trajectory balance are commonly used to ensure consistency between the forward generation policy and an auxiliary backward policy [166, 167].

Crystal-GFN, introduced by Hernandez-Garcia et al., is a foundational GFlowNet model for inorganic crystal structures [168]. It employs a sequential generation scheme, first selecting the space group, then the chemical composition, and finally the lattice parameters of a unit cell. This approach allows for the systematic incorporation of fundamental crystallographic constraints, such as charge neutrality and compatibility between composition, space group, and lattice geometry, at each step. The model is guided by a reward signal from a proxy model predicting formation energy, enabling it to sample structures proportional to their predicted stability.

Building on this, Nguyen et al. proposed a symmetry-aware hierarchical architecture for flow-based traversal (SHAFT) [169]. SHAFT refines the hierarchical generation process by initially selecting a high-level crystallographic prototype (space group), then determining consistent lattice parameters, and finally adding atoms into the unit cell while adhering to these constraints. A key innovation is the explicit exploitation of crystal symmetries; once a space group is chosen, symmetric atomic positions are automatically populated, reducing the search space complexity. Guided by a physics-informed reward function incorporating formation energy and symmetry preferences, SHAFT has shown superior performance in generating valid, diverse, and stable crystal structures compared to non-hierarchical GFlowNets and diffusion-based VAEs.

In another work, Cipcigan et al. proposed MatGfn, which has been used to design reticular materials like MOFs and COFs for carbon dioxide capture, by building string sequences into CIFs based on a gravimetric surface area reward [170].

4.3.4 | Bayesian Flow Networks

In 2023, Graves et al. introduced Bayesian flow networks (BFNs), a novel class of generative neural networks [171]. Similar to diffusion models, BFNs generate new samples through an iterative process. However, they differ by explicitly modeling the underlying data distributions rather than transforming or denoising the data directly. This approach enables a fully continuous and differentiable generation process, even for discrete data. Generation starts from independent, simple prior distributions (e.g., standard normal for continuous data or uniform categorical for discrete data), which are iteratively updated via Bayesian inference using noisy samples and refined through a neural network that captures interdependencies among variables. The final sample is drawn from the resulting distribution after the last step. By updating the parameters of the underlying distributions instead of individual data points, BFNs generate high-quality samples in just a few iterations, offering a more efficient alternative to traditional diffusion and flow-based models [172, 173].

Wu et al. introduced BFNs for crystal structures by extending the Bayesian flow to non-Euclidean manifolds to account for the periodic nature of fractional coordinates [172]. Ruple et al. implemented a symmetry-aware BFN that operates on the asymmetric unit representation, ensuring efficient generation while respecting crystal symmetries [173].

4.3.5 | Transformer Models and Large Language Models (LLMs)

Transformer models, originally introduced in the context of natural language processing, are based on self-attention mechanisms that allow them to learn complex relationships within sequential data. An LLM is based on variants of the transformer model [40], trained on vast corpora of text data to learn the statistical structure of language. LLMs and transformer models can be adapted to material science, either by fine-tuning general-purpose LLMs on text representations of crystal structures or

by using transformers with tokenization schemes specific to crystal structure text representations. The latter approach has shown greater promise in current benchmarks. In this context, tokenization refers to converting structured crystal data (such as CIF files) into a linear sequence of discrete symbols (tokens) that a language model can process. Standard NLP tokenizers (e.g., byte-pair encoding) are often ill-suited to scientific formats, prompting the development of domain-specific tokenizers that better reflect the grammar and symmetries of crystallographic data. This allows full flexibility to capture domain-specific structure and symmetries (e.g., SE(3) invariance), but requires large datasets and significant computational resources.

One of the earliest such methods, XYZTransformer [43], adopts a GPT-style architecture [174], leveraging tokenization derived from compacted CIF files and incorporating data augmentation to learn crystallographic invariances. Similarly, CrystaLLM [44] builds on the nanoGPT architecture [175], employing CIF-specific tokenization and unsupervised pretraining on 2.3 million randomly sampled CIF files. Mat2Seq, also builds on the nanoGPT architecture like CrystaLLM, introduces SE(3)-invariant sequence representations to capture the rotational and translational symmetries inherent in crystal structures [176]. CrystalFormer [177] and WyFormer [178] adopt autoregressive Transformer architectures that explicitly incorporate crystallographic symmetries by conditioning on the space group and applying tokenization schemes based on Wyckoff positions, which also contributes to improved inference speed.

On the other hand, fine-tuning refers to adapting an already pretrained LLM, such as LLaMA or Mistral, to a specific domain by training it further on domain-specific data (e.g., CIF files). This is more efficient, as the model retains useful general representations learned from broad text corpora, and only needs to adjust its weights for the new task. Techniques such as LoRA (Low-Rank Adaptation) and QLoRA (Quantized LoRA) enable parameter-efficient fine-tuning by modifying only a small subset of the model's layers. This reduces computational cost while retaining performance. Crystal-text-LLM [179] marks the first model to fine-tune a LLaMA-2 [180] LLM on CIF files using LoRA. Building on this, FlowLLM [181] and CrysLLMGen [182] combine generation via the fine-tuned LLM with subsequent structural refinement through FlowMM [164] and DiffCSP [142] optimization, respectively. In a similar vein, Choudhary fine-tuned the Mistral-AI model [183] on CIF files using LoRA, while Mohanty et al. fine-tuned the Llama-3.1-8B model with QLoRA to generate structure descriptions [132, 184].

5 | Practical Considerations

5.1 | Constraining and Conditioning Models

While unconstrained *de novo* generation of crystal structures is an interesting challenge, many practical materials discovery tasks demand a more targeted approach: generating structures that satisfy predefined constraints or exhibit specific properties. Targeted generation is central to inverse materials design, where the goal is not merely to explore structural space but to find candidates optimized for function and feasibility. This section outlines possible scenarios in targeted generation, also illustrated in Figure 5.

A first step toward targeted generation is to incorporate structural or compositional constraints into the generation process. Several recent approaches have demonstrated that it is possible to generate crystal structures constrained to a desired space group by leveraging its symmetry operations during generation and explicitly conditioning the model on the target group [64, 144, 173]. This is valuable in real-world settings, as many families of materials, such as perovskites or zeolites, predominantly adopt specific space groups.

Another important constraint is chemical composition. By fixing the atom types, the task reduces from *de novo* generation to CSP. For diffusion models and flow matching, fixing the atom types during generation is straightforward, enabling the model to sample from the distribution of stable configurations for a given composition. Also, inpainting generation based on known host structures is possible [185].

Lastly, given applications define desired target properties, which can be achieved by conditioning models to generate crystal structures that fulfill them, which will be discussed later.

5.1.1 | Symmetry-Constrained Generation

Symmetries play a fundamental role in the modeling of crystal structures (Figure 1). Unlike general molecular systems, crystals must satisfy not only the Euclidean invariances of three-dimensional space, i.e., translations, rotations, and reflections, collectively described by the E(3) group, but also the more restrictive symmetries defined by their specific space groups. Early generative approaches typically only accounted for E(3) invariance but neglected space group symmetries [39, 142, 164]. As a result, these models frequently generated crystals with low symmetry and made conditioning on a target space group challenging or infeasible [64]. To address this limitation, more recent models explicitly incorporate space group information into the generation process. These methods either constrain the placement of atoms to valid Wyckoff positions consistent with the desired space group [144], or focus on generating only the asymmetric unit, which is then expanded into the full periodic crystal during post-processing [64, 173, 177, 178]. This inductive bias improves the physical plausibility of the generated crystal structures and enables the restriction of the generation to specific space groups. Furthermore, models that generate only the asymmetric unit are more computationally efficient and scalable for both iterative and autoregressive architectures, as they operate on a reduced representation while ensuring full structural symmetry.

5.1.2 | Property-Conditioned Generation

In practical, application-driven materials discovery, structural or chemical constraints are not sufficient. Often, the goal is to find materials that possess target properties tailored for specific applications.
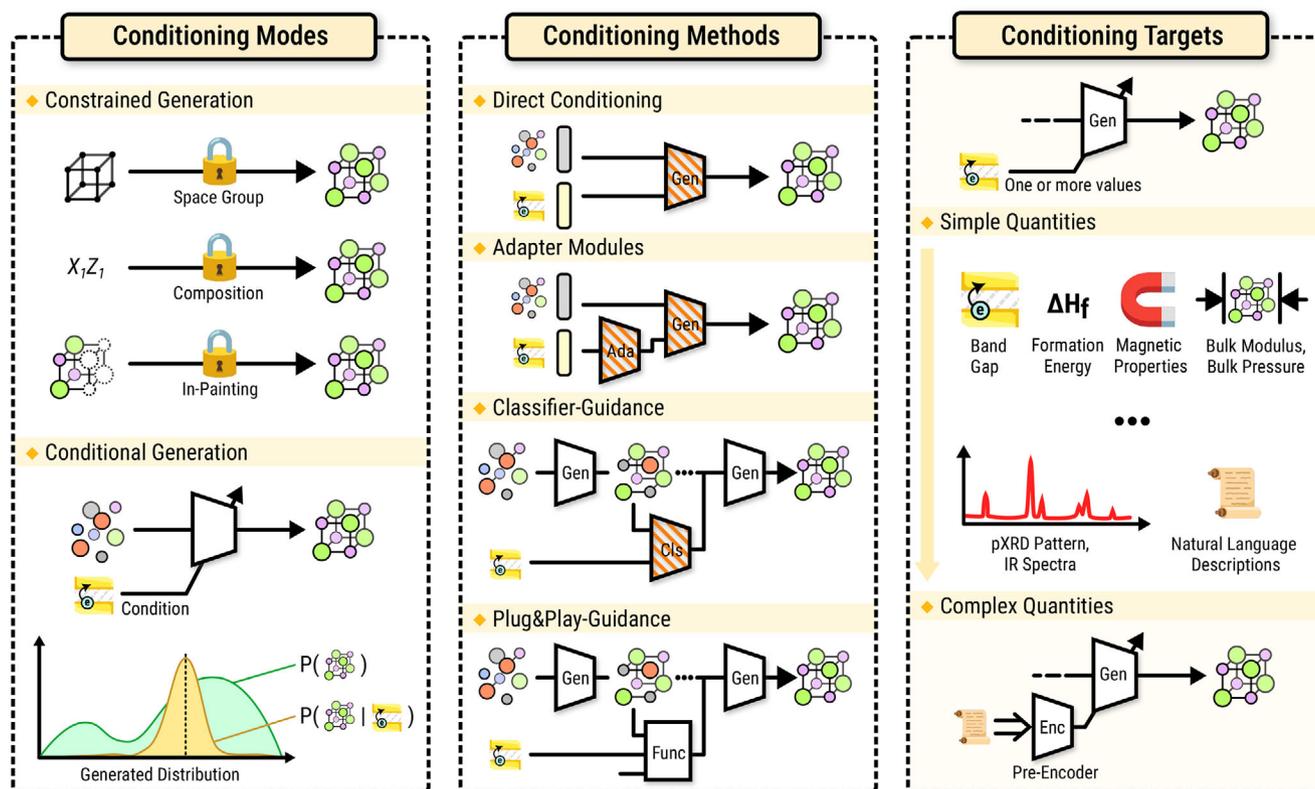
FIGURE 5 | Overview of targeted crystal structure generation. Left: Difference between constrained generation and property-conditioned generation. Middle: Architecture integrations of different conditioning methods. The hatched symbols represent components that are coupled to the specific property and which have to be retrained when switching properties. Right: example properties that can be used as conditioning targets for crystal structure generation. Valid conditioning targets range from simple, single-value properties, such as the band gap, to complex properties, including natural language property descriptions.

Enforcing property constraints during generation is difficult due to the complex relationship between structure and property. Instead, conditional generation is generally approached via soft conditioning, where the model learns to approximate the conditional distribution of structures given a target label. This is typically done by injecting a learned representation of one or more target properties into the architecture [140, 149, 173]. Classifier-free guidance can be used to tune the influence of the target condition [186, 187].

Directly conditioning a generative model requires retraining from scratch for every new target or property of interest, limiting the method's generality and practicality. To address this, MatterGen uses lightweight adapter modules that inject conditioning information additively into the layers of a pre-trained generative model [149]. This decouples the conditioning mechanism from the core architecture and allows models to be first trained generically, then fine-tuned for specific properties with relatively little additional data or compute.

There are many properties suitable for conditional crystal generation, and the choice largely depends on the specific application. First, scalar properties can be used for conditional crystal generation, e.g., electronic properties such as the band gap [140, 149, 173], physical properties such as the bulk modulus [149]/bulk pressure [141], magnetic properties [149], or the thermodynamic

stability given by the formation energy [140, 173]. Recently, Park, Onwuli, and Walsh introduced a text-guided crystal diffusion model, where the generation is conditioned on a text embedding based on the chemical composition and crystal system [188].

Beyond generating functionalized materials with targeted properties, conditional crystal generation has also been used to support experimental analysis. Here, a model generates possible structures conditioned on (non-scalar) experimental measurements such as powder diffractograms [189–192], XANES spectra [193], or IR spectra [194].

From a practical perspective, an interesting approach is to use guided generation without fine-tuning, where a powerful, pre-trained generative model can be guided at inference time. This plug-and-play guidance paradigm greatly expands accessibility, allowing domain experts to specify property targets without retraining or modifying the generative backbone. The first work in this direction has been performed [147, 177]. For example, Yuan and Dordevic guide the generation of superconductors based on reference compounds during inference to discover new families of hypothetical superconductors [147].

In the broader context of materials discovery, targeted generation promises to transform computational design workflows. Traditional screening approaches typically use a broad

proposal distribution that is subsequently filtered in multiple stages. Conditional generative models are promising candidates to accelerate this traditional screening approach, generating materials closer to the requested target and requiring less filtering. Continued progress will require not only architectural innovations but also robust benchmarks, property-aware datasets, and a deeper theoretical understanding of the trade-offs between control, diversity, and fidelity in (conditional) generative modeling.

5.2 | Usability: Software and Code Availability

Access to open-source, ready-to-use software is crucial for researchers looking to apply generative models in practice. To provide a practical starting point, we grouped a selection of implementations that we consider to be scientifically relevant and representative of recent advances in the field. This selection is not complete and balances multiple criteria beyond usual benchmark metrics, including code documentation and usability. Table 3 summarizes this selection of recent generative models for materials, including their associated datasets, code availability, checkpoints, and conditioning options.

5.3 | Speed and Computational Cost

Inference speed is a critical consideration in the practical deployment of generative models for crystal structure generation. A variation in inference time exists among methods, largely due to differences in architectural complexity, data representation, and sampling requirements [173]. Exact inference runtimes depend heavily on hardware configurations, so comparisons between different methods are limited to estimates at the level of orders of magnitude. For iterative models, the number of sampling steps is the primary determinant of runtime, since each step involves one or two computationally expensive neural network evaluations. In many reported settings, diffusion-based generative models exhibit long inference times because achieving state-of-the-art performance often involves hundreds to thousands of iterative denoising steps during sampling [64, 142]. Fine-tuned LLMs also operate at a high computational cost, with inference speeds varying over multiple orders of magnitude depending on model size. This variation reflects a tradeoff between computational efficiency and the quality of the generated structures [179], where the largest LLMs approach inference times comparable to those of diffusion models. Flow-based models operate at a similar speed to the smallest fine-tuned LLMs, as they generally require fewer sampling steps than diffusion models for generation, typically on the order of a few hundred steps [164]. Smaller transformer-based architectures employing symmetry-aware tokenization schemes achieve inference speeds approximately an order of magnitude faster than the smallest fine-tuned LLMs [177]. BFNs improve upon the number of required sampling steps, typically using fewer than 100 steps, reducing inference time by yet another order of magnitude compared to the transformer models. With the experimental setup reported for SymmBFN [173], using 100 sampling steps corresponds to generating a crystal structure in a few milliseconds on NVIDIA A100 GPUs.

TABLE 3 | Overview of selected repositories of generative models for materials. Repositories were selected based on competitive (near state-of-the-art) performance and the availability of well-documented, reproducible code and pretrained checkpoints.

Model	Method	Datasets	Code	Checkpoints	Conditioning
MatterGen [149]	Diffusion	Alex-MP-20, MP-20	GitHub	HuggingFace	Composition, Space group, Band gap, Bulk modulus, Magnetic density, HHI score, Energy above hull
DiffCSP [142]	Diffusion	Carbon, Perov, MP-20	GitHub	GoogleDrive	Composition, Formation Energy
DiffCSP++ [144]	Diffusion	Carbon, Perov, MP-20, MPTS-52	GitHub	GoogleDrive	Space group constrained. Composition
CrysBFN [172]	BFN	Carbon, Perov, MP-20, MPTS-52	GitHub	GoogleDrive	Composition
CrystalFormer [177]	Transformer	Alex-20, MP-20	GitHub	HuggingFace	Space group constrained. Dielectric FoM, Energy above hull (with reinforcement fine-tuning)

5.4 | Metrics and Theoretical Evaluation of the Models

In this section, we review the computational metrics used to benchmark generative models in CSP and *de novo* generation. These metrics, originally developed from an ML perspective, facilitate model comparison but only act as indirect proxies for the actual quality and validity of the generated crystal structures.

5.4.1 | Crystal Structure Prediction

In the CSP task, the model is evaluated on its ability to reconstruct the original crystal structures in the test set using only the atomic species as input. The evaluation metrics, first introduced by Xie et al. [39], are computed using the `StructureMatcher` class from the `pymatgen` library [195]. The `StructureMatcher` defines a match by aligning two structures, comparing all valid lattice transformations and atomic arrangements within specified tolerances. The match rate is defined as the fraction of target structures for which the model generates at least one matching candidate within a fixed number of attempts. Additionally, for all matched structures, the average root-mean-square displacement between corresponding atomic sites is reported to quantify the structural similarity between matched pairs.

However, these metrics have important limitations. Whether a generated structure is classified as a match depends strongly on the chosen tolerance parameters, which can vary between studies and substantially affect reported match rates. Additionally, this evaluation framework considers only exact or near-exact reconstructions of the test structures, potentially disregarding novel but thermodynamically stable configurations that share the same composition yet differ structurally from the reference. To address this limitation, it may be beneficial to additionally assess the thermodynamic stability of the generated structures, providing a more comprehensive evaluation that captures both reconstruction accuracy and the feasibility of the predicted configurations. Martirossyan et al. provided an in-depth analysis of the commonly used metrics in CSP and proposed refined adaptations to address their limitations [196].

5.4.2 | *De Novo* Generation

Generative models for crystal structures aim to generate materials that are (*i*) physically **stable**, (*ii*) **unique** in the sense that a model generates a diverse set of materials, and (*iii*) **novel**, meaning that the generated materials are not part of the training data.

While synthesizability can only be definitively confirmed through experimental validation, such validation is constrained by limited experimental resources. Consequently, many studies instead use thermodynamic **stability** as a proxy, typically evaluating the energy above the convex hull, a metric first adopted for generative models by Zeni et al. [149]. This has since become a standard benchmark for comparing the viability of output crystal structures for generative models, although numerous complementary evaluation methods are required to assess the quality of the generated candidates (see Section 5.5).

The convex hull represents the set of lowest-energy phases at all compositions within a chemical system [197]. A material's position relative to the convex hull indicates its thermodynamic stability: if it lies on the hull, it is considered thermodynamically stable, whereas a positive energy above the hull suggests it may decompose into a combination of more stable phases. To assess the energetic stability of a generated material, its structure is first optimised, and the final energy is computed using either DFT or ML approaches, such as MLIPs. While MLIPs offer significantly faster evaluations, they typically trade off some accuracy compared to DFT. Commonly used MLIPs for evaluation, as described in earlier sections, include M3GNet [92], CHGNet [93], MatterSim [198], and MACE-MP-0 [91]. A comprehensive overview of these and other MLIPs is provided by Jacobs et al. [199], while a systematic benchmark of their performance is presented by Riebesell et al. [200]. The stability rate denotes the proportion of generated materials whose final energies lie on ($E_{\text{hull}} \leq 0.0$ eV/atom), or near (often chosen to be around $E_{\text{hull}} \leq 0.1$ eV/atom), their respective convex hulls. These thresholds are widely used in ML studies for benchmarking generative models [142, 149, 164]. However, experimental workflows often employ more stringent (lower) cutoff values, as described in Section 5.5.

Using thermodynamic stability as an evaluation metric has inherent limitations: the reference convex hulls are often incomplete, particularly for chemical systems containing a large variety of elements. This epistemic uncertainty can result in approximate stability assessments, potentially leading to false positives. Furthermore, comparing stability rates across different sets of generated structures can suffer from statistical fluctuations unless a sufficiently large number of samples, typically at least 10 000, is evaluated. Performing such extensive stability checks directly with DFT is computationally prohibitive; therefore, many recent works first employ MLIPs for pre-screening, followed by high-accuracy DFT calculations for the most promising candidates [172, 179]. Variations in energy above hull thresholds and differences in the computational methods used to calculate the energies of predicted crystal structures often hinder the direct comparability of stability metrics across studies. Finally, a low energy above hull is necessary but not sufficient for synthesis, and therefore cannot serve as a comprehensive metric for the quality of generated structures (see Section 5.5).

In addition to energetic stability, a key objective is to produce a diverse set of novel structures that differ from those seen during training, rather than repeatedly generating identical or similar materials. Therefore, a structure is considered **unique** if no matching structure (determined, e.g., using the `StructureMatcher` class from the `pymatgen` library) exists within the set of generated structures. Similarly, a structure is considered **novel** if no matching structure is found in the training dataset. Both uniqueness and novelty metrics critically depend on the choice of similarity thresholds, which can significantly alter the results by either overestimating diversity or failing to distinguish genuinely distinct structures. A more detailed discussion of the limitations of novelty and uniqueness metrics is provided by Negishi et al. [201].

A generated structure is classified as a *S.U.N.* (stable, unique, and novel) material if it simultaneously satisfies the criteria for

stability, uniqueness, and novelty. The S.U.N. rate represents the proportion of generated structures that qualify as S.U.N. materials. Many studies also report a generation *cost* metric, either in terms of the number of function evaluations for iterative methods [164, 172], or the wall-clock time required to generate a single (S.U.N.) structure.

Earlier studies on generative models for crystal structures introduced a set of *proxy metrics* to evaluate the quality of generated materials [39]. While widely adopted, these metrics are generally less indicative of practical synthetic accessibility than stability-based assessments. They typically include validity checks based on charge neutrality and interatomic distance thresholds, coverage recall, and precision scores computed using structural fingerprints relative to the test set, and comparisons of property distributions, such as elemental diversity and density, between generated structures and the test set. However, because these metrics are usually computed on only 1000 generated structures, they are highly sensitive to sample selection and do not allow for fair method-to-method comparisons. Reporting the variance over multiple random subsets can help capture this uncertainty and improve the interpretability of benchmarking results [64].

5.5 | Post Generation Selection Workflows

Contemporary strategies for new materials development involve (generative) crystal structure prediction followed by experimental synthesis of the target material. However, as previously detailed, not all predicted structures may be realizable. Furthermore, synthesis procedure development is often cited as a bottleneck for new materials development, largely due to a lack of unified chemical synthesis theory [202–204]. Commonly used trial-and-error approaches limit efficient methodology establishment and firmly place synthesis as a resource-intensive field that necessitates scrutiny for synthesis campaign planning. Similarly, as a data-poor field, there has been very limited ML success for predictive chemical synthesis [205]. Despite fragmented theory and data collections, paradigms have emerged in (high-throughput) synthesizability screening [55]. Shown in Figure 6, synthesizability screening post generative crystal structure prediction generally employs filtering for target phase stability and physicality, and is followed by procedures to establish synthetic accessibility, before attempted experimental investigation.

5.5.1 | Evaluation of Target Stability

The minimum assessment for synthesizability is phase stability, agnostic of experimental method and condition. In other words, how likely is it that the generated structure is real? In Figure 6, we show a filtration schema for phase stability assembled from collections of target prediction and synthesizability literature. The criteria of (*i*) energetic stability, (*ii*) chemical heuristics, and (*iii*) structural stability are organized by approximate resource requirements, e.g., least to most computationally and time-intensive. As previously discussed in Section 5.4.2, many screening workflows search initially for favorable (thermodynamic) energetics, often evaluated by a negative formation energy (ΔH_f), energy (E_{hull}) with respect to the convex (composition-

energy) hull, or appearance as the most stable phase on a phase diagram (e.g. composition-temperature or chemical potential diagrams) [197, 206]. The former parameters are easy to integrate into automatic workflows leveraging computational materials databases of bulk crystalline inorganic solids, such as the Materials Project and the Open Quantum Materials Database [3, 71, 207, 208]. Nearly all observed bulk inorganic solids have negative formation energies and are either on the hull or are within $\approx$ 100meV/atom of the convex hull, making E_{hull} a useful tool [209, 210]. However, the threshold for metastability has been shown to be chemistry- and chemical system-dependent and thus prevents it from being a one-size-fits-all metric [210]. Screening for chemical heuristics in proposed structures ensures that critical chemical laws known to constrain ion oxidation state, polyhedra geometries, bonding orders, and coordination numbers are followed. Historically, the evaluation of chemical rules of thumb can be challenging to enforce due to the lack of formal, rigid rules governing bonding and structure for translationally invariant bulk, crystalline solids. However, a number of studies have successfully employed specific heuristic metrics to known and candidate structures, including the Pauling rules, the Crystal Orbital Bond Index, structure-type tolerance factors, oxidation state predictors, and the *mno* counting rules for intermetallic bonding [211–216]. Structural thermal dynamics and stability are important to understand proposed phase feasibility, and, in particular, towards assessing relative stability to a ground state atomic configuration. Structural stability can be measured through Monte Carlo (simulated annealing), molecular dynamics, and phonon calculations [204, 217–220], though computational and time resources are likely to be higher than for other screening metrics. It is important to note that many high-temperature metastable polymorphs may exhibit some lattice instability as the temperature is lowered and are therefore challenging to evaluate with high-throughput 0 K calculations, even when phonons are considered.

5.5.2 | Evaluation of Synthetic Accessibility

In the absence of retrosynthesis models, route-aware synthesizability screening in advance of experimental exploration aims to ensure efficient use of time and laboratory resources when targeting a candidate structure hypothesized to be stable. In Figure 6, a screening workflow outlines (*i*) method selection, (*ii*) method feasibility, and (*iii*) experimental condition prediction as the three steps that can be implemented within a predictive synthesis paradigm. While method selection is the first step in evaluating synthetic accessibility, this choice is often limited by individual expertise and equipment availability. Solid state (ceramic), hydrothermal, sol-gel, flux, sputtering, and deposition are among common preparative techniques [221]. Step *ii* shows options to forecast if the considered method is appropriate for a proposed structure. Domain experts often complete this evaluation with chemical and experimental heuristics and domain knowledge: indeed, Martin Jansen wrote in 2018: "Furthermore, the task of rationally identifying routes of syntheses to the solids predicted… has not yet been systematically tackled… this latter issue is less disturbing in practice, because a well-skilled and experienced chemist will find ways of synthesis for a material that he firmly
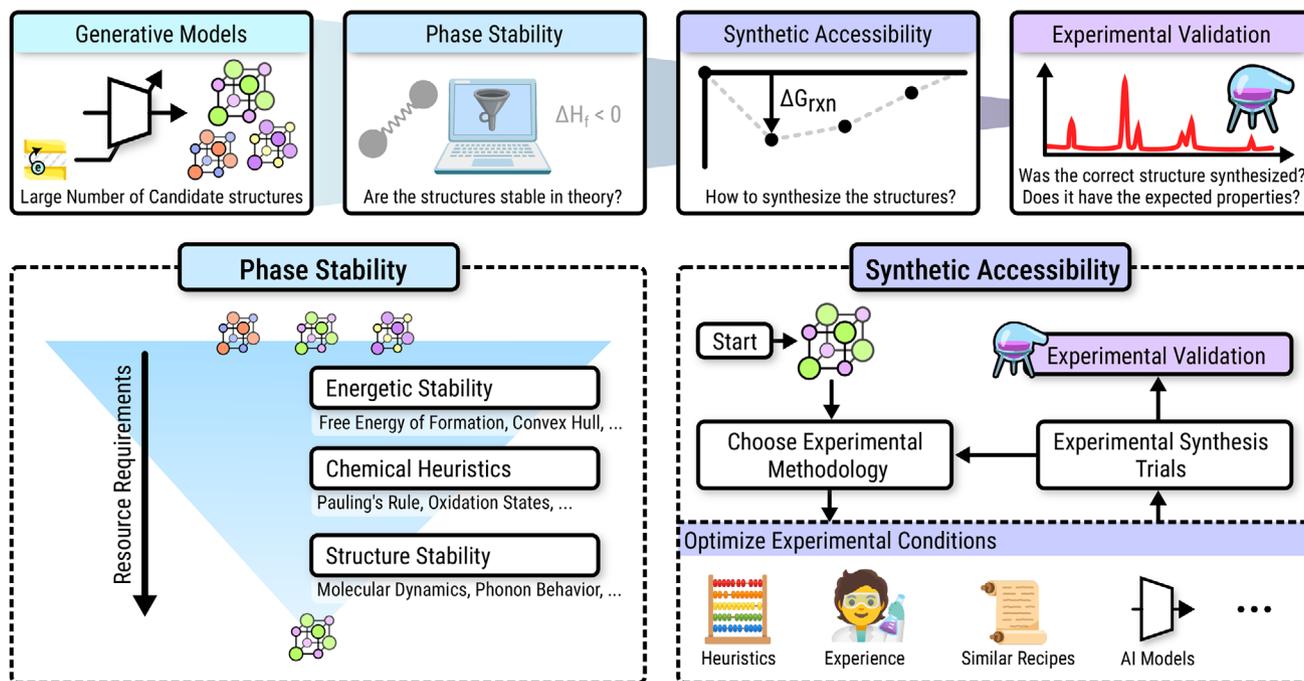
FIGURE 6 | Workflow for computationally guided experimental synthesis of generated crystal structures. Crystal structures created using generative models usually require several post-processing steps before experimental validation, which include evaluation of their phase stability (potentially in multi-step workflows depicted in the lower left), and their synthetic accessibility (lower right).

trusts to be capable of existence" [204]. Standardized models and descriptors to rationalize experimental methods choice are fragmented, but include creation of phase diagrams for assessment of stability at expected synthesis conditions [222], solubility tests for hydrothermal [223] and sol-gel synthesis [224, 225], choice of reaction temperature and precursor selection [226–228], and structure-based descriptors for crystal growth [228–232]. In the third step, suggestions for promising experimental conditions are evaluated. As this is an active area of research in synthesis science multiple models, programs, and theories have been proposed for different synthesis methods, including the use of simple reaction energies, more complex reaction networks ranking precursor systems by a cost function [222, 233]; assessing number of competing phases [228]; phase diagrams such as Pourbaix diagrams to understand precursor solute concentration, temperature and mineralizer pH for hydrothermal and flux based methods [223]; the substrate-target lattice match for epitaxial growth [229, 230]; and data-driven and machine learned models [234–238]. We guide interested readers to an excellent review for current approaches in models and programs for predictive solid state synthesis for battery materials [239]. Experimental investigation following suggested recipe(s) can be optimized via chemical intuition, exhaustive enumeration, or active learning approaches (i.e., Bayesian optimization) [240].

The lack of a theory of synthesizability is particularly challenging for generative structure models, as synthesis feasibility cannot be included as a constraint during the generation, but can at best be evaluated afterwards through explicit calculations of convex hulls and phase diagrams. In principle, requirements for the topology of the convex hull near the generated compound could be included during structure generation, though this is complex due to the non-local nature of hull energy, depending

on energies at many other compositions, and has therefore never been demonstrated as a feasible constraint in generative structure generation.

Despite the use of phase stability and route-aware synthetic accessibility screening, successful experimental realization is not assured [202]; exhausting enumerated synthesis conditions, sometimes for multiple synthetic techniques, is likely required for confidence in inaccessibility.

5.6 | Emerging Topics

5.6.1 | Engineering Defects and Disorder

While most of the previous generative approaches focus on ideal crystalline materials, real solids can contain defects and disorder due to, e.g., vacancies, substitutions, and dislocations. As these features can affect the final properties of the crystals, accounting for them is essential to bridge the gap between theoretical predictions and experimentally realizable materials. ML and generative models for engineering defects and incorporating disorder represent an important design opportunity for broad classes of functional materials, such as 2D materials, ferroelectrics, and small-gap semiconductors, wherein computational crystal structure generation was previously tedious and incomplete. Recent studies have begun addressing this challenge by developing models and frameworks capable of explicitly representing and learning from non-ideal structures. It is worth noting that, to date, most approaches in the literature addressing defects and disorder rely on conventional ML models with the ability to complement generative frameworks through screening, prediction, or post-processing. Only a limited number of works, to the best of

our knowledge, directly incorporate defects or disorder within the structure generation process itself.

Yang et al. introduced DefiNet, a defect-informed EGNN designed to capture the local geometries and interactions in point-defect structures [241]. A key distinction of this approach relative to conventional GNNs is the use of an explicit defect representation rather than relying on defects being implicitly inferred from atomic geometry. Each node in the graph carries additional markers that encode whether the atom corresponds to a pristine lattice site, a substitutional defect, or a vacancy, making defect identity directly available during message passing. This defect-aware formulation allows the network to learn how different defect types locally perturb atomic interactions and structural relaxations. Frey et al. applied deep transfer learning to predict defect properties in 2D materials, identifying numerous dopant–defect combinations with potential for quantum emission and neuromorphic computing [242]. Defect effects are incorporated through engineered feature vectors that explicitly encode local chemical and structural deviations from the pristine lattice, enabling a classifier to predict defect occurrence and a regressor to estimate defect formation energies. Mosquera-Lois et al. used a machine-learning surrogate model to accelerate defect structure relaxation, successfully predicting stable reconstructions for unseen materials and significantly reducing DFT [243]. This framework focuses on post-generation structural refinement around defects, rather than producing full atomic configurations from scratch. Beyond point defects, Jakob et al. introduced ML classifiers to predict the likelihood of crystallographic disorder in inorganic compounds, providing a path toward disorder-aware materials discovery workflows that better mirror experimental realities [244]. Petersen et al. developed Dis-GEN, the first generative model capable of producing symmetry-consistent disordered crystal structures with partial occupancies and vacancies, trained directly on ICSD data [245]. Dis-GEN is a VAE-based framework that explicitly represents occupational disorder and vacancies at the level of crystallographic sites. Disorder is encoded at each Wyckoff position together with a binary indicator specifying whether a site is ordered or disordered. This enables the generation of configurationally disordered structures that more faithfully reflect experimental crystallography. Together, these developments mark an important shift from modeling perfect crystals toward embracing the inherent complexity of real materials by incorporating defects and disorder. However, the widespread adoption of defect- and disorder-aware generative models remains constrained by the scarcity of data on defects and disordered materials, making this a major bottleneck for inverse design workflows.

5.6.2 | Generative Models for Advanced Characterization

As chemistry and materials research laboratories become increasingly automated and autonomous, some of the traditional bottlenecks of synthesis and characterization are gradually shifting toward data analysis and interpretation. Emerging research leverages generative models for crystal structure prediction to assist in advanced characterization and data analysis, including image and structure reconstruction during atomic-scale microscopy and ptychography [246–248], and structure solution approaches based on powder X-ray diffraction data [190, 192, 249, 250]. Embedding generative models for crystal structures as components in data analysis and characterization, as well as other workflows beyond direct materials design, will become increasingly relevant with maturing and usable generative models.

5.6.3 | Incorporating Synthetic Feasibility into Generative Models

Conditioning generative models on synthesizability represents a key upcoming challenge, as synthesis considerations are currently addressed only in post-processing workflows (see Section 5.5). Future progress might draw inspiration from advances in molecular generation, including the use of synthetic accessibility scores. Recent works propose analogous metrics for crystal structures, offering new opportunities to guide or condition generative models directly on synthesizability [251, 252]. Moreover, emerging models that predict viable precursor combinations provide an additional pathway for integrating synthesis considerations, playing a role analogous to retrosynthesis models in molecular design [253, 254]. These developments are further supported by progress in assembling synthesis-focused datasets for inorganic materials [255], which are essential for training and validating such synthesis-aware generative frameworks.

5.6.4 | Model Interpretability and Explainability

In recent years, the question of model interpretability has emerged as a central topic in machine learning research [256]. The ever-increasing complexity of modern models and their resulting incomprehensible black-box character have raised concerns about user trust, technical pitfalls such as artifacts and hallucinations, and societal concerns about model bias and fairness. Research in explainable AI aims to develop methods to mitigate these risks by increasing the transparency of these black box models. While not as well explored as for predictive models, some research exists on the interpretability of generative models [257], centered around meaningful latent space decompositions [258], identification of particularly influential training samples [259], and model parameters [260]. While these topics are being actively explored in the greater context of generative modeling, the aspect of interpretability is still notably underrepresented in crystal structure generation, aside from some isolated discussion of latent space decomposition [261, 262], demonstrating an important direction for future developments of the field.

6 | Summary and Outlook

Advancements in generative models for crystal structures benefited from rapid innovations in generative models in general, from VAEs and GANs to LLMs and diffusion/flow matching/BFNs. Currently, iterative generative models, such as BFNs, and specifically trained transformers with crystal-structure-specific tokenizations, are on par in performance, and hybrid models are not substantially outperforming them. However, there is no reason to expect that these innovation cycles will stop,

so we can expect further improvements for crystal structure generation as well, even though it is difficult to predict what the next innovations will be. Developing faster diffusion and flow matching models that require fewer integration steps is an active area of research and will be transferable to crystal generation. Similarly, many generative models currently rely on equivariant (graph) neural networks, which means that improvements in such architectures will also translate to improvements in crystal structure generation. From 2016 to approximately 2022, innovations in generative models were usually adopted for molecules first before being extended to crystal structures. Since then, development cycles have shortened, and the open challenges in molecular and materials modeling have broadened. New ways of integrating synthesis, e.g., using GFlowNets, are under development, but inherent differences between molecules and crystalline materials are leading to parallel developments to apply such models for synthesis conditioning or even parallel structure-synthesis generation.

Progress in crystal representation has led to notable performance gains, primarily by embedding explicit symmetry information that models previously had to infer. Starting with the transition from Cartesian to fractional coordinates, subsequent innovations such as asymmetric unit representations and symmetry-aware tokenization for transformers have further enhanced generative fidelity. Alternative new crystal structure representations might therefore advance the generation process without innovations on the model side.

Beyond models and representations, available data is a key ingredient for further progress. Most early generative models were trained on simpler subsets of the Materials Project, e.g., perov or MP-20. Recently, benchmarks also include more challenging datasets with larger unit cells, e.g., MPTS-52. As only 10% of the structures in the Materials Project have more than 52 atoms, high metrics on benchmarks like MPTS-52 will cover most of the application-relevant materials regarding unit cell size, but still exclude specific materials classes with larger unit cells, such as metal-organic frameworks. Scaling to larger unit cell sizes benefits from models that work on the asymmetric unit representation. It is currently unclear whether transformer-based architectures with specific tokenization related to asymmetric unit cells will have an advantage over GNN-based models, which typically require full connectivity among all atoms in the asymmetric unit cell.

From the application perspective, generative models should always generate materials with specific property requirements. Most studies on generative models for crystal structures demonstrate the possibility of conditioning generation on simple scalar properties such as formation energy or band gap, or even spectral features. However, enabling targeted multi-objective generation remains an open challenge, as no general off-the-shelf frameworks or sufficiently large datasets currently exist to support it, even though classifier-free guidance emerges as an approach that works on multiple target properties. Most papers so far report proof-of-concept studies regarding conditioning approaches, and there are only a few application studies that show that conditioning can be used to generate application-relevant materials. Recent advances in conditional generative modeling for images and molecules [263],

including training-free inference-time conditioning and non-differentiable conditions, aim to eliminate the need for specialized model retraining, improving conditioning flexibility and applicability. However, this comes at the cost of shifting computational costs from training to inference time. Future adoptions of such advances into crystal generation may have the potential to make complex (multi-)property conditioning more broadly applicable in practice. Advances in generative models conditioned on complex properties will allow their integration in characterization and data analysis workflows for autonomous labs and potentially other applications beyond materials design.

Moving from a complex post-processing workflow to a more end-to-end solution remains a challenge in the area of generative models for materials. Currently, stability assessment of generated crystal structures often requires DFT-based relaxation and energy evaluation to compare with the convex hull. Advancements in the accuracy of generative models and in energy prediction models will continually decrease the need for postprocessing steps. Despite all advances in generative models and their accuracy in benchmark tasks, as well as conditioning on application-relevant properties, there are still fundamental limitations that create a gap to experimental realization: this includes the fact that most generative models only take into account perfect crystal structures, while most application-relevant materials and their properties are at least influenced by, or even fully determined by doping, occupational and structural disorder, defects, and microstructure, posing a range of open challenges to current generative models. Furthermore, even for ideal crystal structures, integrating synthetic accessibility/synthesis prediction in generative models is an emerging topic with the potential to close the gap between progress in ML methods and transfer to experimental validation in relevant applications.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

References

1. A. Agrawal and A. Choudhary, "Perspective: Materials Informatics and Big Data: Realization of the "Fourth Paradigm" of Science in Materials Science," *APL Materials* 4, no. 5 (2016), https://doi.org/10.1063/1.4946894.

2. M. Hellenbrandt, "The Inorganic Crystal Structure Database (ICSD)–Present and Future," *Crystallography Reviews* 10, no. 1 (2004): 17–22, https://doi.org/10.1080/08893110410001664882.

3. A. Jain, S. P. Ong, G. Hautier, et al., "The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation," *APL Materials* 1, no. 1 (2013): 011002, https://doi.org/10.1063/1.4812323.

4. C. Draxl and M. Scheffler, "NOMAD: The FAIR Concept for Big Data-Driven Materials Science," *MRS Bulletin* 43, no. 9 (2018): 676–682, https://doi.org/10.1557/mrs.2018.208.

5. M. Scheffler, M. Aeschlimann, M. Albrecht, et al., "FAIR Data Enabling New Horizons for Materials Research," *Nature* 604, no. 7907 (Apr 2022): 635–642, https://doi.org/10.1038/s41586-022-04501-x.

6. L. Barroso-Luque, M. Shuaibi, X. Fu, et al., "Open Materials 2024 (OMAT24) Inorganic Materials Dataset and Models," *arXiv Preprint* arXiv:2410.12771 (2024), https://doi.org/10.48550/arXiv.2410.12771.

7. J. Schmidt, T. F. Cerqueira, A. H. Romero, et al., "Improving Machine-Learning Models in Materials Science Through Large Datasets," *Materials Today Physics* 48 (2024): 101560, https://doi.org/10.1016/j.mtphys.2024.101560.

8. J. Dagdelen, A. Dunn, S. Lee, et al., "Structured Information Extraction from Scientific Text with Large Language Models," *Nature Communications* 15, no. 1 (2024): 1418, https://doi.org/10.1038/s41467-024-45563-x.

9. B. Sanchez-Lengeling and A. Aspuru-Guzik, "Inverse Molecular Design Using Machine Learning: Generative Models for Matter Engineering," *Science* 361, no. 6400 (2018): 360–365, https://doi.org/10.1126/science.aat2663.

10. J. Noh, J. Kim, H. S. Stein, et al., "Inverse Design of Solid-State Materials via a Continuous Representation," *Matter* 1, no. 5 (Nov. 2019): 1370–1384, https://doi.org/10.1016/j.matt.2019.08.017.

11. J. Schmidt, M. R. Marques, S. Botti, and M. A. Marques, "Recent Advances and Applications of Machine Learning in Solid-State Materials Science," *npj Computational Materials* 5, no. 1 (2019): 83, https://doi.org/10.1038/s41524-019-0221-0.

12. D. T. Speckhard, M. Kuban, C. T. Koch, J. F. Rudzinski, and C. Draxl, "Workflows in Materials Science," (Oct. 2025), https://doi.org/10.26434/chemrxiv-2025-zr8xv.

13. R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, et al., "Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach," *Nature Materials* 15, no. 10 (2016): 1120–1127, https://doi.org/10.1038/nmat4717.

14. A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, "Scaling Deep Learning for Materials Discovery," *Nature* 624, no. 7990 (Nov 2023): 80–85, https://doi.org/10.1038/s41586-023-06735-9.

15. R. Potyrailo, K. Rajan, K. Stoewe, I. Takeuchi, B. Chisholm, and H. Lam, "Combinatorial and High-Throughput Screening of Materials Libraries: Review of State of the Art," *ACS Combinatorial Science* 13, no. 6 (2011): 579–633, https://doi.org/10.1021/co200007w.

16. M. L. Green, C. Choi, J. Hattrick-Simpers, et al., "Fulfilling the Promise of the Materials Genome Initiative with High-Throughput Experimental Methodologies," *Applied Physics Reviews* 4, no. 1 (2017), https://doi.org/10.1063/1.4977487.

17. A. Ludwig, "Discovery of New Materials Using Combinatorial Synthesis and High-Throughput Characterization of Thin-Film Materials Libraries Combined with Computational Methods," *npj Computational Materials* 5, no. 1 (2019): 70, https://doi.org/10.1038/s41524-019-0205-0.

18. R. B. Canty, J. A. Bennett, K. A. Brown, et al., "Science Acceleration and Accessibility with Self-Driving Labs," *Nature Communications* 16, no. 1 (Apr 2025), https://doi.org/10.1038/s41467-025-59231-1.

19. T. Dai, S. Vijayakrishnan, F. T. Szczypiński, et al., "Autonomous Mobile Robots for Exploratory Synthetic Chemistry," *Nature* 635, no. 8040 (Nov 2024): 890–897, https://doi.org/10.1038/s41586-024-08173-7.

20. N. J. Szymanski, B. Rendy, Y. Fei, et al., "An Autonomous Laboratory for the Accelerated Synthesis of Novel Materials," *Nature* 624, no. 7990 (2023): 86–91, https://doi.org/10.1038/s41586-023-06734-w.

21. M. Abolhasani and E. Kumacheva, "The Rise of Self-Driving Labs in Chemical and Materials Sciences," *Nature Synthesis* 2, no. 6 (2023): 483–492, https://doi.org/10.1038/s44160-022-00231-0.

22. G. Tom, S. P. Schmid, S. G. Baird, et al., "Self-Driving Laboratories for Chemistry and Materials Science," *Chemical Reviews* 124, no. 16 (2024): 9633–9732, https://doi.org/10.1021/acs.chemrev.4c00055.

23. P. M. Maffettone, P. Friederich, S. G. Baird, et al., "What Is Missing in Autonomous Discovery: Open Challenges for the Community," *Digital Discovery* 2, no. 6 (2023): 1644–1659, https://doi.org/10.1039/D3DD00143A.

24. C. Bilodeau, W. Jin, T. Jaakkola, R. Barzilay, and K. F. Jensen, "Generative Models for Molecular Discovery: Recent Advances and Challenges," *Wiley Interdisciplinary Reviews: Computational Molecular Science* 12, no. 5 (2022): e1608, https://doi.org/10.1002/wcms.1608.

25. D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *arXiv Preprint* arXiv:1312.6114 (2013), https://doi.org/10.48550/arXiv.1312.6114.

26. R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, et al., "Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules," *ACS Central Science* 4, no. 2 (2018): 268–276, https://doi.org/10.1021/acscentsci.7b00572.

27. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "MolGAN: An Implicit Generative Model for Small Molecular Graphs," *arXiv Preprint* arXiv:1805.11973 (2018), https://doi.org/10.48550/arXiv.1406.2661.

28. N. De Cao and T. Kipf, "MolGAN: An Implicit Generative Model for Small Molecular Graphs," *arXiv Preprint* arXiv:1805.11973 (2018), https://doi.org/10.48550/arXiv.1805.11973.

29. A. Nouira, N. Sokolovska, and J.-C. Crivello, "CrystalGAN: Learning to Discover Crystallographic Structures with Generative Adversarial Networks," arXiv:1810.11203 (2019), https://doi.org/10.48550/arXiv.1810.11203.

30. S. Kim, J. Noh, G. H. Gu, A. Aspuru-Guzik, and Y. Jung, "Generative Adversarial Networks for Crystal Structure Prediction," *ACS Central Science* 6, no. 8 (2020): 1412–1420, https://doi.org/10.1021/acscentsci.0c00426.

31. L. Dinh, D. Krueger, and Y. Bengio, "NICE: Non-Linear Independent Components Estimation," *arXiv* [cs.LG] (2015), https://doi.org/10.48550/arXiv.1410.8516.

32. K. Madhawa, K. Ishiguro, K. Nakago, and M. Abe, "GraphNVP: An Invertible Flow Model for Generating Molecular Graphs," *arXiv Preprint* arXiv:1905.11600 (2019), https://doi.org/10.48550/arXiv.1905.11600.

33. P. Wirnsberger, G. Papamakarios, B. Ibarz, et al., "Normalizing Flows for Atomic Solids," *Machine Learning: Science and Technology* 3, no. 2 (May 2022): 025009, https://doi.org/10.1088/2632-2153/ac6b16.

34. J. Köhler, M. Invernizzi, P. D. Haan, and F. Noe, "Rigid Body Flows for Sampling Molecular Crystal Structures," in *Proceedings of the 40th International Conference on Machine Learning* (PMLR, July 2023), 17301–17326, https://doi.org/10.5555/3618408.3619120.

35. J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Advances in Neural Information Processing Systems* 33 (2020): 6840–6851, https://doi.org/10.5555/3495724.3496298.

36. C. Shi, S. Luo, M. Xu, and J. Tang, "Learning Gradient Fields for Molecular Conformation Generation," in *International Conference on Machine Learning* (PMLR, 2021), 9558–9568, https://doi.org/10.48550/arXiv.2105.03902.

37. M. Xu, L. Yu, Y. Song, C. Shi, S. Ermon, and J. Tang, "GeoDiff: A Geometric Diffusion Model for Molecular Conformation Generation," *arXiv Preprint* arXiv:2203.02923 (2022), https://doi.org/10.48550/arXiv.2203.02923.

38. E. Hoogeboom, V. G. Satorras, C. Vignac, and M. Welling, "Equivariant Diffusion for Molecule Generation in 3D," in *International Conference on Machine Learning* (PMLR, 2022), 8867–8887, https://doi.org/10.48550/arXiv.2203.17003.

39. T. Xie, X. Fu, O.-E. Ganea, R. Barzilay, and T. Jaakkola, "Crystal Diffusion Variational Autoencoder for Periodic Material Generation," *arXiv Preprint* arXiv:2110.06197 (2021), https://doi.org/10.48550/arXiv.2110.06197.

40. A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention Is All You Need," *Advances in Neural Information Processing Systems* 30 (2017), https://doi.org/10.48550/arXiv.1706.03762.

41. P. Schwaller, T. Laino, T. Gaudin, et al., "Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction," *ACS Central Science* 5, no. 9 (2019): 1572–1583, https://doi.org/10.1021/acscentsci.9b00576.

42. V. Bagal, R. Aggarwal, P. Vinod, and U. D. Priyakumar, "MolGPT: Molecular Generation Using a Transformer-Decoder Model," *Journal of Chemical Information and Modeling* 62, no. 9 (2021): 2064–2076, https://doi.org/10.1021/acs.jcim.1c00600.

43. D. Flam-Shepherd and A. Aspuru-Guzik, "Language Models Can Generate Molecules, Materials, and Protein Binding Sites Directly in Three Dimensions as XYZ, CIF, and PDB Files," arXiv:2305.05708 (May 2023), https://doi.org/10.48550/arXiv.2305.05708.

44. L. M. Antunes, K. T. Butler, and R. Grau-Crespo, "Crystal Structure Generation with Autoregressive Large Language Modeling," *Nature Communications* 15, no. 1 (December 2024): 10570, https://doi.org/10.1038/s41467-024-54639-7.

45. X.-Q. Han, X.-D. Wang, M.-Y. Xu, et al., "AI-Driven Inverse Design of Materials: Past, Present and Future," *Chinese Physics Letters* (2025), https://doi.org/10.1088/0256-307X/42/2/027403.

46. A. D. Handoko and R. I. Made, "Artificial Intelligence and Generative Models for Materials Discovery – A Review," (2025), https://doi.org/10.48550/arXiv.2508.03278.

47. H. Park, Z. Li, and A. Walsh, "Has Generative Artificial Intelligence Solved Inverse Materials Design?" *Matter* 7, no. 7 (July 2024): 2355–2367, https://doi.org/10.1016/j.matt.2024.05.017.

48. Z. Wang, H. Hua, W. Lin, M. Yang, and K. C. Tan, "Crystalline Material Discovery in the Era of Artificial Intelligence," arXiv:2408.08044 (2025), https://doi.org/10.48550/arXiv.2408.08044.

49. M. Cheng, C.-L. Fu, R. Okabe, et al., "Artificial Intelligence-Driven Approaches for Materials Design and Discovery," *Nature Materials* (Jan. 2026), https://doi.org/10.1038/s41563-025-02403-7.

50. C.-N. Li, H.-P. Liang, B.-Q. Zhao, S.-H. Wei, and X. Zhang, "Machine Learning Assisted Crystal Structure Prediction Made Simple," *Journal of Materials Informatics* 4, no. 3 (2024), https://doi.org/10.20517/jmi.2024.18.

51. Z. Chen, Z. Meng, T. He, et al., "Crystal Structure Prediction Meets Artificial Intelligence," *The Journal of Physical Chemistry Letters* 16, no. 10 (2025): 2581–2591, https://doi.org/10.1021/acs.jpclett.4c03727.

52. Z. Wang, X. Luo, Q. Wang, et al., "Advances in High-Pressure Materials Discovery Enabled by Machine Learning," *Matter and Radiation at Extremes* 10, no. 3 (2025): 033801, https://doi.org/10.1063/5.0255385.

53. P.-P. D. Breuck, H.-C. Wang, G.-M. Rignanese, S. Botti, and M. A. L. Marques, "Generative AI for Crystal Structures: A Review," arXiv:2509.02723 (2025), https://doi.org/10.48550/arXiv.2509.02723.

54. Z. Li, B. Cao, R. Jiao, et al., "Materials Generation in the Era of Artificial Intelligence: A Comprehensive Survey," (2025), https://doi.org/10.48550/arXiv.2505.16379.

55. H. Park, K. O. Mastej, P. Detrattanawichai, R. Nduma, and A. Walsh, "Closing the Synthesis Gap in Computational Materials Design," *ChemRxiv*, 2025, preprint, https://doi.org/10.26434/chemrxiv-2025-sbc0c-v2.

56. "International Tables for Crystallography," https://doi.org/10.1107/97809553602060000001.

57. S. R. Hall, F. H. Allen, and I. D. Brown, "The Crystallographic Information File (CIF): A New Standard Archive File for Crystallography," *Acta Crystallographica Section A: Foundations of Crystallography* 47, no. 6 (1991): 655–685, https://doi.org/10.1107/S010876739101067X.

58. T. Xie and J. C. Grossman, "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties," *Physical Review Letters* 120, no. 14 (2018), https://doi.org/10.1103/physrevlett.120.145301.

59. R. Ruff, P. Reiser, J. Stühmer, and P. Friederich, "Connectivity Optimized Nested Line Graph Networks for Crystal Structures," *Digital Discovery* 3, no. 3 (2024): 594–601, https://doi.org/10.1039/D4DD00018H.

60. V. G. Satorras, E. Hoogeboom, and M. Welling, "E(n) Equivariant Graph Neural Networks," in *International Conference on Machine Learning* (PMLR, 2021), 9323–9332, https://doi.org/10.48550/arXiv.2102.09844.

61. F. Fuchs, D. Worrall, V. Fischer, and M. Welling, "SE(3)-Transformers: 3D Roto-Translation Equivariant Attention Networks," in *Advances in Neural Information Processing Systems* eds., H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, vol. 33 (Curran Associates, Inc., 2020), 1970–1981, https://doi.org/10.48550/arXiv.2006.10503.

62. T. Long, N. M. Fortunato, I. Opahle, et al., "Constrained Crystals Deep Convolutional Generative Adversarial Network for the Inverse Design of Crystal Structures," *npj Computational Materials* 7, no. 1 (2021): 66, https://doi.org/10.1038/s41524-021-00526-4.

63. D. Mrdjenovich and K. A. Persson, "Crystallographic Map: A General Lattice and Basis Formalism Enabling Efficient and Discretized Exploration of Crystallographic Phase Space," *Physical Review Materials* 8 (2024): 033401, https://doi.org/10.1103/PhysRevMaterials.8.033401.

64. D. Levy, S. S. Panigrahi, S.-O. Kaba, et al., "SymmCD: Symmetry-Preserving Crystal Generation with Diffusion Models," arXiv:2502.03638 (2025), https://doi.org/10.48550/arXiv.2502.03638.

65. D. Zagorac, H. Müller, S. Ruehl, J. Zagorac, and S. Rehme, "Recent Developments in the Inorganic Crystal Structure Database: Theoretical Crystal Structure Data and Related Features," *Journal of Applied Crystallography* 52, no. 5 (2019): 918–925, https://doi.org/10.1107/S160057671900997X.

66. C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward, "The Cambridge Structural Database," *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* 72, no. 2 (2016): 171–179, https://doi.org/10.1107/S2052520616003954.

67. S. Gražulis, D. Chateigner, R. T. Downs, et al., "Crystallography Open Database–An Open-Access Collection of Crystal Structures," *Journal of Applied Crystallography* 42, no. 4 (2009): 726–729, https://doi.org/10.1107/S0021889809016690.

68. S. Curtarolo, W. Setyawan, G. L. Hart, et al., "AFLOW: An Automatic Framework for High-Throughput Materials Discovery," *Computational Materials Science* 58 (2012): 218–226, https://doi.org/10.1016/j.commatsci.2012.02.005.

69. M. M. Ghahremanpour, P. J. Van Maaren, and D. Van Der Spoel, "The Alexandria Library, a Quantum-Chemical Database of Molecular Properties for Force Field Development," *Scientific Data* 5, no. 1 (2018): 1–10, https://doi.org/10.1038/sdata.2018.62.

70. K. Choudhary, K. F. Garrity, A. C. Reid, et al., "The Joint Automated Repository for Various Integrated Simulations (JARVIS) for Data-Driven Materials Design," *npj Computational Materials* 6, no. 1 (2020): 173, https://doi.org/10.1038/s41524-020-00440-1.

71. S. Kirklin, J. E. Saal, B. Meredig, et al., "The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies," *npj Computational Materials* 1, no. 1 (2015): 1–15, https://doi.org/10.1038/npjcompumats.2015.10.

72. S. Haastrup, M. Strange, M. Pandey, et al., "The Computational 2D Materials Database: High-Throughput Modeling and Discovery of Atomically Thin Crystals," *2D Materials* 5, no. 4 (2018): 042002.

73. J. Zhou, L. Shen, M. D. Costa, et al., "2DMatPedia, an Open Computational Database of Two-Dimensional Materials from Top-Down and Bottom-Up Approaches," *Scientific Data* 6, no. 1 (2019): 86, https://doi.org/10.1038/s41597-019-0097-3.

74. C. J. Pickard, "AIRSS Data for Carbon at 10 GPa and the C+N+H+O System at 1 GPa," (Materials Cloud, 2020), https://doi.org/10.24435/MATERIALSCLOUD:2020.0026/V1.

75. I. E. Castelli, D. D. Landis, K. S. Thygesen, et al., "New Cubic Perovskites for One-and Two-Photon Water Splitting Using the Computational Materials Repository," *Energy & Environmental Science* 5, no. 10 (2012): 9034–9043, https://doi.org/10.1039/C2EE22341D.

76. I. E. Castelli, T. Olsen, S. Datta, et al., "Computational Screening of Perovskite Metal Oxides for Optimal Solar Light Capture," *Energy & Environmental Science* 5, no. 2 (2012): 5814–5819, https://doi.org/10.1039/C1EE02717D.

77. L. Chanussot, A. Das, S. Goyal, et al., "Open Catalyst 2020 (OC20) Dataset and Community Challenges," *ACS Catalysis* 11, no. 10 (2021): 6059–6072, https://doi.org/10.1021/acscatal.0c04525.

78. Y. Zhang, S. Itani, K. Khanal, et al., "GPTArticleExtractor: An Automated Workflow for Magnetic Material Database Construction," *Journal of Magnetism and Magnetic Materials* 597 (2024): 172001, https://doi.org/10.1016/j.jmmm.2024.172001.

79. S. Itani, Y. Zhang, and J. Zang, "The Northeast Materials Database for Magnetic Materials," *Nature Communications* 16, no. 1 (2025): 9415, https://doi.org/10.1038/s41467-025-64458-z.

80. M. D. Group, "MDR Supercon Datasheet Ver.220808," https://doi.org/10.48505/nims.3837.

81. T. Sommer, R. Willa, J. Schmalian, and P. Friederich, "3DSC–A Dataset of Superconductors Including Crystal Structures," *Scientific Data* 10, no. 1 (2023): 816, https://doi.org/10.1038/s41597-023-02721-y.

82. L. Sbailò, D. Fekete, L. Ghiringhelli, and M. Scheffler, "The NOMAD Artificial-Intelligence Toolkit: Turning Materials-Science Data into Knowledge and Understanding," *npj Computational Materials* 8 (December 2022), https://doi.org/10.1038/s41524-022-00935-z.

83. M. Scheidgen, L. Himanen, A. N. Ladines, et al., "NOMAD: A Distributed Web-Based Platform for Managing Materials Science Research Data," *Journal of Open Source Software* 8, no. 90 (2023): 5388, https://doi.org/10.21105/joss.05388.

84. A. Le Bail, "Inorganic Structure Prediction with *GRINSP*," *Journal of Applied Crystallography* 38, no. 2 (2005): 389–395, https://doi.org/10.1107/S0021889805002384.

85. C. Chen, W. Ye, Y. Zuo, C. Zheng, and S. P. Ong, "Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals," *Chemistry of Materials* 31, no. 9 (2019): 3564–3572, https://doi.org/10.1021/acs.chemmater.9b01294.

86. K. Choudhary and B. DeCost, "Atomistic Line Graph Neural Network for Improved Materials Property Predictions," *npj Computational Materials* 7, no. 1 (Nov 2021), https://doi.org/10.1038/s41524-021-00650-1.

87. A. Trewartha, N. Walker, H. Huo, et al., "Quantifying the Advantage of Domain-Specific Pre-Training on Named Entity Recognition Tasks in Materials Science," *Patterns* 3, no. 4 (2022), https://doi.org/10.1016/j.patter.2022.100488.

88. A. Niyongabo Rubungo, C. Arnold, B. P. Rand, and A. B. Dieng, "LLM-Prop: Predicting the Properties of Crystalline Materials Using Large Language Models," *npj Computational Materials* 11, no. 1 (2025), https://doi.org/10.1038/s41524-025-01536-2.

89. V. L. Deringer, M. A. Caro, and G. Csányi, "Machine Learning Interatomic Potentials as Emerging Tools for Materials Science," *Advanced Materials* 31, no. 46 (2019): 1902765, https://doi.org/10.1002/adma.201902765.

90. I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner, and G. Csányi, "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields," (2023), https://doi.org/10.48550/arXiv.2206.07697.

91. I. Batatia, P. Benner, Y. Chiang, et al., "A Foundation Model for Atomistic Materials Chemistry," arXiv:2401.00096 (2024), https://doi.org/10.48550/arXiv.2401.00096.

92. C. Chen and S. P. Ong, "A Universal Graph Deep Learning Interatomic Potential for the Periodic Table," *Nature Computational Science* 2, no. 11 (Nov 2022): 718–728, https://doi.org/10.48550/arXiv.2202.02450.

93. B. Deng, P. Zhong, K. Jun, et al., "CHGNet as a Pretrained Universal Neural Network Potential for Charge-Informed Atomistic Modelling," *Nature Machine Intelligence* 5, no. 9 (2023): 1031–1041, https://doi.org/10.1038/s42256-023-00716-3.

94. D. B. Laks, L. Ferreira, S. Froyen, and A. Zunger, "Efficient Cluster Expansion for Substitutional Systems," *Physical Review B* 46, no. 19 (1992): 12587, https://doi.org/10.1103/PhysRevB.46.12587.

95. C. Wolverton and D. de Fontaine, "Cluster Expansions of Alloy Energetics in Ternary Intermetallics," *Physical Review B* 49, no. 13 (1994): 8627, https://doi.org/10.1103/PhysRevB.49.8627.

96. G. Ceder, "Predicting Properties from Scratch," *Science* 280, no. 5366 (1998): 1099–1100, https://doi.org/10.1126/science.280.5366.1099.

97. G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, "Finding Nature's Missing Ternary Oxide Compounds Using Machine Learning and Density Functional Theory," *Chemistry of Materials* 22, no. 12 (2010): 3762–3767, https://doi.org/10.1021/cm100795d.

98. G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, and G. Ceder, "Data Mined Ionic Substitutions for the Discovery of New Compounds," *Inorganic Chemistry* 50, no. 2 (2011): 656–663, https://doi.org/10.1021/ic102031h.

99. H. Chen, G. Hautier, A. Jain, et al., "Carbonophosphates: A New Family of Cathode Materials for Li-Ion Batteries Identified Computationally," *Chemistry of Materials* 24, no. 11 (2012): 2009–2016, https://doi.org/10.1021/cm203243x.

100. H. Chen, Q. Hao, O. Zivkovic, et al., "Sidorenkite ($Na_3MnPO_4CO_3$): A New Intercalation Cathode Material for Na-Ion Batteries," *Chemistry of Materials* 25, no. 14 (2013): 2777–2786, https://doi.org/10.1021/cm400805q.

101. S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The High-Throughput Highway to Computational Materials Design," *Nature Materials* 12, no. 3 (2013): 191–201, https://doi.org/10.1038/nmat3568.

102. J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, and J. K. Nørskov, "Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution," *Nature Materials* 5, no. 11 (2006): 909–913, https://doi.org/10.1038/nmat1752.

103. A. Ahmed, Y. Liu, J. Purewal, et al., "Balancing Gravimetric and Volumetric Hydrogen Density in MOFs," *Energy & Environmental Science* 10, no. 11 (2017): 2459–2471, https://doi.org/10.1039/C7EE02477K.

104. R.-Z. Zhang, K. Chen, B. Du, and M. J. Reece, "Screening for Cu–S Based Thermoelectric Materials Using Crystal Structure Features," *Journal of Materials Chemistry A* 5, no. 10 (2017): 5013–5019, https://doi.org/10.1039/C6TA10607B.

105. C. W. Glass, A. R. Oganov, and N. Hansen, "USPEX–Evolutionary Crystal Structure Prediction," *Computer Physics Communications* 175, no. 11 (2006): 713–720, https://doi.org/10.1016/j.cpc.2006.07.020.

106. Y. Wang, J. Lv, L. Zhu, and Y. Ma, "Crystal Structure Prediction via Particle-Swarm Optimization," *Physical Review B* 82 (2010): 094116, https://doi.org/10.1103/PhysRevB.82.094116.

107. Y. Wang, J. Lv, L. Zhu, and Y. Ma, "CALYPSO: A Method for Crystal Structure Prediction," *Computer Physics Communications* 183, no. 10 (2012): 2063–2070, https://doi.org/10.1016/j.cpc.2012.05.008.

108. C. J. Pickard and R. J. Needs, "Ab Initio Random Structure Searching," *Journal of Physics: Condensed Matter* 23, no. 5 (2011): 053201, https://doi.org/10.1088/0953-8984/23/5/053201.

109. Y. J. Colón, D. A. Gómez-Gualdrón, and R. Q. Snurr, "Topologically Guided, Automated Construction of Metal–Organic Frameworks and Their Evaluation for Energy-Related Applications," *Crystal Growth & Design* 17, no. 11 (2017): 5801–5810, https://doi.org/10.1021/acs.cgd.7b00848.

110. M. A. Addicoat, D. E. Coupry, and T. Heine, "AuToGraFS: Automatic Topological Generator for Framework Structures," *The Journal of Physical Chemistry A* 118, no. 40 (2014): 9607–9614, https://doi.org/10.1021/jp507643v.

111. P. D. Angelis, G. Trezza, G. Barletta, P. Asinari, and E. Chiavazzo, "Energy-GNOME: A Living Database of Selected Materials for Energy Applications," *arXiv* [cond-mat.mtrl-sci] (2024).

112. A. Razavi, A. van den Oord, and O. Vinyals, "Generating Diverse High-Fidelity Images with VQ-VAE-2," in *Advances in Neural Information Processing Systems*, vol. 32 (Curran Associates, Inc., 2019), https://doi.org/10.48550/arXiv.1906.00446.

113. Y. Pu, Z. Gan, R. Henao, et al., "Variational Autoencoder for Deep Learning of Images, Labels and Captions," *Advances in Neural Information Processing Systems* 29 (2016), https://doi.org/10.48550/arXiv.1609.08976.

114. "An Invertible Crystallographic Representation for General Inverse Design of Inorganic Crystals with Targeted Properties," *Matter* 5, no. 1 (2022): 314–335, https://doi.org/10.1016/j.matt.2021.11.032.

115. K. Liu, S. Gao, K. Yang, and Y. Han, "PCVAE: A Physics-Informed Neural Network for Determining the Symmetry and Geometry of Crystals," in *2023 International Joint Conference on Neural Networks (IJCNN)* (2023), pp. 1–8, https://doi.org/10.1109/IJCNN54540.2023.10191051.

116. S. Mal, G. Seal, and P. Sen, "MagGen: A Graph-Aided Deep Generative Model for Inverse Design of Permanent Magnets," *The Journal of Physical Chemistry Letters* 15, no. 12 (2024): 3221–3228, https://doi.org/10.1021/acs.jpclett.4c00068.

117. R. Zhu, W. Nong, S. Yamazaki, and K. Hippalgaonkar, "WyCryst: Wyckoff Inorganic Crystal Generator Framework," *Matter* 7, no. 10 (2024): 3469–3488, https://doi.org/10.1016/j.matt.2024.05.042.

118. E. T. Chenebuah, M. Nganbe, and A. B. Tchagang, "An Evolutionary Variational Autoencoder for Perovskite Discovery," *Frontiers in Materials* 10 (2023), https://doi.org/10.3389/fmats.2023.1233961.

119. E. Chenebuah, M. Nganbe, and A. Tchagang, "Target-Learning the Latent Space of a Variational Autoencoder Model for the Inverse Design of Stable Perovskites," *Proceedings of the Canadian Conference on Artificial Intelligence* (2023), https://doi.org/10.21428/594757db.07402193.

120. R. Winter, F. Noe, and D.-A. Clevert, "Permutation-Invariant Variational Autoencoder for Graph-Level Representation Learning," in *Advances in Neural Information Processing Systems*, vol. 34 (Curran Associates, Inc., 2021), pp. 9559–9573, https://doi.org/10.48550/arXiv.2104.09856.

121. J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," *IEEE Transactions on Knowledge and Data Engineering* 35, no. 4 (2021): 3313–3332, https://doi.org/10.1109/TKDE.2021.3130191.

122. B. Kim, S. Lee, and J. Kim, "Inverse Design of Porous Materials Using Artificial Neural Networks," *Science Advances* 6, no. 1 (2020): eaax9324, https://doi.org/10.1126/sciadv.aax9324.

123. Y. Dan, Y. Zhao, X. Li, S. Li, M. Hu, and J. Hu, "Generative Adversarial Networks (GAN) Based Efficient Sampling of Chemical Composition Space for Inverse Design of Inorganic Materials," *npj Computational Materials* 6, no. 1 (2020): 84, https://doi.org/10.1038/s41524-020-00352-0.

124. Y. Zhao, M. Al-Fahdi, M. Hu, et al., "High-Throughput Discovery of Novel Cubic Crystal Materials Using Deep Generative Neural Networks," *Advanced Science* 8, no. 20 (2021): 2100566, https://doi.org/10.1002/advs.202100566.

125. Y. Zhao, E. M. D. Siriwardane, Z. Wu, et al., "Physics Guided Deep Learning for Generative Design of Crystal Materials with Symmetry Constraints," *npj Computational Materials* 9, no. 1 (2023): 38, https://doi.org/10.1038/s41524-023-00987-9.

126. T. Su, B. Cao, S. Hu, M. Li, and T.-Y. Zhang, "CGWGAN: Crystal Generative Framework Based on Wyckoff Generative Adversarial Network," *Journal of Materials Informatics* 4, no. 4 (2024): 20, https://doi.org/10.20517/jmi.2024.24.

127. E. Zamaraeva, C. M. Collins, D. Antypov, et al., "Reinforcement Learning in Crystal Structure Prediction," *Digital Discovery* 2, no. 6 (2023): 1831–1840, https://doi.org/10.1039/D3DD00063J.

128. S. A. Meldgaard, H. L. Mortensen, M. S. Joergensen, and B. Hammer, "Structure Prediction of Surface Reconstructions by Deep Reinforcement Learning," *Journal of Physics: Condensed Matter* 32, no. 40 (2020), https://doi.org/10.1088/1361-648X/ab94f2.

129. P. Govindarajan, S. Miret, J. Rector-Brooks, M. Phielipp, J. Rajendran, and S. Chandar, "Learning Conditional Policies for Crystal Design Using Offline Reinforcement Learning," *Digital Discovery* 3, no. 4 (2024): 769–785, https://doi.org/10.1039/D4DD00024B.

130. C. Karpovich, E. Pan, and E. A. Olivetti, "Deep Reinforcement Learning for Inverse Inorganic Materials Design," *npj Computational Materials* 10, no. 1 (2024): 287, https://doi.org/10.1038/s41524-024-01474-5.

131. J. Chen, J. Guo, and P. Schwaller, "MatInvent: Reinforcement Learning for 3D Crystal Diffusion Generation," in *AI for Accelerated Materials Design – ICLR 2025* (2025), April, 9.

132. T. Mohanty, M. Mehta, H. M. Sayeed, V. Srikumar, and T. D. Sparks, "CrysText: A Generative AI Approach for Text-Conditioned Crystal Structure Generation Using LLM," ChemRxiv (2025), https://doi.org/10.26434/chemrxiv-2024-gjhpq-v2.

133. H. Park and A. Walsh, "Guiding Generative Models to Uncover Diverse and Novel Crystals via Reinforcement Learning," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2511.07158.

134. J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep Unsupervised Learning Using Nonequilibrium Thermodynamics," in *International Conference on Machine Learning* (PMLR, 2015), pp. 2256–2265, https://doi.org/10.48550/arXiv.1503.03585.

135. L. Yang, Z. Zhang, Y. Song, et al., "Diffusion Models: A Comprehensive Survey of Methods and Applications," *ACM Computing Surveys* 56, no. 4 (2023): 1–39, https://doi.org/10.1145/3626235.

136. A. Alakhdar, B. Poczos, and N. Washburn, "Diffusion Models in De Novo Drug Design," *Journal of Chemical Information and Modeling* 64, no. 19 (2024): 7238–7256, https://doi.org/10.1021/acs.jcim.4c01107.

137. L. Wang, C. Song, Z. Liu, Y. Rong, Q. Liu, and S. Wu, "Diffusion Models for Molecules: A Survey of Methods and Tasks," *arXiv Preprint* (2025), https://doi.org/10.48550/arXiv.2502.09511.

138. Y. Song and S. Ermon, "Generative Modeling by Estimating Gradients of the Data Distribution," *Advances in Neural Information Processing Systems* 32 (2019), https://doi.org/10.48550/arXiv.1907.05600.

139. Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-Based Generative Modeling Through Stochastic Differential Equations," *arXiv Preprint* (2020), https://doi.org/10.48550/arXiv.2011.13456.

140. "Con-CDVAE: A Method for the Conditional Generation of Crystal Structures," *Computational Materials Today* 1 (2024): 100003, https://doi.org/10.1016/j.commt.2024.100003.

141. X. Luo, Z. Wang, P. Gao, et al., "Deep Learning Generative Model for Crystal Structure Prediction," *npj Computational Materials* 10, no. 1 (2024): 254, https://doi.org/10.1038/s41524-024-01443-y.

142. R. Jiao, W. Huang, P. Lin, et al., "Crystal Structure Prediction by Joint Equivariant Diffusion," *arXiv* [cond-mat] (2024), https://doi.org/10.48550/arXiv.2309.04475.

143. V. De Bortoli, E. Mathieu, M. Hutchinson, J. Thornton, Y. W. Teh, and A. Doucet, "Riemannian Score-Based Generative Modelling," *Advances in Neural Information Processing Systems* 35 (2022): 2406–2422, https://doi.org/10.48550/arXiv.2202.02763.

144. R. Jiao, W. Huang, Y. Liu, D. Zhao, and Y. Liu, "Space Group Constrained Crystal Generation," *arXiv* [cs] (2024), https://doi.org/10.48550/arXiv.2402.03992.

145. F. E. Kelvinius, O. B. Andersson, A. S. Parackal, D. Qian, R. Armiento, and F. Lindsten, "WyckoffDiff – A Generative Diffusion Model for Crystal Symmetry," *arXiv* [cond-mat] (2025), https://doi.org/10.48550/arXiv.2502.06485.

146. X. Fu, T. Xie, A. S. Rosen, T. Jaakkola, and J. Smith, "MOFDiff: Coarse-Grained Diffusion for Metal-Organic Framework Design," *arXiv Preprint* (2023), https://doi.org/10.48550/arXiv.2310.10732.

147. S. Yuan and S. Dordevic, "Diffusion Models for Conditional Generation of Hypothetical New Families of Superconductors," *Scientific Reports* 14, no. 1 (2024): 10275, https://doi.org/10.48550/arXiv.2402.00198.

148. J. Park, A. P. S. Gill, S. M. Moosavi, and J. Kim, "Inverse Design of Porous Materials: A Diffusion Model Approach," *Journal of Materials Chemistry A* 12, no. 11 (2024): 6507–6514, https://doi.org/10.1039/D3TA06274K.

149. C. Zeni, R. Pinsler, D. Zügner, et al., "A Generative Model for Inorganic Materials Design," *Nature* 639, no. 8055 (2025): 624–632, https://doi.org/10.1038/s41586-025-08628-5.

150. H. Park, A. Onwuli, et al., "Exploration of Crystal Chemical Space Using Text-Guided Generative Artificial Intelligence," *Nature Communications* 16, no. 1 (2025): 4379, https://doi.org/10.1038/s41467-025-59636-y.

151. I. Kobyzev, S. J. Prince, and M. A. Brubaker, "Normalizing Flows: An Introduction and Review of Current Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, no. 11 (2021): 3964–3979, https://doi.org/10.1109/TPAMI.2020.2992934.

152. G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, "Normalizing Flows for Probabilistic Modeling and Inference," *Journal of Machine Learning Research* 22, no. 1 (2021), https://doi.org/10.48550/arXiv.1912.02762.

153. L. Dinh, J. Sohl-Dickstein, and S. Bengio, "Density Estimation Using Real NVP," *arXiv* [cs.LG] (2017), https://doi.org/10.48550/arXiv.1605.08803.

154. C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, "Neural Spline Flows," in *Advances in Neural Information Processing Systems*, vol. 32 (Curran Associates, Inc., 2019), https://doi.org/10.48550/arXiv.1906.04032.

155. F. Noé, S. Olsson, J. Köhler, and H. Wu, "Boltzmann Generators: Sampling Equilibrium States of Many-Body Systems with Deep Learning," *Science* 365, no. 6457 (2019): eaaw1147, https://doi.org/10.1126/science.aaw1147.

156. P. Wirnsberger, A. J. Ballard, G. Papamakarios, et al., "Targeted Free Energy Estimation via Learned Mappings," *The Journal of Chemical Physics* 153, no. 14 (2020): 144112, https://doi.org/10.1063/5.0018903.

157. R. Ahmad and W. Cai, "Free Energy Calculation of Crystalline Solids Using Normalizing Flow," *Modelling and Simulation in Materials Science and Engineering* 30, no. 6 (2022): 065007, https://doi.org/10.1088/1361-651X/ac7f4b.

158. M. Schebek, M. Invernizzi, F. Noé, and J. Rogal, "Efficient Mapping of Phase Diagrams with Conditional Boltzmann Generators," *Machine Learning: Science and Technology* 5, no. 4 (2024): 045045, https://doi.org/10.1088/2632-2153/ad849d.

159. "Neural Ordinary Differential Equations," in *Advances in Neural Information Processing Systems*, vol. 31 (Curran Associates, Inc., 2018), https://doi.org/10.48550/arXiv.1806.07366.

160. W. Grathwohl, R. T. Q. Chen, J. Bettencourt, I. Sutskever, and D. Duvenaud, "FFJORD: Free-Form Continuous Dynamics for Scalable Reversible Generative Models," *International Conference on Learning Representations* (2019), https://doi.org/10.48550/arXiv.1810.01367.

161. Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le, "Flow Matching for Generative Modeling," *The Eleventh International Conference on Learning Representations* (2023), https://doi.org/10.48550/arXiv.2210.02747.

162. X. Liu, C. Gong, and Q. Liu, "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow," *The Eleventh International Conference on Learning Representations* (2023), https://doi.org/10.48550/arXiv.2209.03003.

163. R. T. Q. Chen and Y. Lipman, "Flow Matching on General Geometries," *The Twelfth International Conference on Learning Representations* (2024), https://doi.org/10.48550/arXiv.2302.03660.

164. B. K. Miller, R. T. Q. Chen, A. Sriram, and B. M. Wood, "FlowMM: Generating Materials with Riemannian Flow Matching," *Forty-First International Conference on Machine Learning* (2024), https://doi.org/10.48550/arXiv.2406.04713.

165. N. Kim, S. Kim, M. Kim, J. Park, and S. Ahn, "MOFFlow: Flow Matching for Structure Prediction of Metal-Organic Frameworks," *arXiv* [q-bio] (2025), https://doi.org/10.48550/arXiv.2410.17270.

166. E. Bengio, M. Jain, M. Korablyov, D. Precup, and Y. Bengio, "Flow Network Based Generative Models for Non-Iterative Diverse Candidate Generation," in *Proceedings of the 35th International Conference on Neural Information Processing Systems* (2021), 27381–27394, https://doi.org/10.48550/arXiv.2106.04399.

167. Y. Bengio, S. Lahlou, T. Deleu, E. J. Hu, M. Tiwari, and E. Bengio, "GFlowNet Foundations," *Journal of Machine Learning Research* 24, no. 210 (2023): 1–55, https://doi.org/10.48550/arXiv.2111.09266.

168. A. Hernandez-Garcia, et al., "Crystal-GFN: Sampling Crystals with Desirable Properties and Constraints," *arXiv* [cs] (2023), https://doi.org/10.48550/arXiv.2310.04925.

169. T. M. Nguyen, et al., "Efficient Symmetry-Aware Materials Generation via Hierarchical Generative Flow Networks," *arXiv* [cs] (2024), https://doi.org/10.48550/arXiv.2411.04323.

170. F. Cipcigan, et al., "Discovery of Novel Reticular Materials for Carbon Dioxide Capture Using GFlowNets," *Digital Discovery* 3, no. 3 (2024): 449–455, https://doi.org/10.1039/D4DD00020J.

171. A. Graves, R. K. Srivastava, T. Atkinson, and F. Gomez, "Bayesian Flow Networks," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2308.07037.

172. H. Wu, Y. Song, J. Gong, et al., "A Periodic Bayesian Flow for Material Generation," *arXiv* [cs] (2025), https://doi.org/10.48550/arXiv.2502.02016.

173. L. Ruple, L. Torresi, H. Schopmans, and P. Friederich, "Symmetry-Aware Bayesian Flow Networks for Crystal Generation," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2502.03146.

174. A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," https://api.semanticscholar.org/CorpusID:49313245.

175. A. Karpathy, "nanoGPT: The Simplest, Fastest Repository for Training/Fine-Tuning Medium-Sized GPTs," https://github.com/karpathy/nanoGPT.

176. K. Yan, X. Li, H. Ling, et al., "Invariant Tokenization of Crystalline Materials for Language Model Enabled Generation," *arXiv* [cs] (2025), https://doi.org/10.48550/arXiv.2503.00152.

177. Z. Cao, X. Luo, J. Lv, and L. Wang, "Space Group Informed Transformer for Crystalline Materials Generation," *arXiv* [cond-mat] (2024), https://doi.org/10.48550/arXiv.2403.15734.

178. N. Kazeev, W. Nong, I. Romanov, et al., "Wyckoff Transformer: Generation of Symmetric Crystals," *arXiv* [cond-mat] (2025), https://doi.org/10.48550/arXiv.2503.02407.

179. N. Gruver, A. Sriram, A. Madotto, A. G. Wilson, C. L. Zitnick, and Z. Ulissi, "Fine-Tuned Language Models Generate Stable Inorganic Materials as Text," *arXiv* [cs] (2024), https://doi.org/10.48550/arXiv.2402.04379.

180. H. Touvron, L. Martin, K. Stone, et al., "Llama 2: Open Foundation and Fine-Tuned Chat Models," *arXiv* [cs.CL] (2023), https://doi.org/10.48550/arXiv.2307.09288.

181. A. Sriram, B. K. Miller, R. T. Q. Chen, and B. M. Wood, "FlowLLM: Flow Matching for Material Generation with Large Language Models as Base Distributions," *arXiv* [cs] (2024), https://doi.org/10.48550/arXiv.2410.23405.

182. S. Khastagir, K. Das, P. Goyal, S.-C. Lee, S. Bhattacharjee, and N. Ganguly, "LLM Meets Diffusion: A Hybrid Framework for Crystal Material Generation," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2510.23040.

183. A. Q. Jiang, A. Sablayrolles, A. Mensch, et al., "Mistral 7B," *arXiv* [cs] (2023), https://doi.org/10.48550/arXiv.2310.06825.

184. K. Choudhary, "AtomGPT: Atomistic Generative Pretrained Transformer for Forward and Inverse Materials Design," *The Journal of Physical Chemistry Letters* 15, no. 27 (2024): 6909–6917, https://doi.org/10.1021/acs.jpclett.4c01126.

185. P. Zhong, X. Dai, B. Deng, G. Ceder, and K. A. Persson, "Practical Approaches for Crystal Structure Predictions with Inpainting Generation and Universal Interatomic Potentials," *arXiv* [cond-mat] (2025), https://doi.org/10.48550/arXiv.2504.16893.

186. J. Ho and T. Salimans, "Classifier-Free Diffusion Guidance," *arXiv* [cs] (2022), https://doi.org/10.48550/arXiv.2207.12598.

187. Q. Zheng, M. Le, N. Shaul, Y. Lipman, A. Grover, and R. T. Q. Chen, "Guided Flows for Generative Modeling and Decision Making," *arXiv* [cs] (2023), https://doi.org/10.48550/arXiv.2311.13443.

188. H. Park, A. Onwuli, and A. Walsh, "Exploration of Crystal Chemical Space Using Text-Guided Generative Artificial Intelligence," *Nature Communications* 16, no. 1 (2025): 4379, https://doi.org/10.1038/s41467-025-59636-y.

189. K. Tsukaue and K. Yasuda, "Crystal Structure Generation Using a Diffusion Model Conditioned on X-Ray Diffraction Intensities Without Label Learning," in *4th Annual AAAI Workshop on AI to Accelerate Science and Engineering* (2025).

190. E. A. Riesel, T. Mackey, H. Nilforoshan, et al., "Crystal Structure Determination from Powder Diffraction Patterns with Generative Machine Learning," *Journal of the American Chemical Society* 146, no. 44 (2024): 30340–30348, https://doi.org/10.1021/jacs.4c10244.

191. Q. Lai, F. Xu, L. Yao, et al., "End-to-End Crystal Structure Prediction from Powder X-Ray Diffraction," *Advanced Science* 12, no. 8 (2025): 2410722, https://doi.org/10.1002/advs.202410722.

192. G. Guo, T. L. Saidi, M. W. Terban, M. Valsecchi, S. J. L. Billinge, and H. Lipson, "Ab Initio Structure Solutions from Nanocrystalline Powder Diffraction Data via Diffusion Models," *Nature Materials* (2025), https://doi.org/10.1038/s41563-025-02220-y.

193. H. Kwon, T. Hsu, W. Sun, et al., "Spectroscopy-Guided Discovery of Three-Dimensional Structures of Disordered Materials with Diffusion Models," *arXiv* [cond-mat] (2023), https://doi.org/10.48550/arXiv.2312.05472.

194. G. Chandan Kanakala, B. Sridharan, and U. D. Priyakumar, "Spectra to Structure: Contrastive Learning Framework for Library Ranking and Generating Molecular Structures for Infrared Spectra," *Digital Discovery* 3, no. 12 (2024): 2417–2423, https://doi.org/10.1039/D4DD00135D.

195. S. P. Ong, W. D. Richards, A. Jain, et al., "Python Materials Genomics (pymatgen): A Robust, Open-Source Python Library for Materials Analysis," *Computational Materials Science* 68 (2013): 314–319, https://doi.org/10.1016/j.commatsci.2012.10.028.

196. M. M. Martirossyan, T. Egg, P. Hoellmer, et al., "All That Structure Matches Does Not Glitter," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2509.12178.

197. C. J. Bartel, "Review of Computational Approaches to Predict the Thermodynamic Stability of Inorganic Solids," *Journal of Materials Science* 57, no. 23 (2022): 10475–10498, https://doi.org/10.1007/s10853-022-06915-4.

198. H. Yang, C. Hu, Y. Zhou, et al., "MatterSim: A Deep Learning Atomistic Model Across Elements, Temperatures and Pressures," *arXiv* [cond-mat.mtrl-sci] (2024), https://doi.org/10.48550/arXiv.2405.04967.

199. R. Jacobs, D. Morgan, S. Attarian, et al., "A Practical Guide to Machine Learning Interatomic Potentials – Status and Future," *Current Opinion in Solid State and Materials Science* 35 (2025): 101214, https://doi.org/10.1016/j.cossms.2025.101214.

200. J. Riebesell, R. E. A. Goodall, P. Benner, et al., "A Framework to Evaluate Machine Learning Crystal Stability Predictions," *Nature Machine Intelligence* 7, no. 6 (2025): 836–847, https://doi.org/10.1038/s42256-025-01055-1.

201. M. Negishi, H. Park, K. O. Mastej, and A. Walsh, "Continuous Uniqueness and Novelty Metrics for Generative Modeling of Inorganic Crystals," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2510.12405.

202. J. H. Montoya, C. Grimley, M. Aykol, et al., "How the AI-Assisted Discovery and Synthesis of a Ternary Oxide Highlights Capability Gaps in Materials Science," *Chemical Science* 15, no. 15 (2024): 5660–5673, https://doi.org/10.1039/d3sc04823c.

203. A. Stein, S. W. Keller, and T. E. Mallouk, "Turning Down the Heat: Design and Mechanism in Solid-State Synthesis," *Science* 259, no. 5101 (1993): 1558–1564, https://doi.org/10.1126/science.259.5101.1558.

204. M. Jansen, "Synergies and Competition Between Current Approaches to Materials Discovery," *Comptes Rendus Chimie* 21, no. 10 (2018): 958–968, https://doi.org/10.1016/j.crci.2018.03.007.

205. W. Sun and N. David, "A Critical Reflection on Attempts to Machine-Learn Materials Synthesis Insights from Text-Mined Literature Recipes," *Faraday Discussions* 256 (2025): 614–638, https://doi.org/10.1039/d4fd00112e.

206. S. P. Ong, L. Wang, B. Kang, and G. Ceder, "Li-Fe-P-O_2 Phase Diagram from First Principles Calculations," *Chemistry of Materials* 20, no. 5 (2008): 1798–1807, https://doi.org/10.1021/cm702327g.

207. M. K. Horton, P. Huck, R. X. Yang, et al., "Accelerated Data-Driven Materials Science with the Materials Project," *Nature Materials* (2025), https://doi.org/10.1038/s41563-025-02272-0.

208. J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, "Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD)," *JOM* 65, no. 11 (2013): 1501–1509, https://doi.org/10.1007/s11837-013-0755-4.

209. W. Sun, S. T. Dacek, S. P. Ong, et al., "The Thermodynamic Scale of Inorganic Crystalline Metastability," *Science Advances* 2, no. 11 (2016), https://doi.org/10.1126/sciadv.1600225.

210. M. Aykol, S. S. Dwaraknath, W. Sun, and K. A. Persson, "Thermodynamic Limit for Synthesis of Metastable Inorganic Materials," *Science Advances* 4, no. 4 (2018), https://doi.org/10.1126/sciadv.aaq0148.

211. P. C. Müller, C. Ertural, J. Hempelmann, and R. Dronskowski, "Crystal Orbital Bond Index: Covalent Bond Orders in Solids," *The Journal of Physical Chemistry C* 125, no. 14 (2021): 7959–7970, https://doi.org/10.1021/acs.jpcc.1c00718.

212. J. George, D. Waroquiers, D. Di Stefano, G. Petretto, G. Rignanese, and G. Hautier, "The Limited Predictive Power of the Pauling Rules," *Angewandte Chemie International Edition* 59, no. 19 (2020): 7569–7575, https://doi.org/10.1002/anie.202000829.

213. J. Simonson and S. Poon, "Applying an Electron Counting Rule to Screen Prospective Thermoelectric Alloys: The Thermoelectric Properties of $YCrB_4$ and Er_3CrB_7-Type Phases," *Journal of Alloys and Compounds* 504, no. 1 (2010): 265–272, https://doi.org/10.1016/j.jallcom.2010.05.110.

214. M. R. Filip and F. Giustino, "The Geometric Blueprint of Perovskites," *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 21 (2018): 5397–5402, https://doi.org/10.1073/pnas.1719179115.

215. N. Fu, J. Hu, Y. Feng, G. Morrison, H. Z. Loye, and J. Hu, "Composition Based Oxidation State Prediction of Materials Using Deep Learning Language Models," *Advanced Science* 10, no. 28 (2023), https://doi.org/10.1002/advs.202301011.

216. M. Thway, A. P. Chen, H. Dai, et al., "Exploring Material Compositions for Synthesis Using Oxidation States," *arXiv* (2024), https://doi.org/10.48550/ARXIV.2405.08399.

217. M. Jansen, "A Concept for Synthesis Planning in Solid-State Chemistry," *Angewandte Chemie International Edition* 41, no. 20 (2002): 3746–3766.

218. O. I. Malyi, K. V. Sopiha, and C. Persson, "Energy, Phonon, and Dynamic Stability Criteria of Two-Dimensional Materials," *ACS Applied Materials & Interfaces* 11, no. 28 (2019): 24876–24884, https://doi.org/10.1021/acsami.9b01261.

219. W. Qian and C. Zhang, "Review of the Phonon Calculations for Energetic Crystals and Their Applications," *Energetic Materials Frontiers* 2, no. 2 (2021): 154–164, https://doi.org/10.1016/j.enmf.2021.03.002.

220. V. Ozoliņš and A. Zunger, "Theory of Systematic Absence of NaCl-Type (β-Sn-Type) High Pressure Phases in Covalent (Ionic) Semiconductors," *Physical Review Letters* 82, no. 4 (1999): 767–770, https://doi.org/10.1103/physrevlett.82.767.

221. G. Brauer, Ed., *Handbook of Preparative Inorganic Chemistry*, 2nd edition (New York, New York: Academic Press, Inc., 1963).

222. M. J. McDermott, B. C. McBride, C. E. Regier, et al., "Assessing Thermodynamic Selectivity of Solid-State Reactions for the Predictive Synthesis of Inorganic Materials," *ACS Central Science* 9, no. 10 (2023): 1957–1975, https://doi.org/10.1021/acscentsci.3c01051.

223. L. N. Walters, C. Zhang, V. P. Dravid, K. R. Poeppelmeier, and J. M. Rondinelli, "First-Principles Hydrothermal Synthesis Design to Optimize Conditions and Increase the Yield of Quaternary Heteroanionic Oxychalcogenides," *Chemistry of Materials* 33, no. 8 (2021): 2726–2741, https://doi.org/10.1021/acs.chemmater.0c02682.

224. S. B. Anantharaman, V. Rajkumar, S. Raghunandan, K. Hari Kumar, R. S. Kumar, and A. S. Gandhi, "Role of Thermodynamic Miscibility Gaps in Phase Selection in Sol-Gel Synthesis of Yttrium Silicates," *Journal of European Ceramic Society* 37, no. 15 (2017): 5001–5007, https://doi.org/10.1016/j.jeurceramsoc.2017.06.040.

225. T. Matsoukas, "Statistical Thermodynamics of Irreversible Aggregation: The Sol-Gel Transition," *Scientific Reports* 5, no. 1 (2015), https://doi.org/10.1038/srep08855.

226. S. Elbashir, M. Broström, and N. Skoglund, "Thermodynamic Modelling Assisted Three-Stage Solid-State Synthesis of High Purity β-$Ca_3(PO_{< pubYearyear = ε4ε > 4 < /pubYear >})_2$," *Materials & Design* 238 (2024): 112679, https://doi.org/10.1016/j.matdes.2024.112679.

227. J. Chen, S. R. Cross, L. J. Miara, J.-J. Cho, Y. Wang, and W. Sun, "Navigating Phase Diagram Complexity to Guide Robotic Inorganic Materials Synthesis," *arXiv* (2023), https://doi.org/10.48550/ARXIV.2304.00743.

228. M. Aykol, J. H. Montoya, and J. Hummelshoej, "Rational Solid-State Synthesis Routes for Inorganic Materials," *Journal of the American Chemical Society* 143, no. 24 (2021): 9244–9259, https://doi.org/10.1021/jacs.1c04888.

229. A. Zur and T. C. McGill, "Lattice Match: An Application to Heteroepitaxy," *Journal of Applied Physics* 55, no. 2 (1984): 378–386, https://doi.org/10.1063/1.333084.

230. H. Ding, S. S. Dwaraknath, L. Garten, P. Ndione, D. Ginley, and K. A. Persson, "Computational Approach for Epitaxial Polymorph Stabilization Through Substrate Selection," *ACS Applied Materials & Interfaces* 8, no. 20 (2016): 13086–13093, https://doi.org/10.1021/acsami.6b01630.

231. L. Yang, Z. Zhu, Y. Feng, et al., "Unveiling Phase Evolution of Complex Oxides Toward Precise Solid-State Synthesis," *Science Advances* 11, no. 30 (2025), https://doi.org/10.1126/sciadv.adx3927.

232. Z. Liu, Y. Jiang, Z. Zhang, et al., "Synthesis of Noble/Non-Noble Metal Alloy Nanostructures via an Active-Hydrogen-Involved Interfacial Reduction Strategy," *Nature Synthesis* 2, no. 2 (2023): 119–128, https://doi.org/10.1038/s44160-022-00217-y.

233. M. J. McDermott, S. S. Dwaraknath, and K. A. Persson, "A Graph-Based Network for Predicting Chemical Reaction Pathways in Solid-State Materials Synthesis," *Nature Communications* 12, no. 1 (2021), https://doi.org/10.1038/s41467-021-23339-x.

234. Y. Chen, H. Ji, M. Lu, et al., "Machine Learning Guided Hydrothermal Synthesis of Thermochromic VO_2 Nanoparticles," *Ceramics International* 49, no. 18 (2023): 30794–30800, https://doi.org/10.1016/j.ceramint.2023.07.035.

235. A. M. Hiszpanski, B. Gallagher, K. Chellappan, et al., "Nanomaterial Synthesis Insights from Machine Learning of Scientific Articles by Extracting, Structuring, and Visualizing Knowledge," *Journal of Chemical Information and Modeling* 60, no. 6 (2020): 2876–2887, https://doi.org/10.1021/acs.jcim.0c00199.

236. T. He, H. Huo, C. J. Bartel, Z. Wang, K. Cruse, and G. Ceder, "Precursor Recommendation for Inorganic Synthesis by Machine Learning Materials Similarity from Scientific Literature," *Science Advances* 9, no. 23 (2023), https://doi.org/10.1126/sciadv.adg8180.

237. D. A. Boiko, R. MacKnight, B. Kline, and G. Gomes, "Autonomous Chemical Research with Large Language Models," *Nature* 624, no. 7992 (2023): 570–578, https://doi.org/10.1038/s41586-023-06792-0.

238. Z. Zhang, Z. Ren, C.-W. Hsu, et al., "A Multimodal Robotic Platform for Multi-Element Electrocatalyst Discovery," *Nature* (2025), https://doi.org/10.1038/s41586-025-09640-5.

239. N. J. Szymanski and C. J. Bartel, "Computationally Guided Synthesis of Battery Materials," *ACS Energy Letters* 9, no. 6 (2024): 2902–2911, https://doi.org/10.1021/acsenergylett.4c00821.

240. B. J. Shields, J. Stevens, J. Li, et al., "Bayesian Reaction Optimization as a Tool for Chemical Synthesis," *Nature* 590, no. 7844 (2021): 89–96, https://doi.org/10.1038/s41586-021-03213-y.

241. Z. Yang, X. Liu, X. Zhang, P. Huang, K. S. Novoselov, and L. Shen, "Modeling Crystal Defects Using Defect-Informed Neural Networks," *arXiv* (2025), https://doi.org/10.48550/ARXIV.2503.15391.

242. N. C. Frey, D. Akinwande, D. Jariwala, and V. B. Shenoy, "Machine Learning-Enabled Design of Point Defects in 2D Materials for Quantum

and Neuromorphic Information Processing," *ACS Nano* 14, no. 10 (2020): 13406–13417, https://doi.org/10.1021/acsnano.0c05267.

243. I. Mosquera-Lois, S. R. Kavanagh, A. M. Ganose, and A. Walsh, "Machine-Learning Structural Reconstructions for Accelerated Point Defect Calculations," *npj Computational Materials* 10, no. 1 (2024), https://doi.org/10.1038/s41524-024-01303-9.

244. K. Jakob, A. Walsh, K. Reuter, and J. T. Margraf, "Learning Crystallographic Disorder: Bridging Prediction and Experiment in Materials Discovery," *ChemRxiv* (2025), https://doi.org/10.26434/chemrxiv-2025-f52qs.

245. M. H. Petersen, R. Zhu, H. Dai, et al., "Dis-Gen: Disordered Crystal Structure Generation," *arXiv* (2025), https://doi.org/10.48550/ARXIV.2507.18275.

246. B. Ma, X. Wei, C. Liu, et al., "Data Augmentation in Microscopic Images for Material Data Mining," *npj Computational Materials* 6, no. 1 (2020), https://doi.org/10.1038/s41524-020-00392-6.

247. W. Huang, Y. Jin, Z. Li, et al., "Auto-Resolving the Atomic Structure at van der Waals Interfaces Using a Generative Model," *Nature Communications* 16, no. 1 (2025), https://doi.org/10.1038/s41467-025-58160-3.

248. A. R. C. McCray, S. M. Ribet, G. Varnavides, and C. Ophus, "Accelerating Iterative Ptychography with an Integrated Neural Network," *Journal of Microscopy* (2025), https://doi.org/10.1111/jmi.13407.

249. F. L. Johansen, U. Friis-Jensen, E. B. Dam, K. M. O. Jensen, R. Mercado, and R. Selvan, "DECIFER: Crystal Structure Prediction from Powder Diffraction Data Using Autoregressive Language Models," *arXiv* (2025), https://doi.org/10.48550/ARXIV.2502.02189.

250. K. Choudhary, "DiffractGPT: Atomic Structure Determination from X-Ray Diffraction Patterns Using a Generative Pretrained Transformer," *The Journal of Physical Chemistry Letters* 16, no. 8 (2025): 2110–2119, https://doi.org/10.1021/acs.jpclett.4c03137.

251. T. Prein, W. O'Leary, A. F. Savvidou, E. Bourneix, and J. E. M. Laulainen, "A Synthesizability-Guided Pipeline for Materials Discovery," *arXiv* [cs.CE] (2025), https://doi.org/10.48550/arXiv.2511.01790.

252. Z. Song, S. Lu, M. Ju, Q. Zhou, and J. Wang, "Accurate Prediction of Synthesizability and Precursors of 3D Crystal Structures via Large Language Models," *Nature Communications* 16, no. 1 (2025): 6530, https://doi.org/10.1038/s41467-025-61778-y.

253. H. Noh, N. Lee, G. S. Na, and C. Park, "Retrieval-Retro: Retrieval-Based Inorganic Retrosynthesis with Expert Knowledge," *arXiv* [cs.LG] (2025), https://doi.org/10.48550/arXiv.2410.21341.

254. T. Prein, E. Pan, S. Haddouti, et al., "Retro-Rank-In: A Ranking-Based Approach for Inorganic Materials Synthesis Planning," *arXiv* [physics.chem-ph] (2025), https://doi.org/10.48550/arXiv.2502.04289.

255. O. Kononova, H. Huo, T. He, et al., "Text-Mined Dataset of Inorganic Materials Synthesis Recipes," *Scientific Data* 6, no. 1 (2019): 203, https://doi.org/10.1038/s41597-019-0224-1.

256. L. Longo, M. Brcic, F. Cabitza, et al., "Explainable Artificial Intelligence (XAI) 2.0: A Manifesto of Open Challenges and Interdisciplinary Research Directions," *Information Fusion* 106 (2024): 102301, https://doi.org/10.1016/j.inffus.2024.102301.

257. J. Schneider, "Explainable Generative AI (GenXAI): A Survey, Conceptualization, and Research Agenda," *Artificial Intelligence Review* 57, no. 11 (2024): 289, https://doi.org/10.1007/s10462-024-10916-x.

258. G. E. Moran and B. Aragam, "Towards Interpretable Deep Generative Models via Causal Representation Learning," *arXiv* (2025), https://arxiv.org/abs/2504.11609v1.

259. T. B. Nguyen, P. Le Nguyen, S. Lucey, and M. Hoai, "Region-Level Data Attribution for Text-to-Image Generative Models," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2025), 18825–18833.

260. Q. H. Nguyen, H. Phan, and K. D. Doan, "Unveiling Concept Attribution in Diffusion Models," *arXiv* (2024), https://arxiv.org/abs/2412.02542v3.

261. B. Wang, S. Zheng, J. Wu, J. Li, and F. Pan, "Inverse Design of Catalytic Active Sites via Interpretable Topology-Based Deep Generative Models," *npj Computational Materials* 11, no. 1 (2025): 147, https://doi.org/10.1038/s41524-025-01649-8.

262. C. J. Court, B. Yildirim, A. Jain, and J. M. Cole, "3D Inorganic Crystal Structure Generation and Property Prediction via Representation Learning," *Journal of Chemical Information and Modeling* 60, no. 10 (2020): 4518–4535, https://doi.org/10.1021/acs.jcim.0c00464.

263. L. Wang, C. Cheng, Y. Liao, Y. Qu, and G. Liu, "Training Free Guided Flow Matching with Optimal Control," *arXiv* (2025), https://doi.org/10.48550/arXiv.2410.18070.