Engineering 39 (2024) 13–17

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Views & Comments

Generative AI for Materials Discovery: Design Without Understanding

Jianjun Hu [a], Qin Li [b], Nihang Fu [a]

[a] Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA
[b] College of Big Data Statistics, Guizhou University of Finance and Economics, Guiyang 550025, China

1. Design as building-block assembling

In the 21st century, society is facing several fundamental challenges, including global climate change, the energy crisis, and public health crises such as cancers and coronavirus disease 2019 (COVID-19). Their solutions share a certain commonality: They all involve the discovery of novel atomic structures such as materials, molecules, proteins, and drugs. Designing such functional atomic structures is challenging due to the astonishing complexity of their inter-atomic interactions, sophisticated physical/chemical/geometric constraints and patterns in the formation of stable structures, and the relation between structures and their functions.

Like most engineering design activities, the mainstream paradigm of materials design is currently the rational design approach, which emphasizes a causal understanding of the structure–function relationship and depends on heuristic expert knowledge and explicit design rules. The typical "tinkering" design process starts with topology design with a limited number of prototypes and ends with parametric design. However, the traditional materials design paradigm is encountering increasing challenges in designing extraordinary functional materials that can effectively meet our needs: It usually leads to sub-optimal solutions in the huge chemical design space due to the limited search capability; it cannot handle the huge amount of implicit knowledge and constraints well and cannot exploit such rules for efficient design space exploration; it requires too many explicit design rules; and it presents difficulties in the design of highly constrained structures such as periodic inorganic crystals.

Here, we argue for a transformative shift from rational materials design to a data-driven deep generative materials design paradigm, in which known materials data are fed to deep generative models in order to enable the models to learn explicit and implicit knowledge of atomic structures and then exploit the models for efficient structure generation. This shift is inspired by two major, recent achievements in artificial intelligence (AI). The first of these is that the data-driven deep learning algorithm AlphaFold2 shows that deep learning models can solve the "finding-a-needle-in-the-haystack" issue inherent in the protein structure prediction problem by exploiting the learned implicit rules and constraints from known protein structures for efficient sampling in the protein structure space. The second achievement is that deep-learning-based artificial-intelligence-generated content (AIGC) technologies have been accelerating in generating authentic images, videos, texts, music, and human voices. Despite the apparent differences between digital artifacts and atomic structures, it can be seen in Table 1 that designing images and texts shares many characteristics with the task of designing proteins, materials, and molecules, in which building blocks of different levels are assembled together to form specific stable or meaningful structures that satisfy diverse grammatical, physical, chemical, or geometric constraints.

Compared with earlier generative design systems [1] that explicitly define the building blocks and generative rules or grammars, the deep generative design paradigm employs deep neural networks to learn the physical or chemical rules for assembling synthesizable and stable structures. Deep generative material design thus offers a new methodology and a philosophy that views materials in terms of dynamic processes and their outcomes and in which neural networks can be used to learn not only static inter-atomic interactions but also self-assembly and self-organization dynamic processes. Just as nature came to use the physical apparatus of DNA as the information carrier of synthesis rules for protein synthesis and biochemistry through evolution, deep neural networks can also be exploited similarly to achieve nature's way of materials design via learning the designing rules from known materials or computational simulations. Just as a female frog can

Table 1
A comparison of designing images and texts with designing proteins, materials, and molecules.

Domain	Building block	Structures
Image/video	Pixel → patches/strips	Objects/shapes
Language	Characters → words → sentences → paragraphs	Text
Proteins	Amino acids (AAs) → AA sequence → secondary structures	Three-dimensional protein structures
Materials	Atoms → bonds → polyhedrons	Crystal structures
Molecules	Atoms → bonds → groups	Organic molecules

https://doi.org/10.1016/j.eng.2024.07.008
2095-8099/© 2024 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

give birth to a frog without knowing how a frog is grown from a zygote through a developmental process, deep generative design can follow a similar design-without-understanding process for creative design.

2. Generative artificial intelligence for design

Generative AI started in the 1950s with Claude Shannon's Markov chains for language generation; it then evolved to Hopfield networks and Boltzmann machines for image and music generation in the 1980s and on to probabilistic graphical models such as hidden Markov models and the Gaussian mixture models in the 1990s. However, it was not until the emergence of generative adversarial networks (GANs) [2] in 2014 that generative AI could really create authentic images, videos, texts, and audio. Trained with a large number of existing data samples, modern deep neural network models can generate strikingly real digital artifacts that are now widely used in the AIGC community by chat generative pre-trained transformer (ChatGPT) or other software. These models can learn the delicate and sophisticated patterns, rhythms, styles, geometrical constraints, and/or interdependencies among building blocks from known samples through their networks and then exploit this implicit knowledge for the effective and efficient generation of new content.

Deep generative models have been increasingly applied to the generative design of DNAs and proteins (sequences and structures), molecules (composition and conformations) [3], materials (composition and structures), and engineering design [4]. Although the tokens or building blocks differ, most of these works share a set of common generative model architectures, as shown in Fig. 1.

The variational autoencoder (VAE) model is composed of an encoder that maximally compresses the raw input information x into a lower-dimension latent space $\boldsymbol{z}$ so that the decoder can reconstruct the input as x' with a minimal amount of reconstruction error; the latent space is regularized by minimizing the Kullback–Leibler divergence between the returned latent distribution and a standard Gaussian distribution. The GAN model consists of two components: a generator and a discriminator. The generator learns to create new samples that mimic the distribution of the training data, while the discriminator learns to distinguish between real and fake samples. The two components are trained together in a process called adversarial training, where the generator attempts to create increasingly realistic materials and the discriminator becomes better at distinguishing between real and fake samples. Compared with the VAE model, which has literal reconstruction loss, GAN models are capable of capturing the semantic information of the training sample distribution and generating diverse new generations.

The diffusion model works by destroying training data through the successive addition of Gaussian noise and then learning to recover the data by reversing this noising process using a denoising neural network. After training, the model can generate data by simply passing randomly sampled noise through the learned denoising process. One of the key advantages of the diffusion model is that it can generate highly realistic and diverse images without the need for complex adversarial training. It can also generate images with controllable attributes by conditioning the diffusion process on additional inputs, such as class labels or semantic embeddings. Another major category of generative models is autoregressive network models such as generative pre-trained

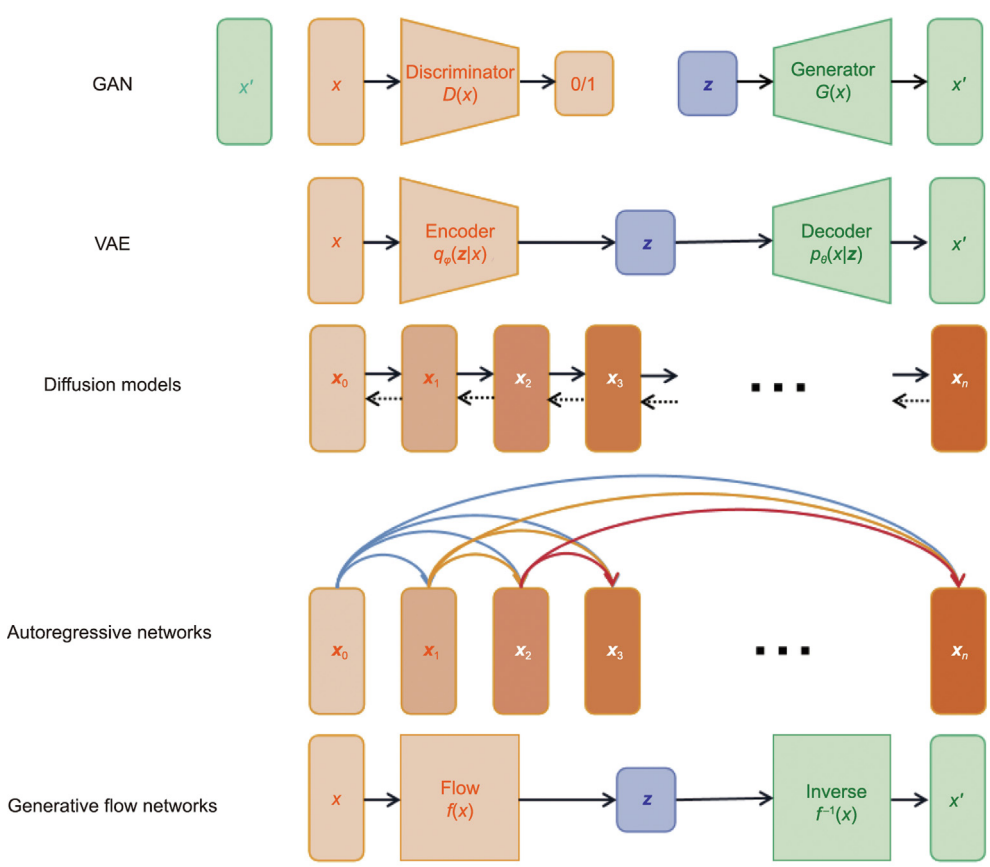

Fig. 1. Deep generative neural network models. x is the training sample; x' is the generated sample; $\boldsymbol{z}$ represents a latent vector; $\boldsymbol{x}_i$ represents latent vectors (i = 1, 2, . . ., n). VAE: variational autoencoder.

transformer (GPT), which are a type of language model used for text generation that generates text by predicting the next word in a sequence based on the previous words in the sequence. This model is trained on a large corpus of text and learns the probability distribution of each word, given its previous words. These models can also be used for image generation in a pixel-by-pixel way.

Generative flow network (GFlowNet) models [5] are probabilistic models that construct objects by iteratively sampling a probability distribution over possible building blocks and adding the next. They build objects at a frequency proportional to the reward. GFlowNet models are trained to mirror the reward function learned by the surrogate model; the generative model can then be used to generate many structures using the GFlowNet, which are then prioritized using the surrogate model. This makes GFlowNet good for sampling intelligently in a large chemical design space.

3. Generative AI for materials discovery

In traditional materials design, researchers rely on trial and error to test new materials, which can be a time-consuming and expensive process. Generative materials design aims to speed up this process by learning and exploiting chemical/geometric/physical constraints for the efficient generation of new materials that meet specific criteria, such as synthesizability, stability, conductivity, or optical properties. These materials can then be synthesized and tested in the lab, potentially leading to new discoveries. Fig. 2 shows representative generative models for materials design.

3.1. Generative design of material compositions

The goal of material composition design is to discover compositions that can be synthesized into stable crystal structures. These hypothetical compositions can be used to guide experimental synthesis or be fed to crystal-structure-prediction algorithms, if available, to obtain their stable structures. Material compositions can also be fed to composition-based machine learning models of material properties such as elastic constants or band gaps for composition screening. However, material composition generation is non-trivial due to three major challenges:

(1) This is a "finding-a-needle-in-the-haystack" problem [6]. The space for three-element materials exceeds 10^9 combinations, and the four-element space exceeds 10^{12} combinations; moreover, a majority of these combinations do not even satisfy basic chemical rules—only 6.7‰ and 7.8‰ satisfy the constraints of charge neutrality and electronegativity balance for four- and five-element compositions, respectively, assuming the number of atoms for each element is ≤ 8.

(2) The relationship between a given composition and its synthesizability and capability to form stable structures is complex due to the many sophisticated chemical and geometric constraints.

(3) Without the structure information, it is difficult to evaluate the quality of generated candidates, such as the synthesizability or structural stability, in order to screen promising candidates.

One of the earliest generative design studies utilized real vectors to encode the numbers of atoms for elements in a composition [7] using both conditional VAE and conditional GAN models. However, the models tended to generate mostly chemically invalid compositions, which could not be easily screened using the basic chemical rules of charge neutrality and electronegativity balance due to their real-valued atom number representation. Realizing the importance of the discrete encoding of the number of atoms of elements, we proposed the MatGAN material composition generator [8], which is based on a GAN generative deep learning model and a one-hot binary matrix representation of material compositions. This encoding scheme greatly facilitates the convolution neural networks of the GAN to learn the sophisticated chemical rules or patterns within the known materials in the Inorganic Crystal Structure Database (ICSD) and Materials Project database. It was found that the percentage of chemically valid (i.e., charge-neutral and electronegativity-balanced) samples out of all generated compositions by MatGAN reached 84.5% when the GAN model was trained with the samples in the ICSD, even though no such chemical rules are explicitly enforced in our GAN model. This indicates MatGAN's capability to learn implicit chemical composition rules, which allows it to exploit learned implicit constraints in generating promising compositions that are more likely to form stable and synthesizable compounds.

A material composition such as $SrTiO_3$ can be naturally represented as a sequence of element symbols, such as Sr Ti O O O;

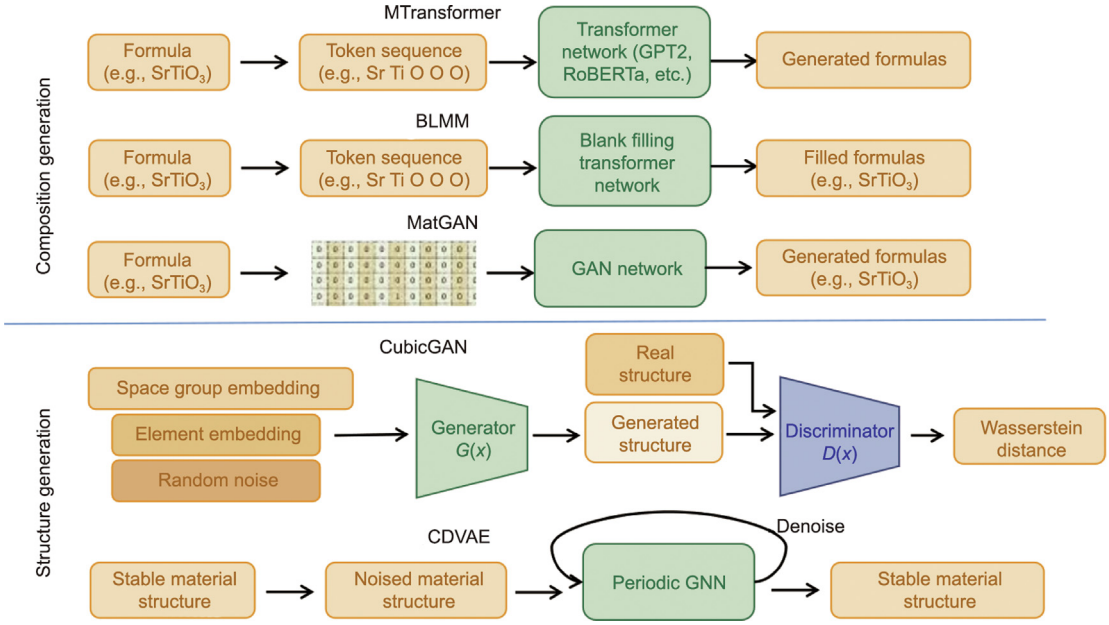

Fig. 2. Representative frameworks for deep generative materials design. RoBERTa: robustly optimized bidirectional encoder representation from transformers (BERT) approach; BLMM: blank language models for materials; CDVAE: crystal diffusion variational autoencoder; GNN: graph neural network.

this inspired us to build composition generators using modern generative language models such as GPT and bidirectional encoder representation from transformers (BERT). These models have achieved huge success in the generation of texts, molecules, and protein sequences. In Ref. [9], we developed and benchmarked seven modern language models (including GPT, GPT-2, GPT-Neo, GPT-J, blank language models for materials (BLMM), bidirectional and autoregressive transformers (BART), and robustly optimized BERT approach (RoBERTa)) as MTransformer algorithms for composition generation. Six different datasets with/without non-charge-neutral or balanced electronegativity samples from the ICSD, Open Quantum Materials Database (OQMD), and Materials Project database were used to train these models. We found that the causal language model (e.g., GPT)-based material transformers could generate chemically valid material compositions of which as high as 97.54% were charge-neutral and 91.40% were electronegativity-balanced, exhibiting an enrichment more than six times higher than a baseline pseudo-random sampling algorithm. This finding demonstrates the capability of language models to capture the implicit chemical rules and constraints for the formation of chemically valid materials compositions. To further improve the interpretability of the learned language model, we applied a blank-filling probabilistic language model to the material composition generation problem [10]. Our crystal transformer algorithm demonstrated the highest generation performance in terms of the percentages of charge neutrality and electronegativity. It also allows the designers to tinker with a given material composition in order to explore the design space based on its learned materials chemistry, which is useful for materials doping.

One of the key decisions in material composition generative design is how to evaluate the generation performance, especially when the structure information is unavailable. While charge neutrality, electronegativity balance, and predicted formation energy can be used as the first level of performance measures, such models can also be evaluated using novelty, uniqueness, and recovery rate, among which the latter is especially useful: If a generator can rediscover most of the leave-out compositions that have been synthesized before, it is a strong indication of its generation power.

3.2. Generative design of crystal structures

The *de novo* generation of novel synthesizable and stable crystal materials is a challenging task due to the highly sophisticated relationships from the composition to stable structures. Unlike organic molecules and other structures, crystal materials tend to have periodic structures with high symmetry, which leads to a highly constrained multimodal design space. The significantly higher diversity of the element types (> 85) and the complex inter-atomic interactions exacerbate the problem.

The field of data-driven crystal structure generative design has recently been emerging rapidly based on a series of deep generative models with a variety of crystal encodings. iMatGen, which is short for image-based materials generator, is one of the earliest algorithms for crystal structure generation. This VAE-based model was trained with structures of the V_xO_y family and was able to discover 40 relatively stable structures with $E_{hull} < 80$ meV per atom (E_{hull}: the energy above hull; 1 meV = 1.602×10^{-22} J). Kim et al. [11] demonstrated that a GAN-based generative model using point clouds as inputs could be used to generate stable Mg–Mn–O ternary compounds. Using a voxelized crystal representation of iMatGen, Court et al. [12] trained a conditional deep-feature-consistent VAE for the generation of new crystals that went beyond a specific chemical system. However, their model needed a U-Net segmentation network to map the predicted electron density into atomic sites, which hindered its performance.

To exploit the symmetry of crystals, we developed CubicGAN [13], a GAN-based generative model for generic cubic crystal structure generation. Our model was used to generate 506 new-prototype stable hypothetical materials, such as Li_6N_6Cl and $CaCO_6$, as verified by phonon dispersion density functional theory (DFT) calculations. We further showed that, by incorporating additional symmetry principles and physics-based constraints into the generative models, the generation performance could be significantly improved, as shown by our physics guided crystal generative model (PGCGM) algorithm [14], which can generate the crystal structures of more than 30 space groups. Another major progress in generative crystal material design is the crystal diffusion variational autoencoder (CDVAE) [15], which trains a decoder that can generate materials in a diffusion process. The neural-network-based diffusion model makes it possible to move atomic coordinates toward a lower energy state with appropriate atom types to satisfy bonding preferences between neighbors. It also models interactions across periodic boundaries and respects permutation, translation, rotation, and periodic invariances, which can further improve its performance. This model has been used to discover thousands of hypothetical two-dimensional (2D) materials; however, its capability in three-dimensional (3D) material generation (especially high-symmetry materials) needs further improvement [14], which has been partially achieved in the newest model, MatterGen [16].

While our review of generative materials design has focused on inorganic materials, the same principles and models have also been widely applied to the generative design of proteins [17], organic materials [3,18], and architected materials, for which both language models and diffusion models are intensively used for forward and inverse design [19].

4. Challenges and opportunities

Generative material design is still in an emerging stage, and there are many significant challenges and related opportunities that must be addressed, ranging from algorithm models to training datasets or design objectives.

Challenge 1: Controlled generative design with multiobjective functions and complex physicochemical or geometric constraints. This challenge is compounded by the fact that the simulation codes for most performances are not even differentiable, which makes it difficult to incorporate those performance objectives into the loss function to guide the model training and generation.

Challenge 2: Mixing large language models (LLMs) with generative models for generative materials design. While deep language models have been combined with generative models for protein [20] and simplified molecular-input line-entry system (SMILES)-based molecule design [18], it is challenging to apply this to crystal material design due to the high symmetry constraints and the difficulty of finding stable and synthesizable crystal structures.

Challenge 3: Fast validation models for screening final candidate materials. Many material design constraints (explicit or implicit) cannot be incorporated into a generative model and can only be validated over generated samples. However, filtering criteria such as synthesizability, mechanical stability, and thermodynamic stability are all very difficult to compute. The consideration of manufacturability makes it even worse. The question of how to improve the hit rate of generating physically and chemically feasible materials and train fast and accurate evaluation models is a key unsolved problem.

Challenge 4: Creative generative design. The objective of design is to find novel materials with exceptional properties while current generative models are trained to generate samples similar

to the training sets. Another obstacle is that neural networks are good at interpolation but not at extrapolation, which makes the performance prediction of out-of-distribution samples difficult and misleads generation models to generate invalid designs.

Challenge 5: Generative design with limited datasets. Due to the high costs of experiments or DFT calculations, most current materials datasets are small in terms of macro material properties such as thermal conductivity and piezoelectricity. However, there are a large number of unlabeled material structures. Opportunities exist here to exploit advanced machine learning techniques such as pretraining and physics-informed neural networks to address these issues. There is also promise in exploring surrogate models trained with mixed fidelity datasets and studying out-of-distribution machine learning models with high generalization performance for novel samples.

With the emergence of more advanced deep generative AI techniques, as demonstrated by OpenAI's Sora for text-to-video generation, generative design research for materials and structures will undergo a transformative revolution in the coming years, which can greatly assist in addressing the global challenges of climate change, energy, and human health.

References

[1] McCormack J, Dorin A, Innocent T. Generative design: a paradigm for design research. In: Redmond J, Durling D, de Bono A, editors. Proceedings of Futureground, Design Research Society International Conference; 2004 Nov 17–21; Melbourne, VIC, Australia. Clayton: Monash University Publishing; 2004.

[2] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM 2020;63(11):139–44.

[3] Bilodeau C, Jin W, Jaakkola T, Barzilay R, Jensen KF. Generative models for molecular discovery: recent advances and challenges. Wiley Interdiscip Rev Comput Mol Sci 2022;12(5):e1608.

[4] Regenwetter L, Nobari AH, Ahmed F. Deep generative models in engineering design: a review. J Mech Des 2022;144(7):071704.

[5] Bengio E, Jain M, Korablyov M, Precup D, Bengio Y. Flow network based generative models for non-iterative diverse candidate generation. In: Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Wortman Vaughan J, editors. Advances in neural information processing systems 34. Red Hook: Curran Associates, Inc.; 2021. p. 27381–94.

[6] Davies DW, Butler KT, Jackson AJ, Morris A, Frost JM, Skelton JM, et al. Computational screening of all stoichiometric inorganic materials. Chem 2016;1(4):617–27.

[7] Sawada Y, Morikawa K, Fujii M. Study of deep generative models for inorganic chemical compositions. 2019. arXiv:1910.11499.

[8] Dan Y, Zhao Y, Li X, Li S, Hu M, Hu J. Generative adversarial networks (GAN) based efficient sampling of chemical composition space for inverse design of inorganic materials. npj Comput Mater 2020;6(1):84.

[9] Fu N, Wei L, Song Y, Li Q, Xin R, Omee SS, et al. Material transformers: deep learning language models for generative materials design. Mach Learn Sci Technol 2023;4(1):015001.

[10] Wei L, Li Q, Song Y, Stefanov S, Siriwardane EMD, Chen F, et al. Crystal transformer: self-learning neural language model for generative and tinkering design of materials. 2022. arXiv:2204.11953.

[11] Kim S, Noh J, Gu GH, Aspuru-Guzik A, Jung Y. Generative adversarial networks for crystal structure prediction. ACS Cent Sci 2020;6(8):1412–20.

[12] Court CJ, Yildirim B, Jain A, Cole JM. 3-D inorganic crystal structure generation and property prediction via representation learning. J Chem Inf Model 2020;60(10):4518–35.

[13] Zhao Y, Al-Fahdi M, Hu M, Siriwardane EMD, Song Y, Nasiri A, et al. High-throughput discovery of novel cubic crystal materials using deep generative neural networks. Adv Sci 2021;8(20):2100566.

[14] Zhao Y, Siriwardane EMD, Wu Z, Fu N, Al-Fahdi M, Hu M, et al. Physics guided deep learning for generative design of crystal materials with symmetry constraints. npj Comput Mater 2023;9(1):38.

[15] Xie T, Fu X, Ganea OE, Barzilay R, Jaakkola T. Crystal diffusion variational autoencoder for periodic material generation. 2021. arXiv:2110.06197.

[16] Zeni C, Pinsler R, Zügner D, Fowler A, Horton M, Fu X, et al. MatterGen: a generative model for inorganic materials design. 2023. arXiv:2312.03687.

[17] Buehler MJ. Generative pretrained autoregressive transformer graph neural network applied to the analysis and discovery of novel proteins. J Appl Phys 2023;134(8):084902.

[18] Luu RK, Wysokowski M, Buehler MJ. Generative discovery of *de novo* chemical designs using diffusion modeling and transformer deep neural networks with application to deep eutectic solvents. Appl Phys Lett 2023;122(23):234103.

[19] Lew AJ, Buehler MJ. Single-shot forward and inverse hierarchical architected materials design for nonlinear mechanical properties using an attention–diffusion model. Mater Today 2023;64:10–20.

[20] Ni B, Kaplan DL, Buehler MJ. ForceGen: end-to-end *de novo* protein generation based on nonlinear mechanical unfolding responses using a language diffusion model. Sci Adv 2024;10(6):eadl4000.