

DiffSyn: A Generative Diffusion Approach to Materials Synthesis Planning

Elton Pan¹, Soonhyoung Kwon², Sulin Liu¹, Mingrou Xie²,
Alexander J. Hoffman¹, Yifei Duan¹, Thorben Prein³,
Killian Sheriff¹, Yuriy Roman-Leshkov², Manuel Moliner⁴,
Rafael Gomez-Bombarelli¹, Elsa Olivetti^{1*}

¹Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

³Department of Chemistry, Technische Universität München, München, 80333, Germany.

⁴Instituto de Tecnología Química, Universitat Politècnica de València-Consejo Superior de Investigaciones Científicas, Valencia, 46022, Spain.

*Corresponding author(s). E-mail(s): elsao@mit.edu;

Contributing authors: eltonpan@mit.edu; s1105hk@mit.edu; sulinliu@mit.edu; mrx@mit.edu; ajhoff29@mit.edu; duanyf99@mit.edu; t.prein@tum.de; ksheriff@mit.edu; yroman@mit.edu; mmoliner@itq.upv.es; rafagb@mit.edu;

Abstract

The synthesis of crystalline materials, such as zeolites, remains a significant challenge due to a high-dimensional synthesis space, intricate structure-synthesis relationships and time-consuming experiments. Considering the *one-to-many* relationship between structure and synthesis, we propose **DiffSyn**, a generative diffusion model trained on over 23,000 synthesis recipes spanning 50 years of literature. **DiffSyn** generates probable synthesis routes conditioned on a desired zeolite structure and an organic template. **DiffSyn** achieves state-of-the-art performance by capturing the multi-modal nature of structure-synthesis relationships. We apply **DiffSyn** to differentiate among competing phases and generate optimal synthesis routes. As a proof of concept, we synthesize a UFI material

using DiffSyn-generated synthesis routes. These routes, rationalized by density functional theory binding energies, resulted in the successful synthesis of a UFI material with a high Si/Al_{ICP} of 19.0, which is expected to improve thermal stability and is higher than that of any previously recorded.

Keywords: Generative models, Materials synthesis, Porous materials

1 Introduction

Materials discovery lays the foundation for modern technologies, from catalysis to electronics [1]. Recent large-scale computational searches of chemical composition and structures [2–5] have uncovered millions of potentially stable, synthesizable materials (*what* to synthesize) [2, 6–8]. However, finding viable synthesis routes remains a bottleneck in materials discovery (*how* to synthesize) [9–12] because there are many synthesis parameters (composition, conditions, etc.) that interact in complex ways. Moreover, the compute required for atomistic simulations scales poorly with system size, precluding accurate modeling of the underlying physical phenomena in complex materials (i.e., thermodynamics and kinetics) [13]. Consequently, there is continued interest in machine learning (ML) approaches to directly learn from experimental synthesis data to predict materials synthesis parameters at lower computational cost [10, 11, 14].

Materials synthesis prediction presents a unique challenge for ML for several reasons. First, structure-synthesis relationships are *one-to-many*—i.e., a single target structure may form through *multiple* possible synthesis recipes. Second, the inverse relationship (synthesis-structure) is also one-to-many—i.e., a single recipe may result in the formation of a mixture of products (competing phases) due to the complex interplay of thermodynamic and kinetic pathways [13]. Capturing this phase competition is crucial to selectively synthesize single-phase materials instead of mixtures. Third, complex non-linear interactions exist between synthesis parameters, such as temperature and time (Fig. A10) [11], requiring approaches that model joint probabilities across multiple synthesis parameters. Predictions must capture relationships among variables to make tradeoffs between parameters leveraging physical information about materials synthesis (e.g., crystallization kinetics).

Previous ML approaches to predict synthesis have predominantly used regression approaches [11, 15–17], which deterministically map a representation (e.g., composition [10, 11], structural features [16], graphs [15]) of a material to its synthesis parameters. These approaches are limited because the deterministic mapping is incompatible with the one-to-many nature of structure-synthesis relationships and assume independence between synthesis parameters (Fig. A4) [18, 19]. These factors limit the predictive accuracy of regression approaches and motivate a shift to generative models, which can sample a complex distribution that accounts for the non-linear interactions between parameters in high-dimensional synthesis space.

To address these challenges, we introduce a diffusion model for materials synthesis. Diffusion models are a powerful class of generative models that were demonstrated

to generate novel, high-quality images conditioned on text [20, 21]. Diffusion models can be guided at each step of the denoising process toward a specific objective (e.g., target material) [22, 23]. Unlike generative adversarial networks that suffer from mode collapse [24], diffusion models can generate diverse outputs because they are trained to denoise data (e.g., a synthesis route) that have been corrupted with noise. Other generative approaches, namely variational autoencoders and normalizing flows, have limited expressivity due to one-step decoding and affine invertible layers, respectively. In contrast, the iterative denoising process renders diffusion models highly expressive, which enables high sample quality [23]. This high expressivity may enable diffusion models to capture boundaries in synthesis space between competing phases.

We propose **DiffSyn**, a diffusion approach to materials synthesis prediction and demonstrate it on zeolites, which are crystalline, microporous materials with applications in catalysis, adsorption, and ion exchange. Zeolite synthesis is challenging due to its high dimensionality (Fig. 1a), with numerous variables influencing the synthesis outcome (Fig. A1). Moreover, multiple modes of valid synthesis routes exist for a given structure (Fig. A5). Progress in zeolite synthesis has focused on trial-and-error experiments guided by domain heuristics [25]. We demonstrate the first application of a guided diffusion model for materials synthesis prediction, which exhibits state-of-the-art performance compared to regression-based and other deep generative approaches. We show that the performance of **DiffSyn** arises from its ability to capture the one-to-many and multi-modal nature of structure-synthesis relationship in materials. We experimentally validate our approach by synthesizing the UFI zeolite based on **DiffSyn**-generated synthesis routes. We rationalize these routes using density functional theory (DFT) calculations of inorganic cations that guide UFI synthesis. Together, these results indicate that **DiffSyn** learns the underlying chemistry that influences synthesis outcomes implicitly from published synthesis recipes.

2 Results

2.1 DiffSyn framework for materials synthesis planning

2.1.1 Chemically guided diffusion model

DiffSyn leverages a chemically guided diffusion model for predicting materials synthesis routes with the target zeolite structure c_{zeo} and the organic structure-directing agent (OSDA) c_{OSDA} as inputs (Fig. 1b). An OSDA is an organic molecule that templates the zeolite's pores (Fig. A1e), directing the synthesis toward the formation of a specific structure. Prior work has shown that the optimal OSDA to synthesize a given structure can be identified from atomistic simulations [26–31]. The goal is to learn $p(x_{\text{comp}}, x_{\text{cond}} | c_{\text{zeo}}, c_{\text{OSDA}})$ to generate an *ensemble* of synthesis routes consisting of gel compositions $\{x_{\text{comp}}\}$ and synthesis conditions $\{x_{\text{cond}}\}$ given a target structure and OSDA, as shown in Fig. A2 (also known as synthesis parameters, defined in Table A1). Fig. A4 shows an example of the predicted vs. ground truth synthesis parameters (green points). The prediction of synthesis parameters is an under-determined problem with multiple possible valid synthesis routes $\{x_{\text{comp}}, x_{\text{cond}}\}$ for each c_{zeo} and c_{OSDA} .

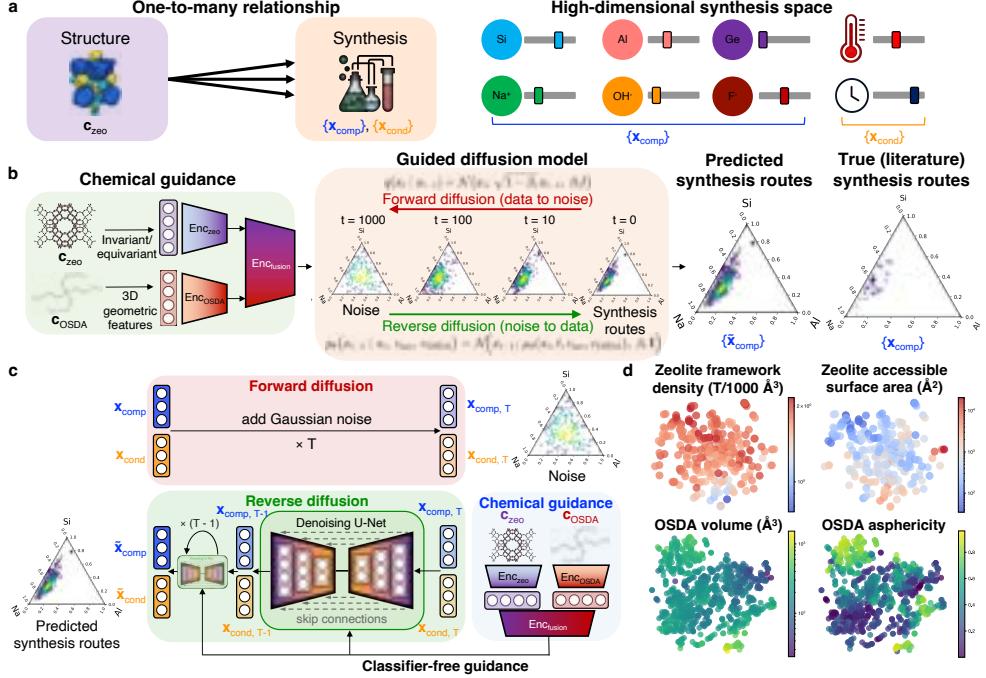


Fig. 1 **DiffSyn: A diffusion approach to materials synthesis planning.** (a) Materials, such as zeolites, often have an *one-to-many* structure-synthesis relationship, where a structure c_{zeo} can be synthesized via multiple possible synthesis routes in a high-dimensional synthesis space (compositions $\{x_{\text{comp}}\}$ and conditions $\{x_{\text{cond}}\}$). (b) Overview of DiffSyn: Given a desired zeolite structure c_{zeo} and an organic structure directing agent COSDA , the model $p_\theta(x_{\text{comp}}, x_{\text{cond}} | c_{\text{zeo}}, \text{COSDA})$ generates an *ensemble* of synthesis routes $\{\tilde{x}_{\text{comp}}\}$ and $\{\tilde{x}_{\text{cond}}\}$ via reverse diffusion (green arrow) that matches ground-truth synthesis routes $\{x_{\text{comp}}\}$, accurately capturing the aforementioned one-to-many structure-synthesis relationship with which previous methods struggle. Note: Only $\{\tilde{x}_{\text{comp}}\}$ is shown. (c) Model architecture: Two separate encoders embed the zeolite material c_{zeo} and organic template COSDA before their latent representations are concatenated. The joint representation steers the generation process from noise ($x_{\text{comp},T}, x_{\text{cond},T} \sim \mathcal{N}(0, \mathbf{I})$) to realistic synthesis routes specific to the zeolite and OSDA via classifier-free guidance by learning $\mu_\theta(x_t, t, c_{\text{zeo}}, \text{COSDA}) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} \epsilon_\theta(x_t, t, c_{\text{zeo}}, \text{COSDA}) \right)$ where $\epsilon_\theta(\cdot)$ is a conditional U-Net trained to predict the noise at time t (Section 4.2). Note that x_{comp} and x_{cond} are jointly noised and denoised. (d) PCA of learned representations of zeolites and OSDAs. The colorbar refers to a specific physical property (as shown in each title).

During training, the forward diffusion process (red arrow in Fig. 1b) adds Gaussian noise to x_{comp} and x_{cond} , progressively mapping them to a Gaussian distribution (noise). During inference, the reverse diffusion process (green arrow in Fig. 1b) starts from Gaussian noise and iteratively denoises using a U-Net [32] conditioned on chemical guidance (see Section 2.1.2) via classifier-free guidance (Fig. 1c). After T timesteps of denoising, the model generates synthesis routes for a desired structure. This denoising process can be seen in the improvement of generation metrics (e.g., Wasserstein

distance and COV-P, defined in Section 4.4.1) throughout the reverse diffusion process (Fig. A7). We train **DiffSyn** on the ZeoSyn dataset [14], which consists of 23,961 synthesis recipes, 233 zeolite topologies, and 921 OSDAs (Supplementary Section A).

2.1.2 Representation learning of materials

DiffSyn integrates a dual-encoder approach consisting of separate encoders (Enc_{zeo} and Enc_{OSDA}) for the zeolite structure and OSDA, respectively (Fig. 1b). We use two representations of the zeolite structure: invariant geometric features and an equivariant graph neural network (EGNN). The invariant geometric features are physical descriptors (e.g., pore volume) calculated from the zeolite structure using the `Zeo++` package. The EGNN encoder directly learns a representation from a graph of the zeolite crystal structure (Supplementary Section D). For the OSDA, we perform molecular geometry relaxation and calculate its physicochemical descriptors (e.g., volume and shape) (Section 4.1).

Fig. 1d shows that the respective encoders learn smooth and continuous latent spaces with respect to the properties of zeolites and OSDAs. A comprehensive set of properties plotted in embedding space can be found in Fig. A8 and A9, indicating chemically meaningful representations of zeolites and OSDAs. These representations are concatenated before a fusion encoder ($\text{Enc}_{\text{fusion}}$) learns a joint representation. We refer to the joint representation as chemical guidance (Fig. 1b). Chemically meaningful representations are pivotal in steering the diffusion model to generate realistic synthesis routes for a desired materials structure. This representation enables **DiffSyn** to generate synthesis parameters that reflect synthesis routes unseen in training, which have been reported in the literature (Fig. 2e).

2.1.3 Influence of chemical guidance in diffusion model

Classifier-free guidance [22] is a critical component of **DiffSyn**, where the chemical guidance steers the generation process by reweighing the unconditional score function with a conditional score function (Section 4.3). We probe the influence of two key hyperparameters related to classifier-free guidance: the probability of the chemical guidance being omitted in score estimation during training (p_{uncond}) and the guidance strength that weighs the conditional score relative to the unconditional score during inference (w ; Eq. 7).

Higher values of p_{uncond} and w amplify the conditional score, making the sampling process more dependent on the chemical guidance. This increased dependence on chemical guidance can result in over-constrained outputs, reducing diversity and potentially missing valid synthesis routes. Lower values of p_{uncond} and w reduce the influence of the chemical guidance, potentially decreasing specificity and generating synthesis recipes that are less tailored to the target structure, but improving the diversity of recipes. We find that the balance between diversity and quality of generated synthesis routes occurs at $p_{\text{uncond}} = 0.1$ and $w = 1.0$ (Fig. A6).

2.2 Modeling structure-synthesis relationships

We evaluate **DiffSyn** against a suite of previously published approaches in materials synthesis planning. These baseline models fall into three categories: regression-based approaches (average minimum distance (AMD) [16] and Bayesian neural networks (BNN) [33]); classical generative models (Gaussian mixture model (GMM) [34]); and deep generative models (conditional generative adversarial network (GAN) [35], normalizing flow (NF) [36], and variational autoencoder (VAE) [11, 19]). More information about these baselines is included in Section 4.4.2. We compare approaches using Wasserstein distance (lower is better), which measures the distance between the generated and ground-truth distributions of literature reported synthesis parameters for unseen zeolite-OSDA systems. Additionally, we propose the coverage metric COV-F1. The model should maximize both COV-P (precision) and COV-R (recall) simultaneously. Therefore, their harmonic mean (COV-F1) measures the degree of generated recipes being both realistic and diverse (ranges from 0–1, higher is better). Detailed explanations of these metrics are in Section 4.4.1 and Supplementary Section C.

2.2.1 Generative approaches better model structure-synthesis relationships

Wasserstein distances show that deep generative models such as GAN, NF, VAE and **DiffSyn** outperform the classical approaches, with **DiffSyn** outperforming the next best baseline (VAE) by over 25%. Classical generative approaches like GMM do not perform much better than a random baseline (Fig. 2a) while a probabilistic regression model (BNN) performs better than GMM.

The COV-F1 of 12 synthesis parameters are shown in Fig. 2b. The models perform better on synthesis parameters related to heteroatoms (Si/Al, Al/P, Si/Ge, Si/B), cations (Na^+/T , K^+/T) and anions (F^-/T , OH^-/T). However, they struggle to predict crystallization time, which could be attributed to anthropogenic factors; crystallization times are subject to human bias, where experimentalists test and report “rounded” numbers [37]. This bias results in the ground-true time distribution peaking at specific intervals, hence rendering time prediction more challenging.

Deep generative approaches (VAE, NF, and **DiffSyn**) outperform regression-based approaches (AMD and BNN). We hypothesize that generative models have superior performance due to better recall (higher COV-R, Fig. A3). Meanwhile, **DiffSyn** outperforms other deep generative models due to higher precision (higher COV-P, Fig. A3), where the diffusion model generates higher quality synthesis routes. Interestingly, **DiffSyn** achieves the lowest mean absolute error (MAE) for 10/12 synthesis parameters (Fig. 2c), despite not being explicitly trained on the MAE objective like the regression-based models.

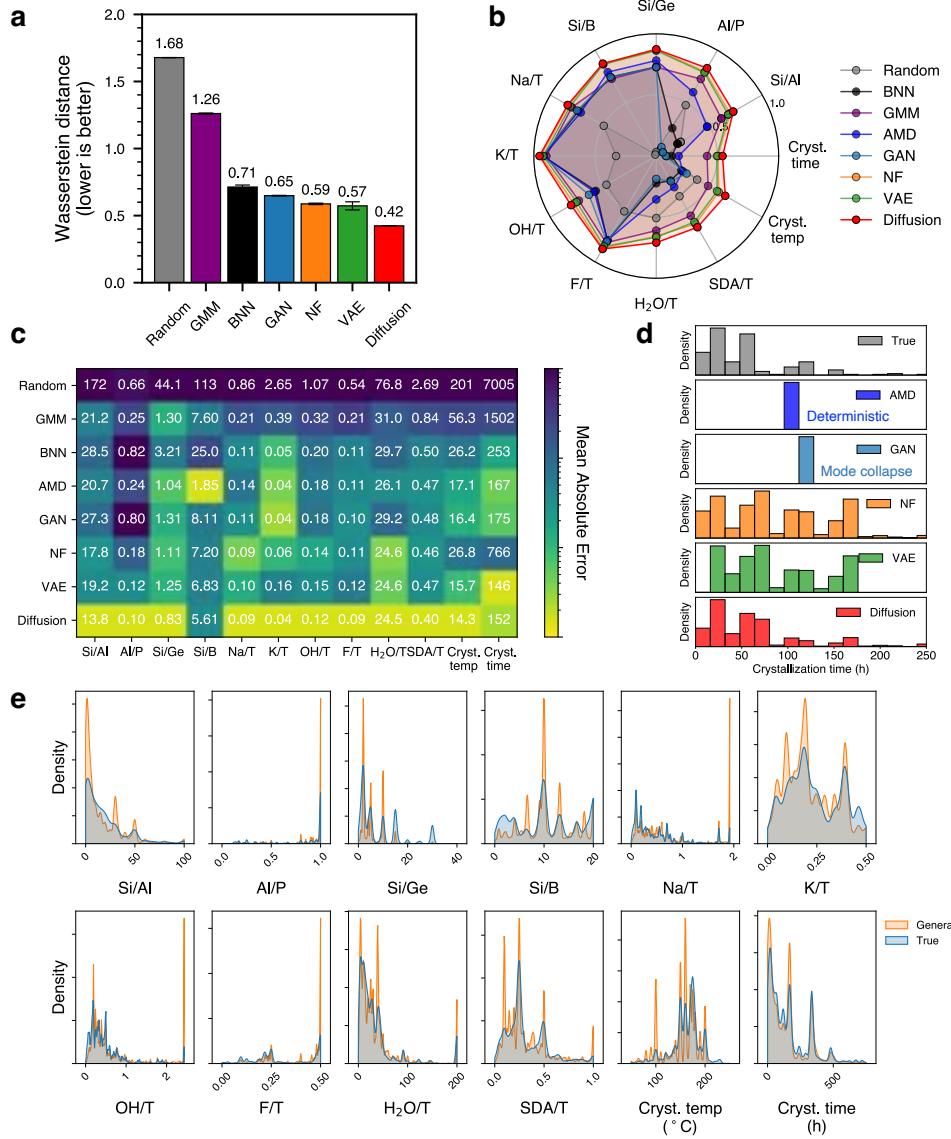


Fig. 2 DiffSyn achieves state-of-the-art performance in materials synthesis prediction. **(a)** Wasserstein distance (lower is better) between generated and literature synthesis routes. Metric is averaged across all test systems. Error bars correspond to standard deviation across 3 independent training runs. **(b)** COV-F1 (higher is better, ranging from 0 to 1) of individual synthesis parameters. **(c)** Mean absolute error between the means of distribution of generated and literature synthesis parameters. **(d)** Distributions of predicted and true synthesis parameters across various different modeling approaches. AMD (purple) is regression-based (outputs deterministic, single-point prediction), while GAN (light blue) suffers from mode collapse. NF and VAE output distributions do not match the ground-truth (grey). DiffSyn (red) accurately captures the true distribution (grey) of the synthesis parameters. **(e)** DiffSyn-generated synthesis routes (orange) and distribution of synthesis routes reported in literature (blue). Synthesis parameters are aggregated across all unseen zeolite-OSDA systems.

2.2.2 Rationalizing the superior performance of generative approaches

For a given target material, there are a range of possible values for each synthesis parameter (e.g., temperature) at which a material can be synthesized (i.e. *synthesis window*). We show the distribution of crystallization times from published syntheses (grey, Fig. 2d).

This distribution of synthesis parameters renders structure-synthesis relationship *one-to-many* instead of one-to-one. Regression-based AMD is deterministic, and thus outputs a point prediction that comes from a weighted average of the distribution (Fig. 2d). Generative models like GAN also output a point prediction as they often suffer from mode collapse [24]. Although generative models such as NF and VAE address mode collapse, they lack expressivity and fail to accurately capture the ground-truth literature distribution. In contrast, DiffSyn captures the ground-truth distribution. Furthermore, we compare the predicted and true joint distributions of *multiple* synthesis parameters for all of these approaches for the AEL zeolite (Fig. A4). Only deep generative approaches (NF, VAE, and DiffSyn) capture the ground-truth joint distribution of crystallization temperatures and times for the AEL structure, with DiffSyn most accurately capturing that joint distribution. DiffSyn captures most of the ground truth points, including some outliers; however, DiffSyn fails to predict points in a minor mode (bottom right of Guided Diffusion panel in Fig. A4), which are extreme outliers (e.g., low crystallization temperature and long crystallization time).

The distribution of synthesis parameters is also *multi-modal*. We plot the PCA of all synthesis parameters for the aforementioned AEL structure in Fig. A5, which shows that the true distribution has multiple modes—2 in this case. Regression-based models (AMD, BNN) predict only one of the modes. GMM predicts synthesis routes that are far out of distribution. GAN suffers from mode collapse to one of the modes. NF and VAE capture both modes, but also generate a large number of false positives. This behavior arises from the low expressivity of VAEs and NFs, which use one-step decoding and affine invertible layers, respectively. In contrast, DiffSyn accurately predicts the true distribution because it generates high-quality and diverse outputs. Consequently, DiffSyn generates synthesis routes that overlap with unseen literature-reported synthesis parameters (Fig. 2e). For a discussion on diversity of generated samples, please refer to Supplementary Section B.

2.2.3 Learning chemically meaningful relationships

We perform an unsupervised, hierarchical clustering of zeolite structures based on their learned representations (Fig. 3a), and observe distinct clusters according to their corresponding structural features (e.g., number of channels, largest free sphere diameter). The clustering indicates that the zeolite encoder has learned to separate structurally distinct materials. Consequently, the chemical guidance (Fig. 1b), which requires learning good representations of the zeolite and OSDA, guides the generative process towards the desired target material.

Given that DiffSyn learns the joint distribution of multiple synthesis parameters, we inspect two synthesis parameters (crystallization temperature and time) in Fig.

[A10](#) for two unseen zeolite-OSDA systems. An inverse relationship is observed between generated temperatures and times. This observation aligns well with the Arrhenius equation

$$k = Ae^{-\frac{E_a}{RT}}, \quad (1)$$

where crystallization time (related to rate k) is inversely related to temperature T . In addition, generated H₂O/T (here, T refers to tetrahedral framework atoms or T-atoms, rather than temperature) and framework density (FD_{Si}) of zeolite structure correlate positively (Spearman's coefficient: 0.673; Fig. [A11](#)). This finding agrees with Villaescusa's rule [\[38\]](#), which states that denser phases (higher FD_{Si}) are favored at lower concentrations of T-atoms (higher H₂O/T), showing that **DiffSyn** has learned domain-specific rules in materials synthesis. The model predictions also follow the thermodynamics of zeolite formation, where the generated crystallization temperature and framework density (FD_{Si}) of zeolite structure positively correlate (Spearman's coefficient: 0.931; Fig. [A12](#)). This finding agrees with the thermodynamic argument from Ostwald's rule of stages, which states that higher temperatures enable the synthesis to overcome the activation barrier to form more stable structures with higher framework densities [\[14, 39\]](#).

2.3 Case studies

We compare **DiffSyn**-generated synthesis routes to literature-reported synthesis routes for diverse zeolite-OSDA systems that are synthetically interesting and industrially useful. The generated routes for these unseen systems (MWW, BEC, and a pair of competing phases—FAU and LTA) provide evidence of **DiffSyn** learning meaningful synthesis-structure relationships (Fig. 3). Other systems (MTT and ATO) are analyzed in the SI (Fig. [A13](#) and [A14](#), respectively).

We first consider the MWW phase, a unique two-dimensional structure with 10-membered rings and large cavities, with applications including isomerization [\[40\]](#) and aromatization [\[41\]](#). The generated OH⁻/T, K⁺/T, H₂O/T, SDA/T, and crystallization temperature/time overlap significantly with ground-truth synthesis parameters (Fig. 3b). We also test the model on a significantly different structure: BEC, a large-pore zeolite. BEC has a three-dimensional pore topology with intersecting 12-membered ring channels, with applications in isomerization [\[42\]](#) and epoxidation [\[43\]](#). **DiffSyn**-generated synthesis parameters closely agree with synthesis parameters reported in the literature, particularly Si/Ge, F⁻/T, and crystallization temperature/time (Fig. 3c). This prediction aligns with reports that Ge and F⁻ stabilize the *d_{4r}* composite building unit of the BEC structure during synthesis [\[44\]](#). This finding suggests that **DiffSyn** learns how particular heteroatoms or synthesis conditions favor the formation of specific building units within zeolites. However, the generated synthesis parameters do not always fully recall the ground-truth for BEC synthesis. For example, the model fails to predict the full range of possible SDA/T values that have been identified in past recipes (SDA/T = 0.15–0.25; Fig. 3c).

Typically, materials synthesis aims to produce a single framework; however, if two phases during synthesis compete, some recipes result in two or more phases. Here, we applied **DiffSyn** to predict OSDA-free synthesis routes for FAU and LTA zeolites [\[45\]](#). Synthesis routes generated by **DiffSyn** align closely with literature-reported recipes

(Fig. 3d). Notably, **DiffSyn** accurately predicts the phase boundary region (green) between FAU and LTA in OSDA-free conditions, delineating the synthesis space under which the competing phases form. This result shows that **DiffSyn** not only accurately captures the forward relationship (structure-synthesis), but also the decision boundaries of the *inverse* relationship (synthesis-structure), hence demonstrating its potential to enable phase-selective synthesis. Similarly, we demonstrate this capability for another pair of competing phases (ERI and KFI) in Fig. A15. Taken together, these case studies illustrate **DiffSyn**'s ability to generalize to a variety of zeolite frameworks and their corresponding chemistries.

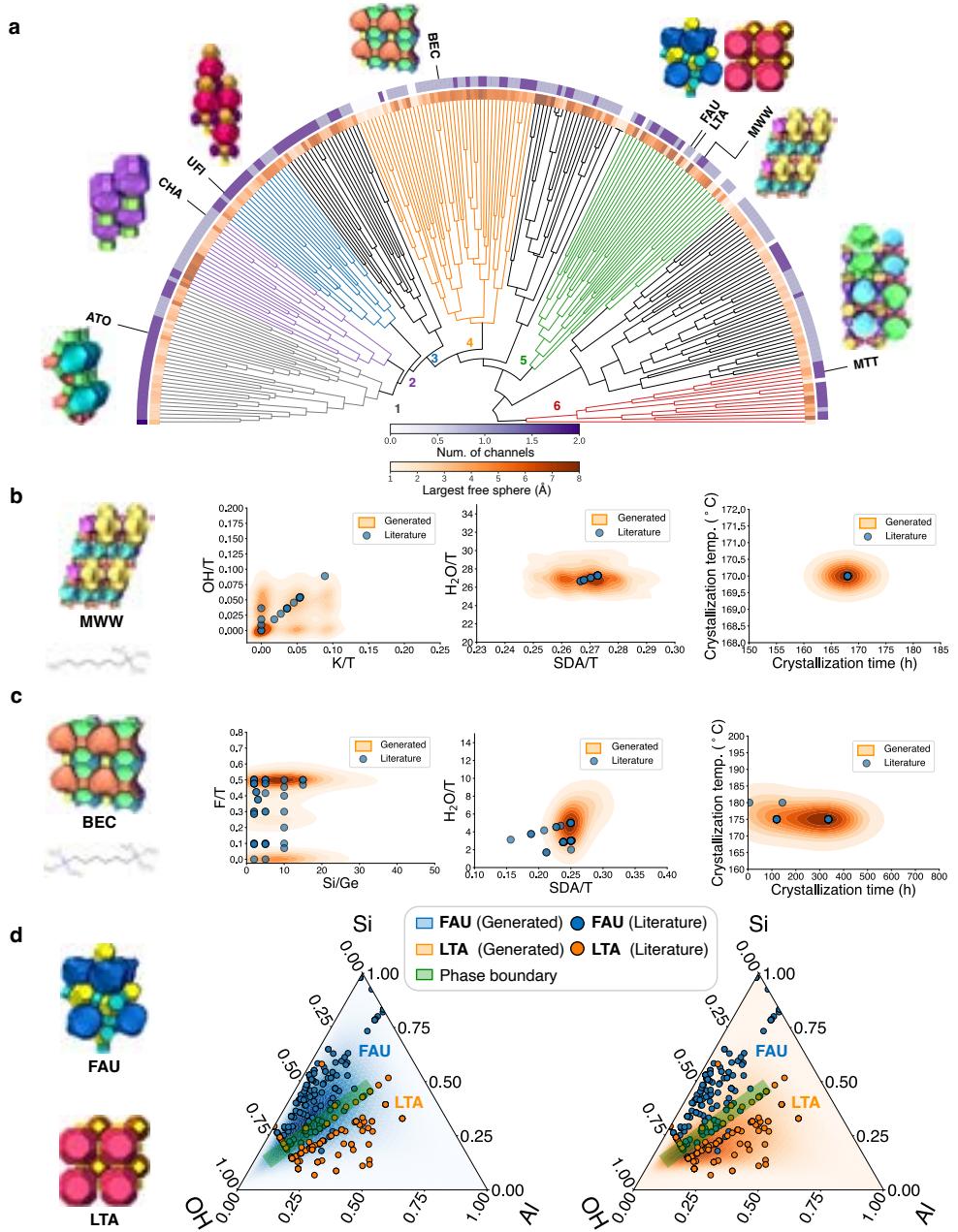


Fig. 3 Case studies on unseen materials systems. (a) Hierarchical clustering of zeolite structures. Splits are obtained via agglomerative hierarchical clustering of learned zeolite representations. Each structure (with a 3-letter code as its name) is colored by its number of channels (NC) in purple, and its largest free sphere (LFS) in orange. This leads to several distinct clusters: (1) high NC, low LFS; (2) low NC, low+high LFS; (3) high NC, low+high LFS; (4) low NC, high LFS; (5) low+high NC, high LFS; (6) high NC, high LFS. Generated synthesis routes for unseen materials systems (b) MWW structure (in cluster 5) with N,N,N-trimethylhexan-1-aminium as OSDA. (c) BEC structure (in cluster 4) with pentane-1,5-diyl-bis(trimethylammonium) as OSDA. Orange heatmaps refer to synthesis routes generated by *DiffSyn*, while blue points refer to literature-reported synthesis routes. (d) Competing phases FAU and LTA (in cluster 5). Heatmaps refer to generated routes, while points refer to literature-reported synthesis routes. Notice that the model accurately predicts the phase boundary (green shaded region) between FAU and LTA.

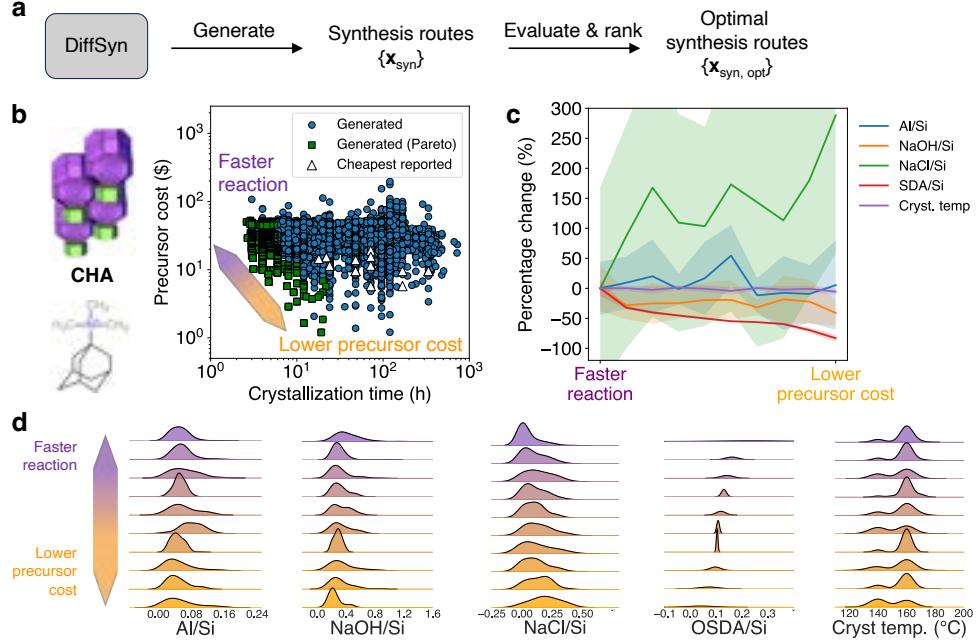


Fig. 4 Generating optimal synthesis routes. (a) DiffSyn generates synthesis routes $\{x_{\text{syn}}\}$, which are then evaluated based on some metric of interest (cost, crystallization time) and ranked to find an optimal synthesis route. (b) Precursor cost and crystallization time of synthesis routes for the CHA structure (in cluster 2 of Fig. 3a) with TMAda as the OSDA. Circles refer to DiffSyn-generated routes, triangles refer to 20 cheapest routes reported in literature. Notice the Pareto front (green squares) trades off cost and time. (c) Percentage change in generated synthesis parameters along the Pareto front. Solid lines and shaded regions refer to mean and standard deviation of the synthesis parameter, respectively. (d) Distributions of generated synthesis parameters along the Pareto front.

2.4 Generating optimal synthesis routes

We employ DiffSyn to generate more feasible synthesis routes (Fig. 4a), by evaluating and ranking the generated synthesis routes based on precursor cost and crystallization time (see Table A2 for details). We apply this approach to the synthesis of CHA, using trimethyladamantyl ammonium (TMAda) as the OSDA [46] and generate an ensemble of synthesis routes with corresponding precursor costs and crystallization times (Fig. 4b). Among the generated routes, there are Pareto optimal routes that trade off low precursor cost and fast synthesis time. These two objectives are optimal in different regions in the synthesis space—the OSDAs that produce CHA crystals faster tend to be more expensive (Fig. A16). Some of these Pareto-optimal routes generated by DiffSyn have lower crystallization time and lower precursor cost compared to the 20 least expensive synthesis routes reported in literature.

We inspect the evolution of synthesis parameters on the Pareto front. Notably, Al/Si and NaOH/Si remain relatively constant (Fig. 4c). In contrast, SDA/Si decreases as we prioritize lower cost over faster reaction as the OSDA typically drives precursor

cost. The distributions of generated synthesis parameters also change along the Pareto front (Fig. 4d). This analysis reveals that raising the crystallization temperature from 140 to 160 °C while simultaneously increasing NaOH/Si and OSDA/Si would favor faster reaction (Fig. 4d). This assessment shows how varying the joint distributions of synthesis parameters can accelerate synthesis (synthesizing desired structure in shorter time at higher temperature).

2.5 Experimental and DFT validation

We validate **DiffSyn** by experimentally synthesizing a UFI zeolite from recipes generated by our model with Kryptofix 222 (K222) as the OSDA (Fig. 5) [47]. UFI has potential applications in industrially relevant reactions (e.g., selective catalytic reduction of NO_x [48]). The UFI-K222 system has not been reported in prior literature and is not present in the training dataset, and hence serves as a test of **DiffSyn**'s out-of-distribution generalization.

Fig. 5a shows the PCA of 1000 **DiffSyn**-generated synthesis routes for UFI (orange), which constitutes a subspace of literature-reported synthesis routes for all zeolites (grey). We retrieve the k -nearest neighbors ($k = 5$) of generated UFI synthesis from the literature-reported synthesis routes (Fig. A17). Among the retrieved synthesis, the top 2 most similar frameworks (PAU and RHO) share a common *lta* composite building unit (CBU) with UFI (Fig. A18). Beyond this, the remaining frameworks do not share any common CBUs with UFI. This observation is supported by a previous work [49] reporting that majority of competing phases do not share any CBUs, suggesting that structure-synthesis relationships are complex and cannot be rationalized with building units alone.

We visualize individual synthesis parameters of generated routes in Fig. 5e (orange histograms), including gel compositions (Si/Al, Ge/Si, B/Si, Na⁺/Si, K⁺/Si, H₂O/Si, F⁻/Si) and reaction conditions (crystallization temperature and time). For inorganic cations, **DiffSyn** recommends high Na⁺/Si and low K⁺/Si for UFI synthesis (Fig. 5e). We rationalize this observation by calculating binding energies (ΔE_b) of inorganic cations (Na⁺ and K⁺) in building units of UFI (*wbc* and *rth*) using density functional theory (DFT) as shown in Fig. 5b (details in Section 4.6). The calculations reveal that Na⁺ binds more strongly to *wbc* and *rth* ($\Delta E_b^{sol} = -70$ and +13 kJ/mol) compared to K⁺ ($\Delta E_b^{sol} = -50$ and +63 kJ/mol), suggesting that Na⁺ ions stabilize building units of UFI better than K⁺. These binding energies explain the high Na⁺/Si and low K⁺/Si in **DiffSyn**-generated routes.

We experimentally test the **DiffSyn**-generated routes for UFI, resulting in four successful experimental syntheses of the UFI material (Fig. 5a, blue). Powder X-ray diffraction (XRD) patterns for synthesized UFI samples closely match the simulated diffraction pattern (Fig. 5c), confirming that the crystallized material has the UFI structure. The resultant crystals had a measured Si/Al_{ICP} of 19.0, the highest yet recorded in UFI synthesis (see Section 4.5.4 for details on Si/Al_{ICP}). This composition is important because higher Si/Al_{ICP} is associated with higher thermal stability under catalytic conditions [50]. Additionally, the crystals exhibit a house-of-cards morphology (Fig. 5d), making them highly promising catalysts for a diverse range of applications.

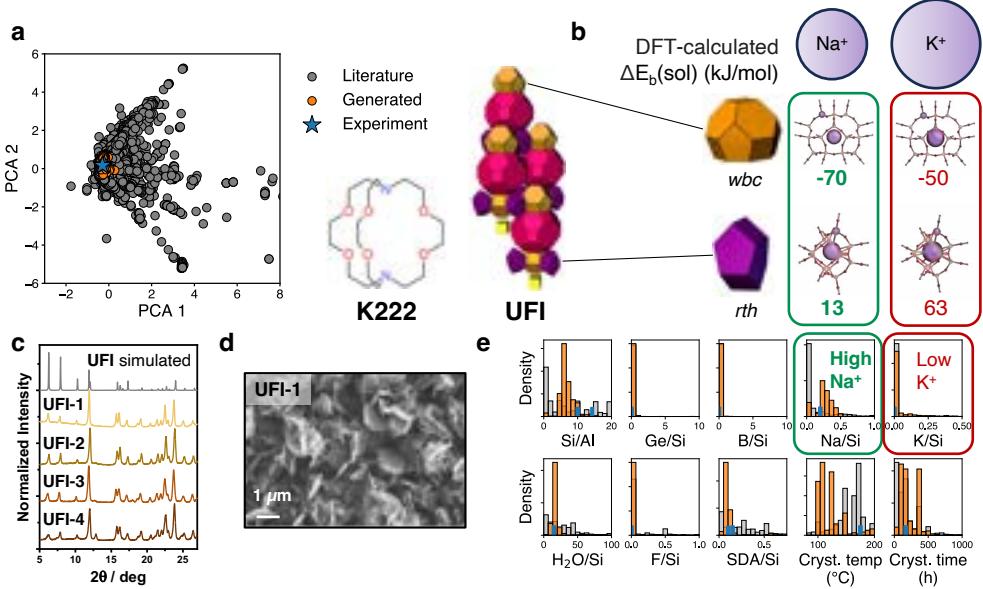


Fig. 5 Experimental and DFT validation. We leverage **DiffSyn** to generate synthesis routes for the UFI structure, which resulted in successful crystallization of the desired structure (a) PCA of synthesis parameters of all routes reported literature (grey), **DiffSyn**-generated for UFI (orange) and experimentally-verified (blue) in this work. (b) DFT-calculated binding energies (ΔE_b) between cations (Na^+ , K^+) and building units of UFI (*rth*, *wbc*). Noticed Na^+ has stronger (more negative) ΔE_b than K^+ , which suggests that Na^+ (compared to K^+) favors UFI formation. This rationalizes why **DiffSyn** suggested synthesis routes with high Na^+/Si and low K^+/Si for UFI (in (e)). (c) Simulated and measured XRD patterns of the synthesized UFI samples. (d) SEM image of UFI. (e) Distributions of individual synthesis parameters. Blue are experimentally-verified synthesis routes, orange is **DiffSyn**-generated, and grey is all literature-reported routes.

While the experimentally verified synthesis routes fall within the generated distributions, the model generated unusually low crystallization temperatures with a major mode at 100–150°C (Fig. 5e). The generated minor mode at 175°C was chosen based on domain expertise because higher temperatures accelerate crystallization kinetics and higher temperatures (>170°C) avoid the formation of the undesired LTA competing phase. The higher temperature allowed us to use a relatively shorter crystallization time of 7 days (168 h). These considerations underscore a powerful synergy between model and human expertise to achieve the desired synthesis outcome. Nonetheless, **DiffSyn** generated 9/10 synthesis parameters as the major mode. The successful synthesis of UFI, enabled by **DiffSyn**, is a testament to **DiffSyn** learning the intricacies of zeolite synthesis without explicit chemical encodings and highlights its ability to recommend suitable synthesis parameters for unseen materials systems.

3 Discussion

Zeolite synthesis is a complex task with a high-dimensional space where synthesis parameters interact. This complexity underscores the need for generative models that capture the intricacies of materials synthesis and circumvent expensive first-principles approaches. By capturing the one-to-many and multi-modal nature of structure-synthesis relationships, **DiffSyn** outperforms previous methods for zeolite synthesis prediction.

DiffSyn requires the OSDA to be known *a priori*. Including the OSDA as an input ensures the model captures the influence of the organic template. Recent work on OSDA design has enabled the selection of suitable OSDA inputs into **DiffSyn** [29, 30, 51]. The combination of OSDA design tools and **DiffSyn** could enable an end-to-end predictive workflow for zeolite synthesis. While our work focuses solely on continuous variables, the synthesis outcome can also be influenced by categorical/discrete variables, such as the precursor choice [9, 52, 53] and presence of seed crystals. These discrete variables present a compelling opportunity for future work in ML for synthesis modeling—e.g., discrete diffusion [54], autoregressive models [55–58], reinforcement learning [59, 60] for sampling discrete synthesis parameters.

For diffusion models, the inference speed is slower than other generative approaches due to sequential denoising. However, since the bottleneck of materials synthesis is experimental synthesis (e.g., reaction time of weeks), inference speed is not a significant problem. Even if sampling speed becomes a bottleneck, techniques such as denoising diffusion implicit model (DDIM) sampling [61] can trade off sampling quality for faster sampling. Flow matching models [62] could be an alternative to balance quality and speed. Notably, training **DiffSyn** requires optimizing hyperparameters for classifier-free guidance such as p_{uncond} and w as shown in Fig. A6, which may be challenging if compute is limited.

The successful synthesis of UFI enabled by **DiffSyn** exemplifies the potential of our approach to guide experimental zeolite synthesis. However, a comprehensive validation of **DiffSyn** would ideally also demonstrate that recipes substantially diverging from the generated ones fail to yield the target phase; such an evaluation is beyond the scope of this study and is reserved for future work. We hypothesize diffusion models could be applied to other materials systems, particularly when data are plentiful. This approach signifies a shift from regression to generative models for materials synthesis, where the latter is needed to effectively model one-to-many structure-synthesis relationships. This work has enabled the prediction of probable and diverse synthesis pathways of microporous materials, signifying a step toward bridging computational materials design (*what* to synthesize) and synthesis planning (*how* to synthesize).

4 Methods

4.1 Zeolite and OSDA representations

Zeolite We adopt two different approaches of encoding the zeolite structure. First, invariant features (e.g., ring sizes, largest included sphere) are retrieved from the International Zeolite Association (IZA) database [63]. These features serve as inputs into a

MLP encoder. Second, an equivariant graph neural network (EGNN) [64] encodes the zeolite as a graph (see Section D). **OSDA** We optimized each OSDA in the gas phase with the MMFF94 force field from RDKit-generated conformers [65]. Each OSDA is featurized using its physicochemical descriptors (e.g., molecular volume and 2D shape descriptors) averaged over all conformers [28, 29]. The OSDA features are defined in Table A3. The zeolite embedding is concatenated with the OSDA embedding and further encoded using the fusion encoder before the joint embedding guides the reverse diffusion process to generate synthesis routes (Fig. 1). Model performance across the denoising diffusion trajectory can be found in Fig. A7.

4.2 Denoising diffusion probabilistic models

Denoising diffusion probabilistic models (DDPMs) [66] are generative models that use a diffusion process to generate data by reversing a forward process that incrementally adds noise to the data. The forward process gradually corrupts the data x_0 into a noisy sample x_t by adding Gaussian noise in a Markov chain:

$$q(x_t | x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t \mathbf{I}\right) \quad (2)$$

where β_t is the variance at timestep t . It is often useful to sample x_t directly from x_0 :

$$q(x_t | x_0) = \mathcal{N}\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t) \mathbf{I}\right) \quad (3)$$

where $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$. The generative process learns to reverse this corruption step by step (Fig. 1b), sampling from a distribution $p_\theta(x_{t-1} | x_t)$ parameterized by a neural network. This reverse process can also be represented as a Gaussian distribution

$$p_\theta(x_{t-1} | x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \quad (4)$$

Here, we learn the mean while fixing the variance as

$$\mu_\theta(x_t, t) = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(x_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right), \quad \Sigma_\theta(x_t, t) = \beta_t \mathbf{I} \quad (5)$$

where $\epsilon_\theta(x_t, t)$ is a neural network trained to predict the noise ϵ . In this work, we use a U-Net (Fig. 1c) [32]. The training objective is to minimize a variational bound $\mathbb{E}_{t, x_0, \epsilon} [\|\epsilon - \epsilon_\theta(x_t, t)\|^2]$, which results in the diffusion model

$$p_\theta(x_0) = \int p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t) dx_{1:T} \quad (6)$$

4.3 Chemically guided diffusion model

In standard guided diffusion models, a classifier is used to guide the generation process by adjusting the score to steer the model towards specific target classes. In contrast,

classifier-free guidance [22] eliminates the need for a separate classifier by conditioning the diffusion model directly on the desired attributes. During training, the score function \tilde{s}_θ is trained both with and without conditioning c using a null token \emptyset . This training is done by randomly setting c to the unconditional null token \emptyset with some probability p_{uncond} . Sampling is then performed using a linear combination of the conditional and unconditional score estimates:

$$\tilde{s}_\theta(x_t, t, c) = (1 + w)s_\theta(x_t, t, c) - ws_\theta(x_t, t, \emptyset) \quad (7)$$

where c refers to the chemical guidance from zeolite and OSDA embeddings shown in Fig. 1, while w is the strength of the chemical guidance. Supplementary Section D contains the implementation details.

4.4 Model evaluation

4.4.1 Metrics

For each test zeolite-OSDA system, we sample 1000 synthesis routes using the model and compute the following metrics with reference to *unseen* synthesis routes reported in literature. **Wasserstein distance** measures the distance between two probability distributions by finding the minimum cost to move probability mass from one distribution to another [67]. The Wasserstein metric captures differences in both the location and the shape of distributions (including spread and the presence of multiple modes). While Kullback–Leibler (KL) divergence may be used, the Wasserstein distance is chosen as it fulfills all requirements of a metric: non-negativity, identity of indiscernibles, symmetry and triangle inequality. Moreover, a significant drawback of KL divergence is its behavior with distributions that do not have overlapping support. If there is any point where one distribution has a zero probability and the other has a non-zero probability, the KL divergence can become infinite or undefined. Wasserstein distance, in contrast, provides a meaningful and finite distance even for distributions with non-overlapping supports. Wasserstein distance has been widely adopted in generative modeling (e.g., WGANs [68] employ a form of Wasserstein distance) because it provides a smooth, sensitive measure of distributional differences. Small improvements in how well the model captures tails or secondary modes are reflected by a lower Wasserstein distance, guiding us during model development to favor settings that capture the full distribution. **Coverage** Inspired by Xu [69], we use two coverage metrics, COV-R (recall) and COV-P (precision), to measure the similarity between sets of generated and literature-reported synthesis for each zeolite-OSDA system. Intuitively, COV-R measures the fraction of literature synthesis routes being correctly predicted, and COV-P measures the fraction of generated synthesis routes being probable. COV-F1 is computed as the harmonic mean of COV-R and COV-P. Refer to Supplementary Section C for a detailed justification of the metrics.

4.4.2 Baselines

Random A random dummy baseline corresponds to picking a random point in synthesis space. **AMD** Schwalbe-Koda et al. proposed a deterministic regression-based

approach using average minimum distance (AMD) for zeolite structural featurization for synthesis prediction task [16]. **BNN** Bayesian neural networks [33] extend standard neural networks by incorporating Bayesian inference by treating network weights as probability distributions, enabling a distribution of outputs. We implement a classical generative approach, Gaussian mixture model (**GMM**) [34], which models data probabilistically as a sum of Gaussians (each with its mean and covariance). We also implement deep generative baselines: conditional variational autoencoder (**VAE**) [11, 19], conditional generative adversarial network (**GAN**) [35] and conditional normalizing flow (**NF**). For NF, we use real-valued non-volume preserving (RealNVP) transformations [36].

4.5 Experimental methods

4.5.1 Synthesis of UFI

We used colloidal silica (Ludox AS-40, 40 wt%, Aldrich), aluminum hydroxide (Al(OH)_3 , SPI Pharmacy), sodium hydroxide (NaOH, 50%, Aldrich), with Kryptofix 222 (98%, Ambeed) and tetramethylammonium hydroxide pentahydrate ($\text{TMAOH}\cdot\text{5H}_2\text{O}$, 97%, Aldrich) as OSDAs to synthesize UFI. The OSDAs are selected by the method described in [47]. The synthesis mixture had the molar composition 1.0 SiO_2 , 0.033 Al_2O_3 , 0.1 Na_2O , 15 H_2O , 0.15 Kryptofix 222, 0.025 TMA_2O . 0.021 g of aluminum hydroxide (Al(OH)_3) was dissolved in a solution containing 0.64 g of NaOH, 6.9 g of deionized water, 0.36 g of $\text{TMAOH}\cdot\text{5H}_2\text{O}$. After the solution became clear, 6 g of Ludox AS-40 and 2.31 g of Kryptofix 222 were added. 4 wt% of H-UFI seed material was included. The mixture was stirred and aged at room temperature for 24 h, then transferred to Teflon-lined 23 mL autoclaves and heated at 175 °C under dynamic conditions for 7 days. The resulting zeolites were recovered by centrifugation (12,000 rpm, 10 min), washed with deionized water three times, and dried overnight at 110 °C. Samples were then heated at a ramp rate of 1 °C/min to 580 °C in flowing dry air (100 mL/g of zeolite) for 6 h to remove OSDA molecules.

4.5.2 X-ray diffraction

Zeolite crystal structures were analyzed using powder X-ray diffraction (XRD) with a Bruker D8 diffractometer (Cu-K α radiation, $\lambda = 1.5418 \text{ \AA}$, 40 kV, 40 mA). Diffraction patterns were recorded over a 2θ range of 5°–45°, with an angular step size of 0.02° and a scanning rate of 4°/min to confirm phase purity.

4.5.3 Scanning electron microscopy

The morphology of the calcined zeolite crystals was observed using a Zeiss Merlin High-Resolution scanning electron microscope (SEM). Zeolite samples were prepared as fine powders and mounted on carbon tape. Images were collected at 2.0 kV, 100 pA, and a working distance of 6.7 mm using the HE-SE2 detector in High-Resolution Column mode.

4.5.4 Elemental analysis

The elemental composition of silicon, and aluminum ($\text{Si}/\text{Al}_{\text{ICP}}$) of synthesized UFI zeolites was determined using inductively coupled plasma atomic emission spectroscopy (ICP-AES, Agilent 5100). $\text{Si}/\text{Al}_{\text{ICP}}$ is an important property of zeolites as high $\text{Si}/\text{Al}_{\text{ICP}}$ is correlated with high thermal stability [50]. Samples (10 mg) were digested in 100 μL of hydrofluoric acid (48 wt%, Sigma-Aldrich) for 24 hours, followed by dilution to 10 g with 2 wt% aqueous nitric acid (GFS Chemicals). Calibration curves were constructed with six-point standards using ICP solutions of 1,000 ppm Si, Al, and Na in 2 wt% HNO_3 (Sigma-Aldrich, TraceCERT).

4.6 DFT methods

Density functional theory (DFT) calculations were conducted using the ORCA package (v5.0.4) [70, 71]. For all calculations, the $\omega\text{B97X-D}$ hybrid functional [72, 73] was employed along with the def2 triple- ζ basis set with polarization (def2-TZVP) [74]. During electronic optimization, wavefunctions were converged when energy changes were less than 10^{-8} Ha. Geometry optimizations were performed until the energy varied by less than 5×10^{-6} Ha and the forces on all atoms were below 3×10^{-4} Ha bohr $^{-1}$, adhering to the default ORCA optimization protocol. Where indicated, the SMD solvation model was applied to estimate solvation effects on the energies of each structure, using water as the solvent with a dielectric constant (ϵ) of 80.4 [75].

Binding energies were calculated relative to an isolated ion in vacuum and an empty Al-substituted composite building units (CBUs):

$$\Delta E_b = E[\text{M} \bullet \text{CBU}] - E[\text{M}^{n+}] - E[\text{CBU}^{n-}] \quad (8)$$

where $E[\text{M} \bullet \text{CBU}]$ represents the energy of the metal ion positioned within the Al-substituted CBU, $E[\text{M}^{n+}]$ is the energy of the isolated ion, and $E[\text{CBU}^{n-}]$ is the energy of the empty Al-substituted CBU containing n Al atoms. When CBUs offered multiple potential single Al positions or several configurations of two Al atoms, we selected the ΔE_b corresponding to the most stable Al arrangement, as this stability likely reflects the thermodynamically preferred configuration. Given adequate synthesis time, Al atoms tend to occupy stable framework positions, as demonstrated in recent CHA studies [76]. Atoms within CBUs, including terminal hydroxyl groups, were kept fixed during calculations. Energies for empty, Al-substituted CBUs and isolated ions (in vacuum or with implicit solvation) were determined via single-point calculations in ORCA with appropriate charges.

The optimization of the metal-docked ion within the CBU allowed movement of the ion but maintained fixed positions for the CBU atoms, replicating the constraints imposed by a crystal environment where surrounding framework atoms would restrict movement; unconstrained optimization could lead to significant atomic displacements in these molecular forms. This study aims to understand how these ions influence specific CBU structures, given that the OSDA likely directs *lta* cage formation in these zeolites. Thus, CBU atoms were fixed to preserve a shape close to their zeolite-based structures, and binding energies were calculated with and without solvation effects from water.

Zeolite crystal structures for materials synthesized in this work were obtained from the International Zeolite Association (IZA) database [63]. Molecular models of individual CBUs for each synthesized framework were extracted by isolating them from their crystalline counterparts, with frameworks and corresponding CBUs. CBUs were derived from selected zeolite frameworks, as listed in the IZA database. Terminal SiOH groups were added to undercoordinated Si atoms to maintain tetrahedral coordination. These SiOH groups were oriented with Si–O–H bond angles of 180° to minimize hydrogen bonding (H-bonding) interactions between nearby SiOH groups (Fig. A19).

5 Data and code availability

The corresponding code and dataset will be available at: https://github.com/eltonpan/zeosyn_gen

6 Acknowledgments

The authors acknowledge funding from the Generalitat Valenciana through the Prometeo Program (CIPROM/2023/34), CSIC through the I-link+ Program (ILINK24035) and MIT International Science and Technology Initiatives (MISTI) Seed Funds between MIT and CSIC. The authors also acknowledge partial funding from the National Science Foundation DMREF Awards 1922090, 1922311, 1922372, the Office of Naval Research (ONR) under contract N00014-20-1-2280, Kwanjeong Educational Fellowship, ExxonMobil, and the Agency for Science, Technology and Research. The authors thank Di (Daniel) Du, Allen Burton, Xiaochen Du, Soojung Yang and Serena Khoo for helpful discussions and suggestions.

References

- [1] Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
- [2] Merchant, A. *et al.* Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023).
- [3] Sriram, A. *et al.* The open dac 2023 dataset and challenges for sorbent discovery in direct air capture (2024).
- [4] Zhu, Y. *et al.* Uncovering the network modifier for highly disordered amorphous li-garnet glass-ceramics. *Advanced Materials* **36**, 2302438 (2024).
- [5] Kim, K. J., Balaish, M., Wadaguchi, M., Kong, L. & Rupp, J. L. Solid-state li–metal batteries: challenges and horizons of oxide and sulfide solid electrolytes and their interfaces. *Advanced Energy Materials* **11**, 2002689 (2021).
- [6] Zeni, C. *et al.* Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687* (2023).

- [7] Barroso-Luque, L. *et al.* Open materials 2024 (omat24) inorganic materials dataset and models. *arXiv preprint arXiv:2410.12771* (2024).
- [8] Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *Jom* **65**, 1501–1509 (2013).
- [9] He, T. *et al.* Precursor recommendation for inorganic synthesis by machine learning materials similarity from scientific literature. *Science advances* **9**, eadg8180 (2023).
- [10] Huo, H. *et al.* Machine-learning rationalization and prediction of solid-state synthesis conditions. *Chemistry of Materials* **34**, 7323–7336 (2022).
- [11] Karpovich, C., Pan, E., Jensen, Z. & Olivetti, E. Interpretable machine learning enabled inorganic reaction classification and synthesis condition prediction. *Chemistry of Materials* **35**, 1062–1079 (2023).
- [12] Mahbub, R. *et al.* Text mining for processing conditions of solid-state battery electrolytes. *Electrochemistry Communications* **121**, 106860 (2020).
- [13] Bianchini, M. *et al.* The interplay between thermodynamics and kinetics in the solid-state synthesis of layered oxides. *Nature materials* **19**, 1088–1095 (2020).
- [14] Pan, E. *et al.* Zeosyn: A comprehensive zeolite synthesis dataset enabling machine-learning rationalization of hydrothermal parameters. *ACS Central Science* **10**, 729–743 (2024).
- [15] Prein, T. *et al.* Reaction graph networks for inorganic synthesis condition prediction of solid state materials (2024). URL <https://openreview.net/forum?id=VGsXQOTs1E>.
- [16] Schwalbe-Koda, D., Widdowson, D. E., Pham, T. A. & Kurlin, V. A. Inorganic synthesis-structure maps in zeolites with machine learning and crystallographic distances. *Digital Discovery* **2**, 1911–1924 (2023).
- [17] Luo, Y. *et al.* Mof synthesis prediction enabled by automatic data mining and machine learning. *Angewandte Chemie International Edition* **61**, e202200242 (2022).
- [18] Pan, E. *et al.* A chemically-guided generative diffusion model for materials synthesis planning (2024).
- [19] Karpovich, C., Jensen, Z., Venugopal, V. & Olivetti, E. Inorganic synthesis reaction condition prediction with generative machine learning. *arXiv preprint arXiv:2112.09612* (2021).

- [20] Saharia, C. *et al.* Photorealistic text-to-image diffusion models with deep language understanding. *Advances in neural information processing systems* **35**, 36479–36494 (2022).
- [21] Ramesh, A. *et al.* Zero-shot text-to-image generation, 8821–8831 (Pmlr, 2021).
- [22] Ho, J. & Salimans, T. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598* (2022).
- [23] Dhariwal, P. & Nichol, A. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems* **34**, 8780–8794 (2021).
- [24] Zhang, Z., Li, M. & Yu, J. in *On the convergence and mode collapse of gan* 1–4 (2018).
- [25] Moliner, M., Román-Leshkov, Y. & Corma, A. Machine learning applied to zeolite synthesis: the missing link for realizing high-throughput discovery. *Accounts of chemical research* **52**, 2971–2980 (2019).
- [26] Lewis, D., Freeman, C. & Catlow, C. Predicting the templating ability of organic additives for the synthesis of microporous materials. *The Journal of Physical Chemistry* **99**, 11194–11202 (1995).
- [27] Zones, S., Nakagawa, Y., Yuen, L.-T. & Harris, T. Guest/host interactions in high silica zeolite synthesis:[5.2. 1.02. 6] tricyclodecanes as template molecule. *Journal of the American Chemical Society* **118**, 7558–7567 (1996).
- [28] Jensen, Z. *et al.* Discovering relationships between osdas and zeolites through data mining and generative neural networks. *ACS Central Science* **7**, 858–867 (2021).
- [29] Schwalbe-Koda, D. *et al.* A priori control of zeolite phase competition and intergrowth with high-throughput simulations. *Science* eabh3350 (2021).
- [30] Hoffman, A., Xie, M. & Gómez-Bombarelli, R. Learning descriptors to predict organic structure-directing agent applicability in zeolite synthesis (2024).
- [31] Xie, M. *et al.* An exhaustive mapping of zeolite-template chemical space (2024).
- [32] Ronneberger, O., Fischer, P. & Brox, T. *U-net: Convolutional networks for biomedical image segmentation*, 234–241 (Springer, 2015).
- [33] Jospin, L. V., Laga, H., Boussaid, F., Buntine, W. & Bennamoun, M. Hands-on bayesian neural networks—a tutorial for deep learning users. *IEEE Computational Intelligence Magazine* **17**, 29–48 (2022).
- [34] Reynolds, D. A. *et al.* Gaussian mixture models. *Encyclopedia of biometrics* **741** (2009).

- [35] Mirza, M. & Osindero, S. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014).
- [36] Dinh, L., Sohl-Dickstein, J. & Bengio, S. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803* (2016).
- [37] Jia, X. *et al.* Anthropogenic biases in chemical reaction data hinder exploratory inorganic synthesis. *Nature* **573**, 251–255 (2019).
- [38] Cambor, M. A., Villaescusa, L. A. & Diaz-Cabanas, M. Synthesis of all-silica and high-silica molecular sieves in fluoride media. *Topics in Catalysis* **9**, 59–76 (1999).
- [39] Le, T. *et al.* Process regulation of microwave intensified synthesis of y-type zeolite. *Microporous and Mesoporous Materials* **284**, 476–485 (2019).
- [40] Corma, A. *et al.* Proposed pore volume topology of zeolite mcm-22 based on catalytic tests. *Applied Catalysis A: General* **115**, 121–134 (1994).
- [41] Ma, D. *et al.* Mo/hmcm-22 catalysts for methane dehydroaromatization: a multi-inuclear mas nmr study. *The Journal of Physical Chemistry B* **105**, 1786–1793 (2001).
- [42] Zhang, G. *et al.* Single isomerization selectivity of glucose in methanol over sn-bece zeolite of homogenous sn distribution. *Microporous and Mesoporous Materials* **247**, 158–165 (2017).
- [43] Moliner, M. *et al.* Synthesis of the ti- silicate form of bec polymorph of β -zeolite assisted by molecular modeling. *The Journal of Physical Chemistry C* **112**, 19547–19554 (2008).
- [44] Villaescusa, L. *et al.* Pure silica large pore zeolite itq-7: synthetic strategies, structure-directing effects, and control and nature of structural disorder. *Chemistry of Materials* **19**, 1601–1612 (2007).
- [45] Oleksiak, M. D. *et al.* Organic-free synthesis of a highly siliceous faujasite zeolite with spatially biased q4 (nal) si speciation. *Angewandte Chemie* **129**, 13551–13556 (2017).
- [46] Di Iorio, J. R. & Gounder, R. Controlling the isolation and pairing of aluminum in chabazite zeolites using mixtures of organic and inorganic structure-directing agents. *Chemistry of Materials* **28**, 2236–2247 (2016).
- [47] Lee, H. *et al.* Data-driven sda design and high-throughput experimentation for the synthesis of high-silica small-pore zeolites containing lta-cavities (2025).
- [48] Jo, D. *et al.* Synthesis of high-silica lta and ufi zeolites and nh3-scr performance of their copper-exchanged form. *Acs Catalysis* **6**, 2443–2447 (2016).

- [49] Schwalbe-Koda, D., Jensen, Z., Olivetti, E. & Gómez-Bombarelli, R. Graph similarity drives zeolite diffusionless transformations and intergrowth. *Nature materials* **18**, 1177–1181 (2019).
- [50] Cruciani, G. Zeolites upon heating: Factors governing their thermal stability and structural changes. *Journal of Physics and Chemistry of Solids* **67**, 1973–1994 (2006).
- [51] Xie, M. *et al.* A comprehensive mapping of zeolite-template chemical space. *Nature Computational Science* 1–14 (2025).
- [52] Prein, T. *et al.* Retro-rank-in: A ranking-based approach for inorganic materials synthesis planning. *arXiv preprint arXiv:2502.04289* (2025).
- [53] Noh, H., Lee, N., Na, G. S. & Park, C. Retrieval-retro: Retrieval-based inorganic retrosynthesis with expert knowledge. *Advances in Neural Information Processing Systems* **37**, 25375–25400 (2025).
- [54] Lou, A., Meng, C. & Ermon, S. Discrete diffusion language modeling by estimating the ratios of the data distribution (2023).
- [55] Kim, S., Jung, Y. & Schrier, J. Large language models for inorganic synthesis predictions. *Journal of the American Chemical Society* **146**, 19654–19659 (2024).
- [56] Okabe, R. *et al.* Large language model-guided prediction toward quantum materials synthesis. *arXiv preprint arXiv:2410.20976* (2024).
- [57] Prein, T. *et al.* Language models enable data-augmented synthesis planning for inorganic materials. *arXiv preprint arXiv:2506.12557* (2025).
- [58] Song, Z., Lu, S., Ju, M., Zhou, Q. & Wang, J. Accurate prediction of synthesizability and precursors of 3d crystal structures via large language models. *Nature Communications* **16**, 6530 (2025).
- [59] Karpovich, C., Pan, E. & Olivetti, E. A. Deep reinforcement learning for inverse inorganic materials design. *npj Computational Materials* **10**, 287 (2024).
- [60] Pan, E., Karpovich, C. & Olivetti, E. Deep reinforcement learning for inverse inorganic materials design. *arXiv preprint arXiv:2210.11931* (2022).
- [61] Song, J., Meng, C. & Ermon, S. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502* (2020).
- [62] Lipman, Y., Chen, R. T., Ben-Hamu, H., Nickel, M. & Le, M. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747* (2022).
- [63] Baerlocher, C. & McCusker, L. B. Database of zeolite structures. <http://www.iza-structure.org/databases/> (2021).

- [64] Satorras, V. G., Hoogeboom, E. & Welling, M. *E* (n) equivariant graph neural networks, 9323–9332 (PMLR, 2021).
- [65] Halgren, T. A. Mmff vi. mmff94s option for energy minimization studies. *Journal of computational chemistry* **20**, 720–729 (1999).
- [66] Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems* **33**, 6840–6851 (2020).
- [67] Rubner, Y., Tomasi, C. & Guibas, L. J. A metric for distributions with applications to image databases, 59–66 (IEEE, 1998).
- [68] Arjovsky, M., Chintala, S. & Bottou, L. Wasserstein generative adversarial networks, 214–223 (PMLR, 2017).
- [69] Xie, T., Fu, X., Ganea, O.-E., Barzilay, R. & Jaakkola, T. Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197* (2021).
- [70] Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2**, 73–78 (2012).
- [71] Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **8**, e1327 (2018). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1327>.
- [72] Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615–6620 (2008). Publisher: The Royal Society of Chemistry.
- [73] Chai, J.-D. & Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *The Journal of Chemical Physics* **128**, 084106 (2008).
- [74] Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005). Publisher: The Royal Society of Chemistry.
- [75] Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **113**, 6378–6396 (2009). Publisher: American Chemical Society.
- [76] Lee, S. *et al.* Evolution of Framework Al Arrangements in CHA Zeolites during Crystallization in the Presence of Organic and Inorganic Structure-Directing Agents. *Crystal Growth & Design* **22**, 6275–6295 (2022). Publisher: American Chemical Society.

- [77] Sheriff, K., Cao, Y. & Freitas, R. Chemical-motif characterization of short-range order with e (3)-equivariant graph neural networks. *arXiv preprint arXiv:2405.08628* (2024).
- [78] Muraoka, K., Chaikittisilp, W. & Okubo, T. Multi-objective de novo molecular design of organic structure-directing agents for zeolites using nature-inspired ant colony optimization. *Chemical Science* **11**, 8214–8223 (2020).
- [79] Higgins, I. *et al.* *beta-vae: Learning basic visual concepts with a constrained variational framework* (2017).
- [80] Sugiyama, M. *et al.* *Conditional density estimation via least-squares density ratio estimation*, 781–788 (JMLR Workshop and Conference Proceedings, 2010).

Supplementary Information for DiffSyn: A Generative Diffusion Approach to Materials Synthesis Planning

Elton Pan¹, Soonhyoung Kwon², Sulin Liu¹, Mingrou Xie², Alexander J. Hoffman¹, Yifei Duan¹, Thorben Prein³, Killian Sheriff¹, Yuriy Roman-Leshkov², Manuel Moliner⁴, Rafael Gomez-Bombarelli¹, Elsa Olivetti^{1*}

¹ Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

² Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

³ Department of Chemistry, Technische Universität München, München, 80333, Germany.

⁴ Instituto de Tecnología Química, Universitat Politècnica de València-Consejo Superior de Investigaciones Científicas, Valencia, 46022, Spain.

Appendix A ZeoSyn dataset and preprocessing

Dataset The ZeoSyn dataset [14] contains comprehensive synthesis information on zeolites including gel composition, reaction conditions (crystallization time/temperature), precursors, and OSDAs. The dataset also includes the resulting zeolite structures formed (or lack thereof e.g., dense/amorphous phases) for each synthesis route. The dataset consists of 23,961 synthesis routes from 3,096 journal articles spanning the years 1966–2021. It contains data on 921 unique OSDA molecules, 233 zeolite topologies, and 1,022 unique materials. The gel compositions are a combination of 51 different gel components including Si, Al, P, Na⁺, K⁺, OH⁻, F⁻, Ge, Ti, B, Ga, V, OSDA, H₂O, and additional solvents. In the context of zeolite synthesis, the OSDA refers to an organic molecule that act as a ‘template’ that guides the arrangement of inorganic building blocks (e.g. Si, Al) to form the zeolite material.

Dataset split For data splits, a 80/20 train/test split via zeolite-OSDA systems. This is to prevent data leakage across systems and ensure the model generalizes to unseen chemical systems. This results in 1856 and 464 systems in the training and test sets, respectively. Furthermore, the training set is further partitioned via a 87.5/12.5 train/validation random split for hyperparameter tuning. Given most synthesis parameters do not follow the Gaussian distribution (see blue distributions in Fig. 2e), we perform quantile transformation using `sklearn.preprocessing.QuantileTransformer` to map the raw synthesis parameters to a uniform distribution.

Appendix B Additional discussion on generalization and diversity of DiffSyn

Generalization to completely out-of-distribution systems Given the train/test split described in Appendix A, it may be possible that *either* the zeolite framework or

the OSDA may be present in both train and test set. To more rigorously test **DiffSyn**, we evaluate **DiffSyn** on only test systems where *both* zeolite and OSDA are not seen during training. On these strict out-of-distribution test systems that measures the true explorative power of **DiffSyn**, the Wasserstein distance metric (lower is better) is worse (0.45 for out-of-distribution test systems vs. 0.42 for original test systems). This is expected, due to the difficulty of these systems. However, this deterioration in performance is rather small. For reference, even if evaluated on this harder subset of test systems, **DiffSyn** still outperforms all previous methods (Fig. 22a). This shows that **DiffSyn** is robust to out-of-distribution test systems.

Diversity of generated samples We observe that **DiffSyn** may place a higher density of generated samples on major modes, which results in the generated distribution having more spiky peaks compared to ground truth distribution (Fig. 2e). This is because the data on which **DiffSyn** is trained on (ZeoSyn dataset) is mined from the scientific papers, which is prone to anthropogenic factors such as researchers reporting the most common/convenient value (e.g., crystallization temperature of 150°C instead of 153°C). We observe this “spiky peak” phenomenon for synthesis parameters such as crystallization temperature. However, while this phenomenon may decrease the diversity of generated samples, it does not entail mode collapse. For example, **DiffSyn** does not suffer the mode collapse issue unlike conditional GAN, while most accurately capturing the ground truth distribution compared to other generative approaches such as VAEs and normalizing flows (Fig. A4). Moreover, additional case studies in Fig. 3b, 3b, 3d, A14, A13 all show that mode collapse is not a major problem, with **DiffSyn** maintaining output diversity. We do acknowledge that, however, reduced diversity may impede **DiffSyn**’s ability to capture minor modes, which have been discussed in the main text in Sections 2.2.2 and 2.3 for AEL and BEC, respectively.

Appendix C Coverage metric (COV-F1)

We define 2 metrics to compare two distributions of synthesis parameters: Synthesis routes generated by a model $\{\mathbf{S}_k\}_{k \in [1..K]}$ and synthesis routes reported in literature $\{\mathbf{S}_l\}_{l \in [1..L]}$. This method is inspired by [69].

For each synthesis parameter, we calculate the Euclidean distance between generated and literature synthesis parameter, which we call $D_{\text{syn}}(\mathbf{S}_k, \mathbf{S}_l)$. Following the well-known classification metrics of precision and recall, we define the coverage metrics as:

$$\text{COV-R (Recall)} = \frac{1}{L} \left| \{l \in [1..L] : \exists k \in [1..K], D_{\text{syn}}(\mathbf{S}_k, \mathbf{S}_l) < \delta_{\text{syn}}\} \right| \quad (\text{C1})$$

$$\text{COV-P (Precision)} = \frac{1}{K} \left| \{k \in [1..K] : \exists l \in [1..L], D_{\text{syn}}(\mathbf{S}_l, \mathbf{S}_k) < \delta_{\text{syn}}\} \right| \quad (\text{C2})$$

where COV is Coverage. The recall metrics measure how many ground synthesis routes are correctly recalled by the generated synthesis routes, while the precision

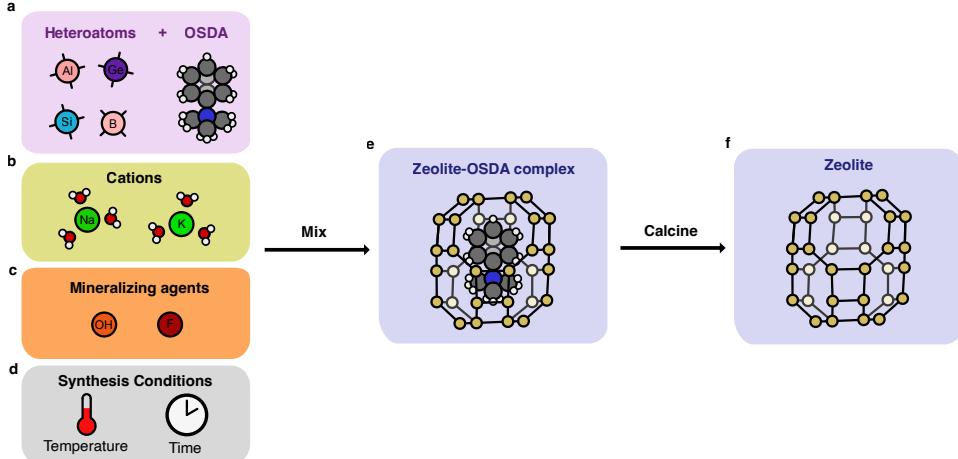


Fig. A1 Complexity of zeolite synthesis. The synthesis of zeolite involves (a) heteroatoms and organic structure-directing agent (OSDA) (b) inorganic cations (c) mineralizing agents, mixed under (d) synthesis conditions to form (e) zeolite-OSDA complex, where the OSDA "templates" the resulting zeolite's pore structure. Calcination then removes the OSDA to yield the (f) zeolite structure.

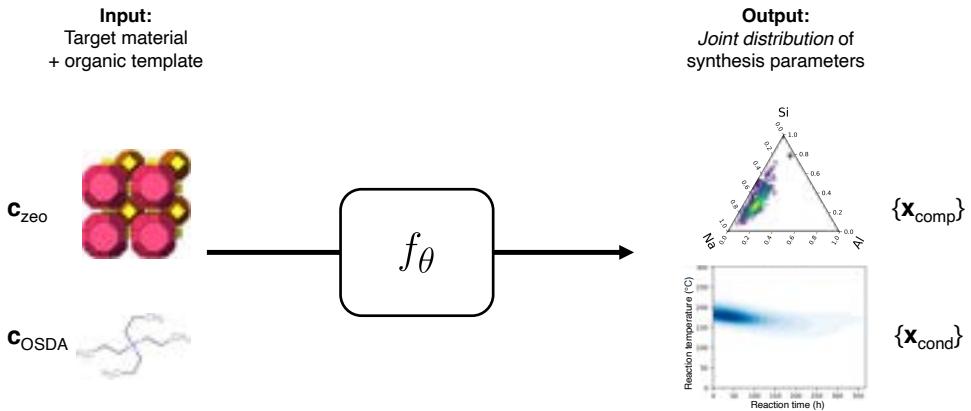


Fig. A2 Materials synthesis prediction task. Given a target zeolite material c_{zeo} and an OSDA c_{OSDA} , the task is to predict an *ensemble* of synthesis routes consisting of gel compositions $\{x_{\text{comp}}\}$ (e.g. Si/Al) and reaction conditions $\{x_{\text{cond}}\}$ (e.g. reaction temperature/time).

metrics measure how many generated synthesis routes are probable (close to literature-reported routes). The threshold δ_{syn} is defined by a domain expert in zeolite synthesis in Table A1. Finally, the overall metric is COV-F1, which is defined as the harmonic mean of COV-R and COV-P i.e. $\text{COV-F1} = \frac{2}{\frac{1}{\text{COV-R}} + \frac{1}{\text{COV-P}}}$, which ranges from 0 – 1 (higher is better). An ideal generative model for materials synthesis planning should maximize its quality of generated synthesis routes (high precision) while maximizing

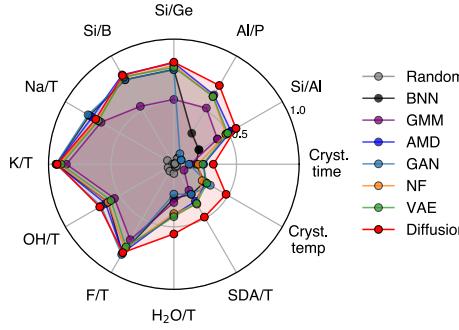
their diversity (high recall). As such, COV-F1 is used as the metric to evaluate the models.

Appendix D Model implementation

Diffusion model training and sampling We train a conditional DDPM with U-Net [32] as the score function, with an input dimension of 12, followed by subsequent layers dimensions multipliers of 128, 64 and 32 for downsampling and the reverse for upsampling. The score function is trained with $T = 1000$, with a $p_{\text{uncond}} = 0.1$ for classifier-free guidance with exponential moving average ($\beta = 0.99$), batch size of 32 at a constant learning rate of 4×10^{-4} for 1M epochs. For sampling, $w = 1.0$ was found to be optimal (see Fig. A6). Models are trained on a single NVIDIA RTX A5000 GPU with a wall time of 51 hours.

Equivariant graph encoder We implement an equivariant graph neural network (EGNN) using the `e3nn` package (inspired by [77]) to encode the structure of the zeolite as a graph. Here, we use a single layer EGNN with $l_{\max} = 2$, $n_{\text{neighbors}} = 10$ and $n_{\text{nodes}} = 140$. In this case, we assume a pure-silica zeolite structure obtained from the IZA database [63], and initialize node embeddings as one-hot according to the element identity (Si or O). We found that the EGNN did not outperform the invariant features (Wasserstein distance of 0.60 for EGNN vs. 0.53 for invariant features). A plausible reason for this result is as follows: The invariant geometric descriptors (e.g., pore volume, cage diameters, framework density) are physically meaningful and capture key aspects of the zeolite structure; these hand-crafted features may already provide a sufficiently informative representation for the synthesis prediction task. In contrast, the EGNN must learn these structural features from scratch. Given the limited size of the dataset (233 distinct zeolite frameworks), the EGNN-learned representation may not generalize better than the invariant descriptors. Considering that the EGNN encoder takes significantly longer time during training and inference, we use invariant features for DiffSyn and all baseline models.

COV-P



COV-R

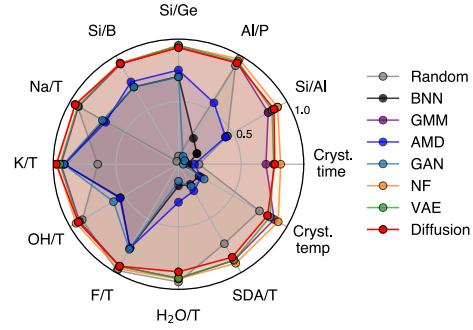


Fig. A3 COV-P (precision) and COV-R (recall) for all models. COV-F1 is the mean of COV-P and COV-R. Notice that high COV-F1 for generative models such as diffusion and VAE (vs. AMD) largely driven by higher COV-R, while high COV-F1 for diffusion (vs. other generative models) largely driven by higher COV-P.

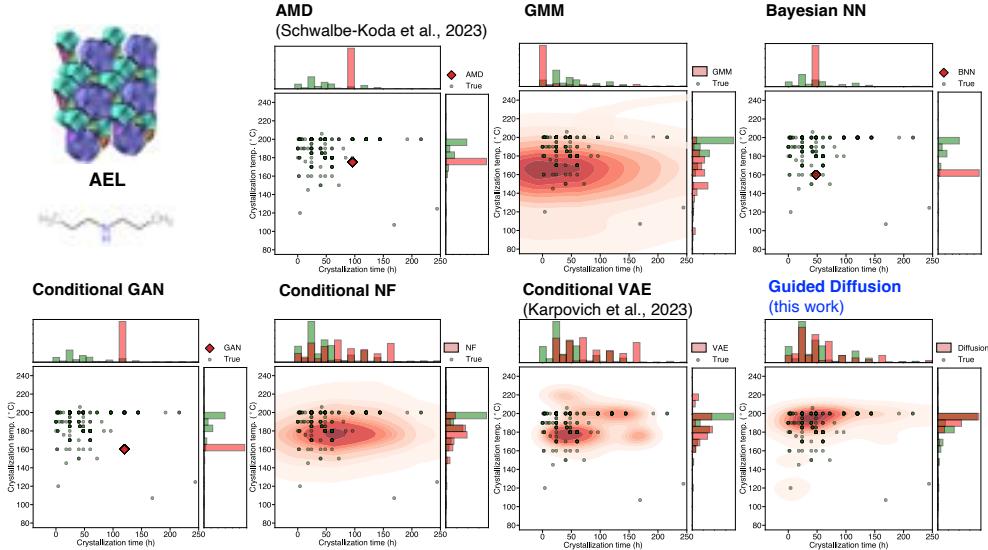


Fig. A4 Predicted $\{x_{\text{cond}}^{\text{gen}}\}$ vs. ground-truth $\{x_{\text{cond}}^{\text{true}}\}$ reaction conditions across different models on an *unseen* zeolite-OSDA system (AEL, dipropylamine). Here, the reaction conditions are crystallization temperature and time. As shown, prior methods based on regression models (AMD) and generative models (Conditional GAN, normalizing flow and VAE) are unable to accurately capture the ground-truth. In contrast, our guided diffusion method (blue, bottom right) is able to generate reaction conditions (red heatmap) that accurately captures the distribution of literature-reported conditions (green points). However, a minority (see 2 points at bottom right) are not captured by the diffusion model. This two points are outliers (very low crystallization temperature and very long crystallization time).

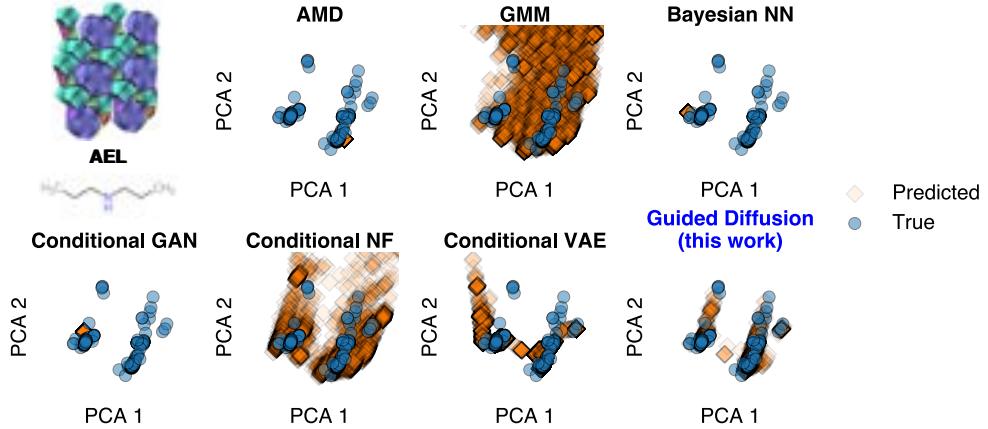


Fig. A5 Principal component analysis of predicted vs. ground-truth synthesis routes across different models on an *unseen* zeolite-OSDA system (AEL, dipropylamine). This plot demonstrates the *multi-modal* nature of synthesis routes for a given structure, where the true distribution (blue) have 2 major modes. Regression-based models (AMD, Bayesian NNs) predict only one of the major modes while missing the other. GAN suffers from mode collapse to one of the modes. NF and VAE are able to predict both modes, but also generated a large number of false positives. In contrast, only guided diffusion (DiffSyn) is able to accurately predict the true distribution of synthesis parameters.

Table A1 Expert-defined thresholds δ_{syn} . Values are based on practical utility and experimental error in zeolite synthesis.

Synthesis parameter	δ_{syn} (unitless unless otherwise specified)
Si/Al	1.5
Al/P	0.05
Si/Ge	0.05
Si/B	0.05
Na^+/T	0.05
K^+/T	0.05
OH^-/T	0.05
F^-/T	0.05
$\text{H}_2\text{O}/\text{T}$	5
SDA/T	0.05
Crystallization temperature	5 °C
Crystallization time	12 h

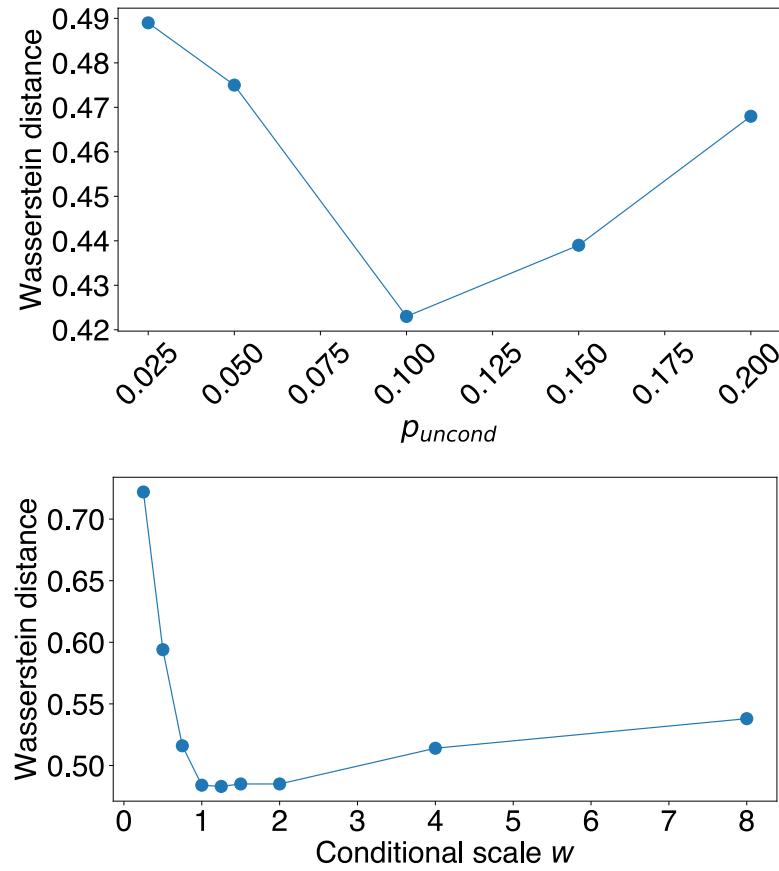


Fig. A6 Influence of classifier-free guidance hyperparameters on model performance in terms of p_{uncond} and guidance strength w .

Table A2 Cost of precursors used in synthesis.
Precursor cost = $\sum_i \eta_i$ (Source: [78])

Reagent	Cost parameter η
Colloidal silica (LUDOX AS-40)	0.036
Aluminum hydroxide	0.052
NaOH	0.0087
NaCl	0.0069
OSDA	190

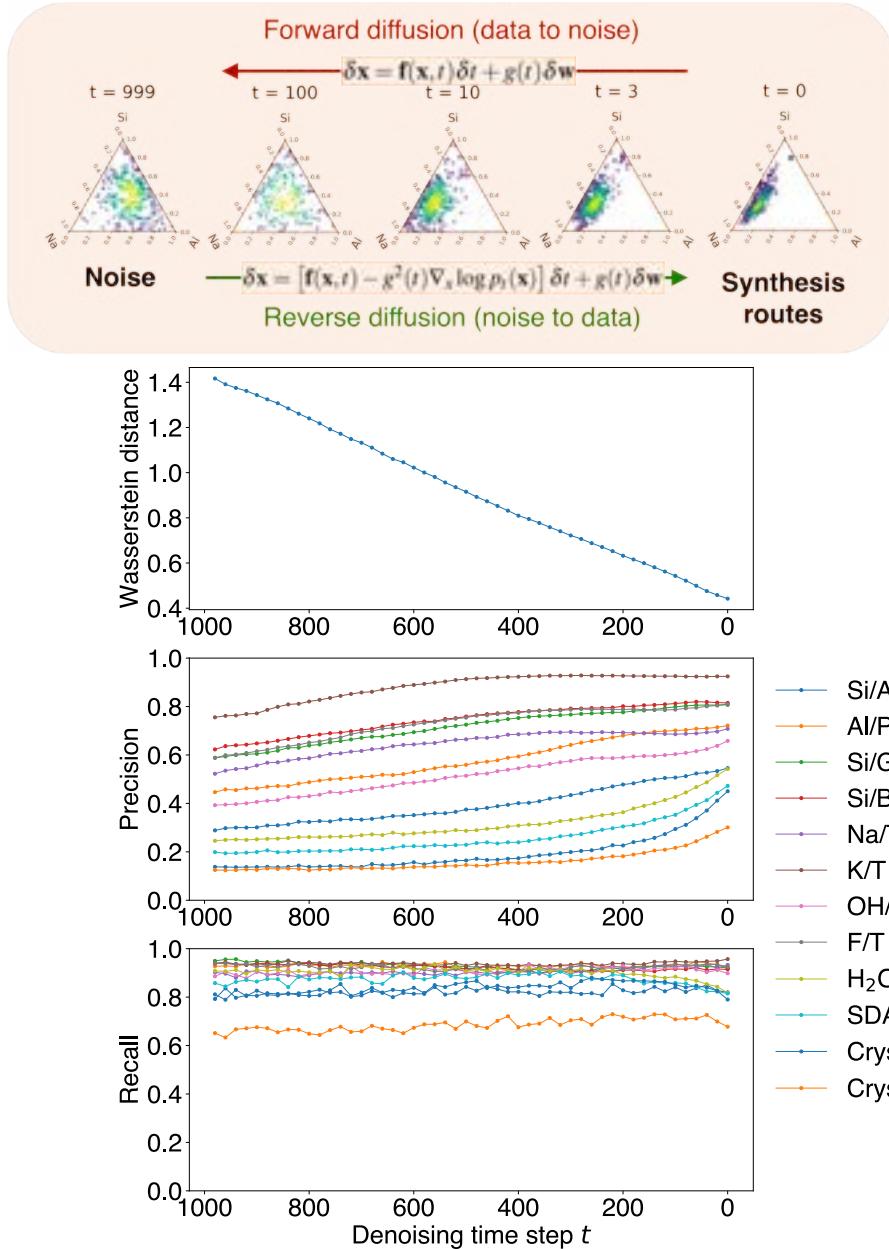


Fig. A7 Model performance across denoising diffusion trajectory measured by Wasserstein distance (↓), precision (↑) and recall (↑). At inference, reverse diffusion (green arrow) maps noise (at $t = T$) to probable synthesis routes (at $t = 0$), leading to a monotonic improvement in Wasserstein distance (between generated and ground truth distributions) and precision (COV-P). In contrast, recall (COV-R) remains relatively constant across the trajectory. This shows that, in reverse time, the diffusion model is able to generate higher quality synthesis recipes while maintaining their diversity, hence capturing the *one-to-many* structure-synthesis relationship for zeolite materials.

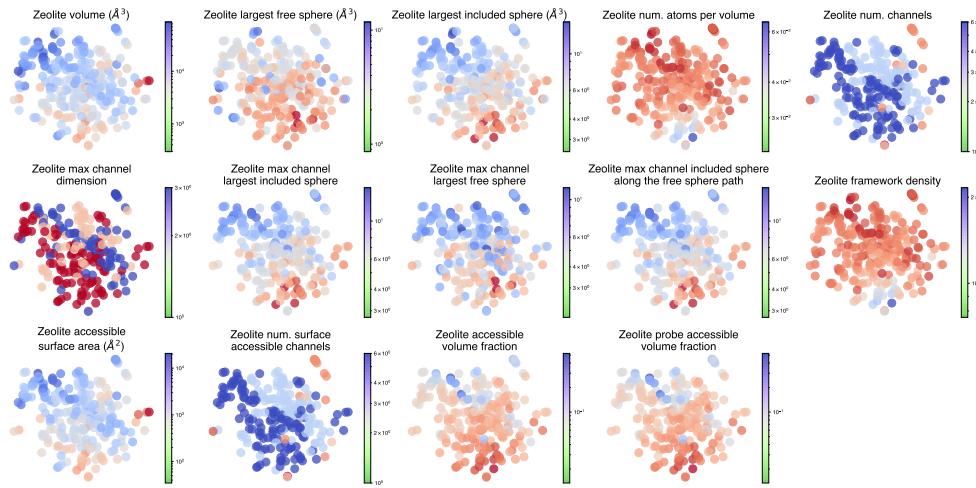


Fig. A8 Learned zeolite representations are chemically meaningful. Latent space is smooth and continuous with respect to zeolite physical properties (e.g., accessible surface area).

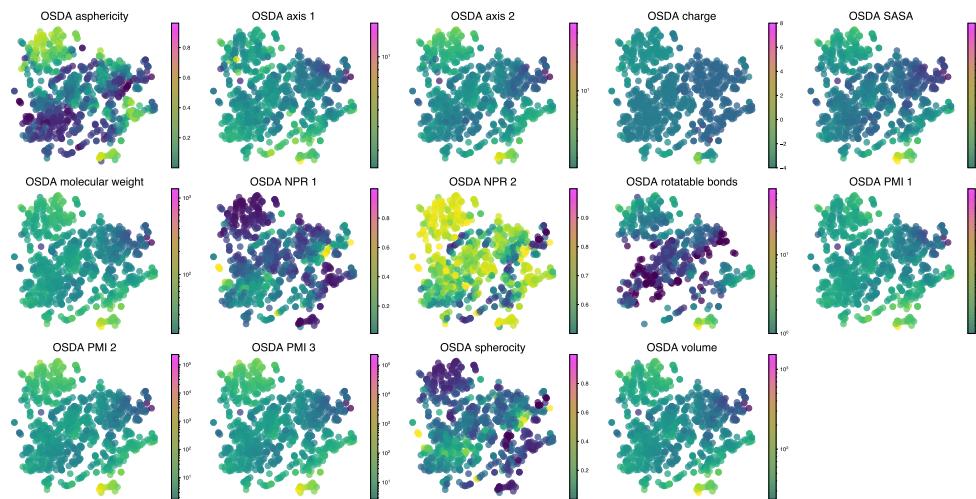


Fig. A9 Learned OSDA representations are chemically meaningful. Latent space is smooth and continuous with respect to OSDA physical properties (e.g., volume).

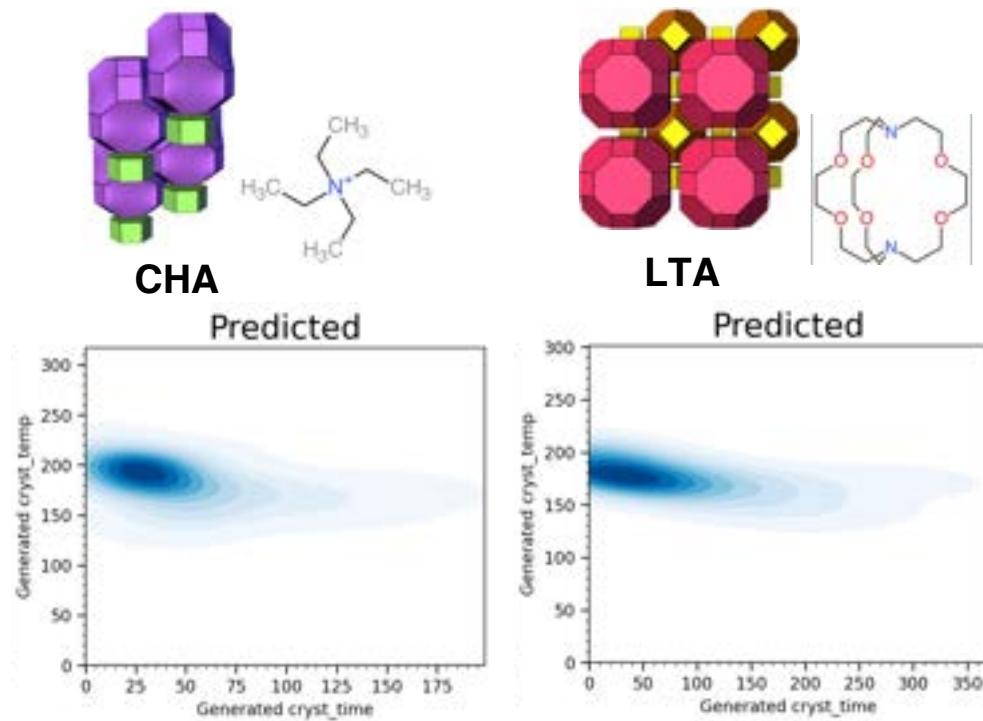


Fig. A10 Model captures physical relationships between synthesis parameters. In the above 2 examples, generated crystallization temperatures and times follow an inverse relationship, which aligns with the Arrhenius equation $k = Ae^{-\frac{E_a}{RT}}$, where rate constant k (determines crystallization time) is inversely related to synthesis temperature. A , E_a , R and T refer to the pre-exponential factor, activation energy and temperature, respectively.

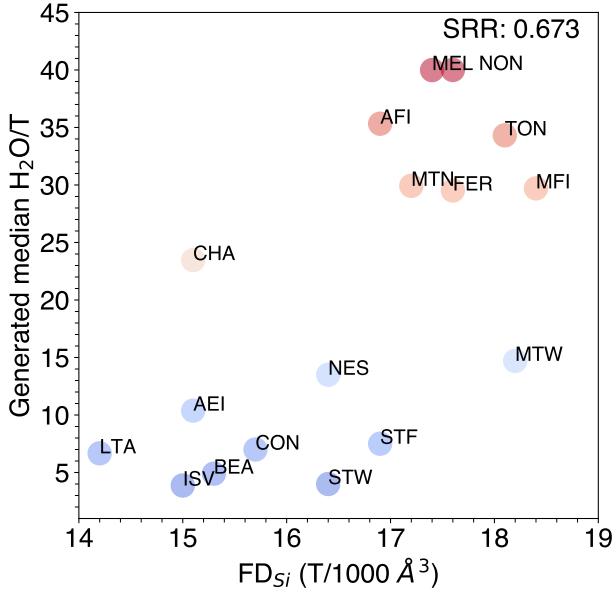


Fig. A11 Model captures domain-specific heuristics. A positive correlation (Spearman's rank coefficient: 0.673) exists between median diffusion-generated H_2O/T and framework density (FD_{Si}) of zeolite structure for fluoride-mediated synthesis of high silica ($Si/Al > 30$) aluminosilicates. This agrees with Villaescusa's rule [38], which states that denser phases (higher FD_{Si}) are favored at the less concentrated conditions (higher H_2O/T), showing that the model has learned domain-specific rules in zeolite synthesis.

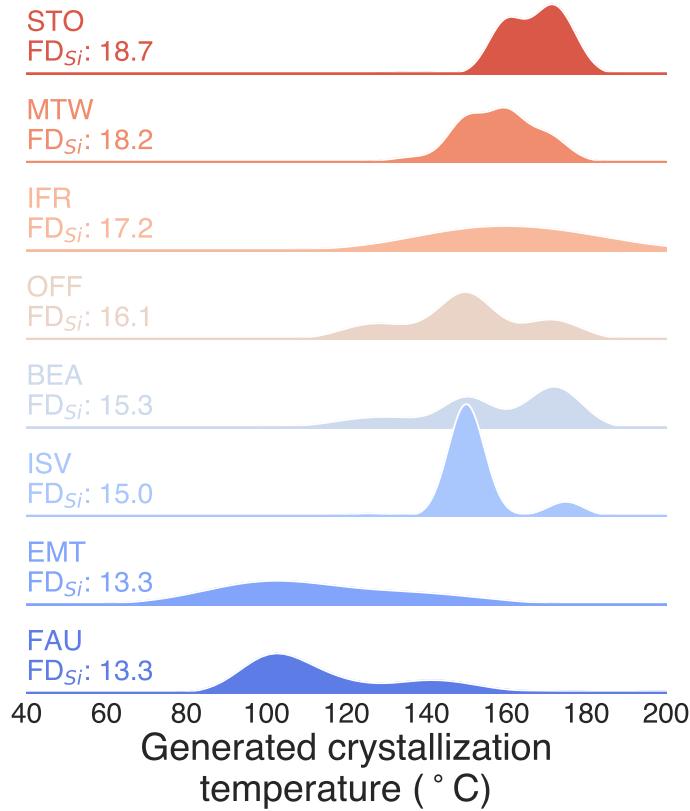


Fig. A12 Model predictions follow thermodynamics of zeolite formation. A positive correlation (Spearman's rank coefficient: 0.931) exists between median diffusion-generated crystallization temperature and framework density (FD_{Si}) of zeolite structure for large-pore aluminosilicates. This can be seen in an rightward shift in distribution of generated temperatures as FD_{Si} increases. This agrees with the thermodynamic argument that higher crystallization temperatures enable the synthesis to overcome the energy activation barrier to form more stable structures with higher framework density [14, 39]. Furthermore, this observation aligns with Ostwald's rule of stages, where the zeolite passes through metastable states before reaching the most thermodynamically favorable framework.

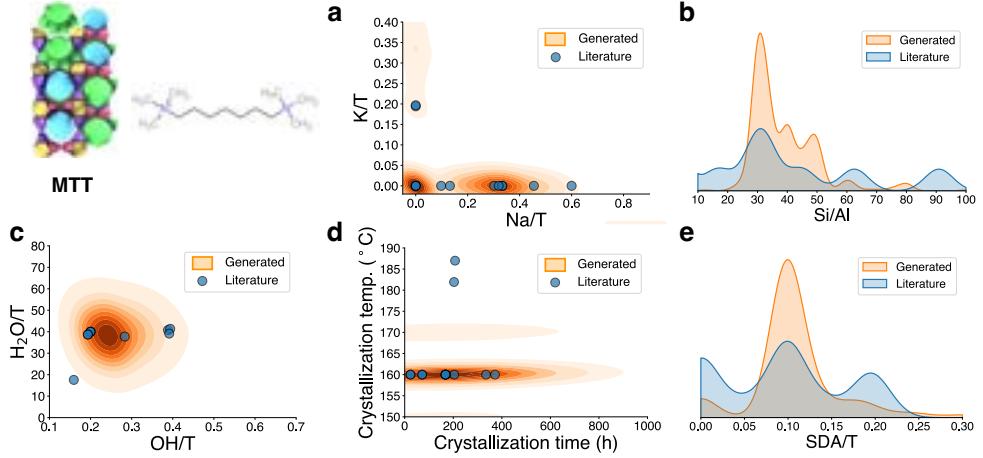


Fig. A13 Case study: MTT synthesis. Generated synthesis parameters for MTT (in cluster 6 of Fig. 3a) overlap closely with unseen synthesis parameters reported in literature. While the model is accurate for most parameters, it is unable to capture very low and high values of Si/Al for MTT (see (b)).

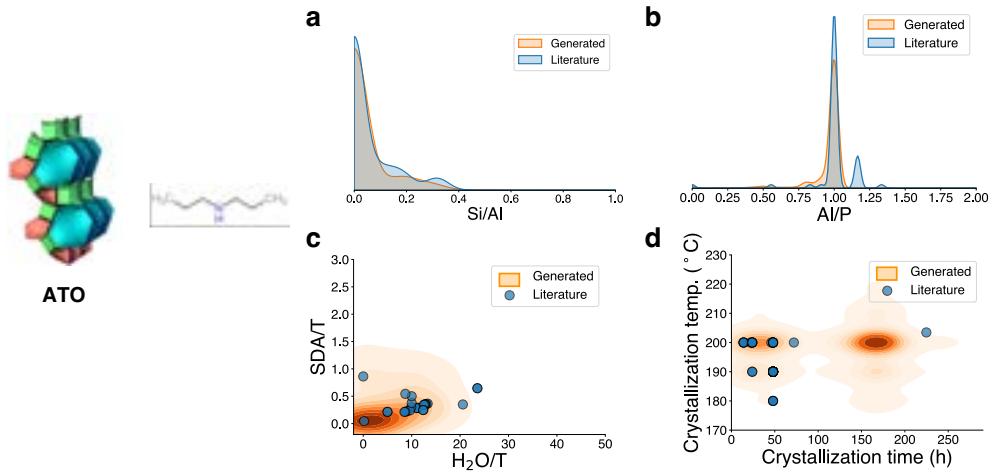


Fig. A14 Case study: ATO synthesis. Generated synthesis parameters for ATO (in cluster 1 of Fig. 3a) overlap closely with unseen synthesis parameters reported in literature. While the model is accurate for most parameters, it predicts an extra peak at crystallization time = 170h and crystallization temperature = 200°C that has weak overlap with only 1 literature datapoint (see (d)). This demonstrates a potential drawback of using literature reported synthesis routes as validation, as they are partial ground-truths i.e. A specific synthesis route that is not being observed in literature is *not a sufficient* condition that the route will fail. Consequently, this will increase the false positive rate, and reduce the COV-R metric (see Fig. A3) of our model. As such, the reported values of the COV-R metric in Fig. A3 are lower bounds.

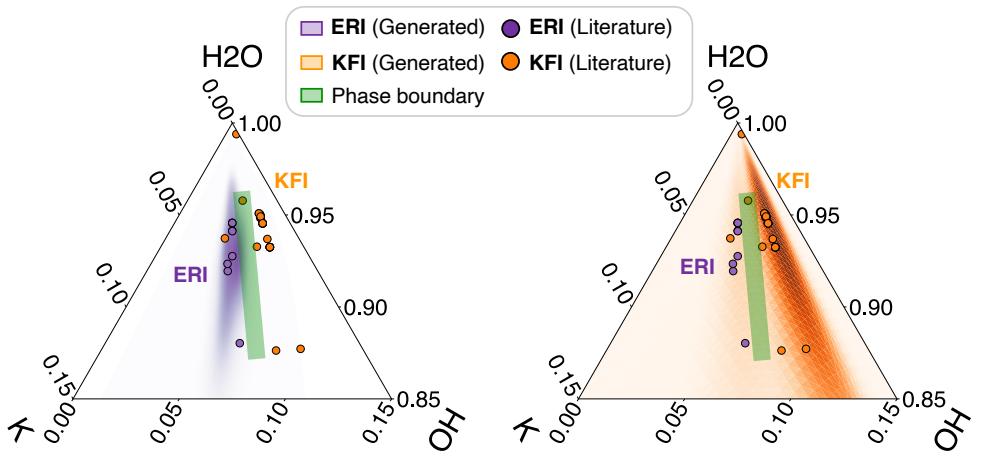


Fig. A15 Competing phases ERI and KFI. Heatmaps refer to generated routes, while points refer to literature-reported synthesis routes. Notice that the model accurately predicts the phase boundary (green shaded region) between ERI and KFI.

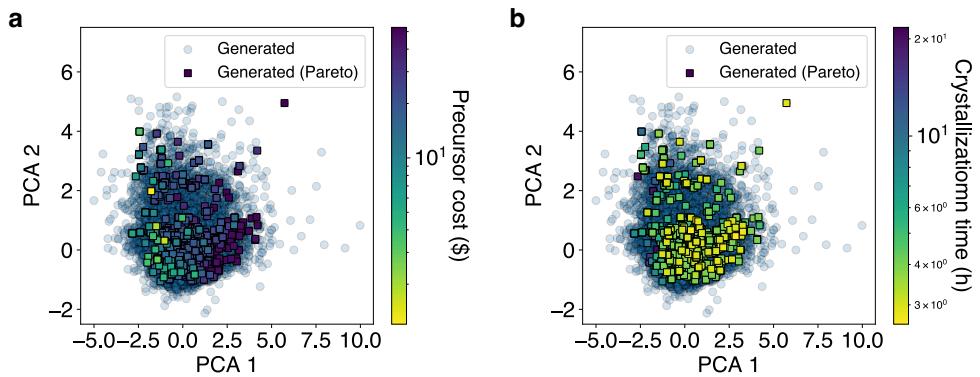


Fig. A16 PCA of synthesis space for CHA-TMAda. Circles refer to all generated synthesis routes. Squares refer to generated synthesis routes at/near the Pareto front in discussed in Fig. 4b and Section 2.4. Each point on the Pareto front is colored by (a) precursor cost (b) crystallization time. Notice that majority of Pareto-optimal synthesis routes reside only in a bottom-half region, and the each of the two objectives (precursor cost and crystallization time) are optimal on *opposite* sides (left for low precursor cost, right for low crystallization time) of the synthesis space.

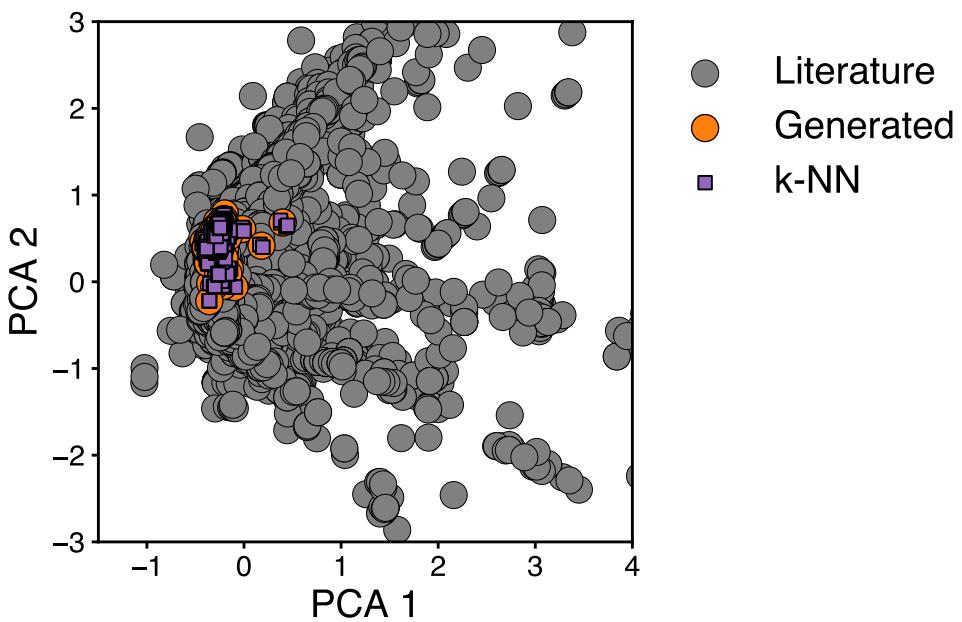


Fig. A17 PCA of synthesis parameters of all routes reported in literature (grey), DiffSyn-generated for UFI (orange) and k -nearest neighbors ($k = 5$) of generated UFI synthesis in literature-reported synthesis.

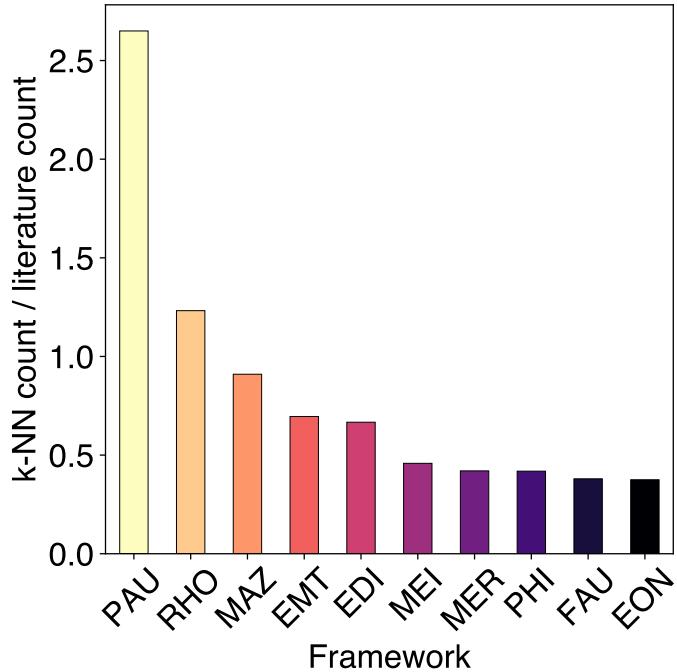


Fig. A18 Normalized frequency of k -NN recipes (purple squares in Fig. A17). Note, it is important to normalize the k -NN count via dividing by the framework's frequency in literature. Larger values (k -NN count / literature count) suggest higher ‘synthesis similarity’ to the DiffSyn-generated UFI recipes. Notably, the top 2 most similar frameworks (PAU and RHO) share a common composite building unit (*lta*) with UFI. Beyond this, we did not observe a clear trend in terms of common CBUs for the remaining frameworks. This may suggest that there is not a clear correlation between structure and synthesis.

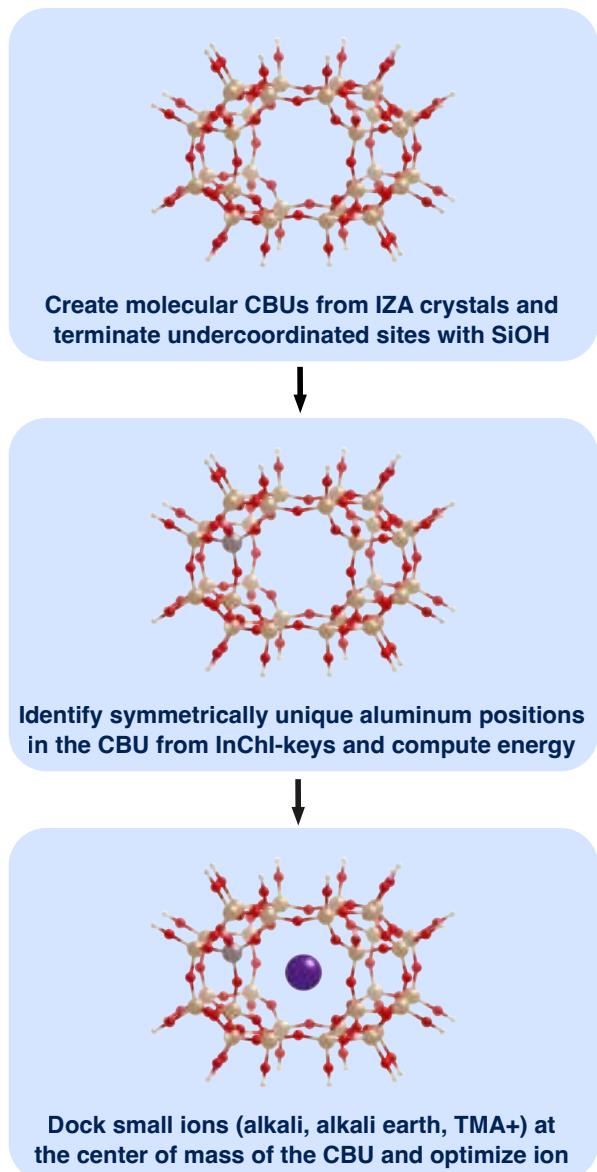


Fig. A19 DFT workflow for docking ions in zeolite composite building units (CBUs). First, CBUs are extracted from zeolite structures (from IZA database [63]) with SiOH capping under-coordinated sites. Second, an exhaustive enumeration of unique aluminum occupancies in the CBU is performed, followed by DFT energy calculations. Third, ions are docked in the CBU to calculate the binding energy between ions and CBU.

Table A3 Physicochemical descriptors of OSDAs. (Source: [14]).

OSDA descriptor	Description
Asphericity	An anisometry descriptor for the deviation from the spherical shape
Axis 1	Two-dimensional (2D) shape descriptors of molecule calculated by projecting the atomic coordinates into a 2D space based on a principal component analysis (PCA) of the positions. The range of the distribution of points in each principal component is reported as the axis of the conformer. Axis 1 is reported as the larger axis, whereas Axis 2 is the smaller axis
Axis 2	See above
Charge	Formal charge of molecule
SASA	Solvent-accessible surface area (SASA) is the surface area of a molecule that is accessible to a solvent
Molecular weight	Molecular mass of molecule
NPR 1	Normalized principal moments ratio ($\frac{I_1}{I_3}$) where I is principal moment of inertia
NPR 2	Normalized principal moments ratio ($\frac{I_2}{I_3}$) where I is principal moment of inertia
Rotatable bonds	Number of rotatable bonds in the molecule. A measure of molecular flexibility.
PMI 1	Principal moments of inertia (PMI) are physical quantities related to the rotational dynamics of a molecule
$I = \sum_{i=1}^A m_i \cdot r_i^2 \quad (D3)$	
	where A is the number of atoms, and m_i is the atomic mass and r_i is the perpendicular distance from the chosen axis of the i th atom of the molecule
PMI 2	See above
PMI 3	See above
Sphericity	Sphericity index of molecule. A measure of how closely the shape of an object resembles that of a perfect sphere
Volume	Molecular volume calculated by using a grid-encoding of the molecular shape using a grid spacing of 0.2 Å and 2.0 Å of margin for the boxes

Table A4 VAE hyperparameters. Here, we train a β -VAE [79].

Hyperparameter	Value
epochs	5000
batch size	2048
learning rate	0.0001
β	0.01
latent dim	2
encoder dims	[4096, 2048, 1024, 512]
decoder dims	[512, 1024, 2048, 4096]

Table A5 NF hyperparameters. Here, we train a normalizing flow with real-valued non-volume preserving (RealNVP) transformations [36].

Hyperparameter	Value
epochs	3000
batch size	2048
learning rate	0.0001
num. flows	32
hidden dim	1024

Table A6 GAN hyperparameters.

Hyperparameter	Value
epochs	400
batch size	2048
learning rate	0.0001
latent dim	2
generator dims	[64, 128, 256]
discriminator dims	[256, 128, 64]

Table A7 BNN hyperparameters.

Hyperparameter	Value
epochs	10000
batch size	8192
learning rate	0.0001
prior μ	0
prior σ	0.1

Table A8 AMD hyperparameters. Here, we train a multi-task MLP with zeolite representation from Schwalbe-Koda et al. [16].

Hyperparameter	Value
epochs	5000
batch size	4096
learning rate	0.0001
hidden dims	[512, 128]

Table A9 GMM hyperparameters. Here, we use a conditional density estimator via least-squares density ratio estimation. [80]. This can be implemented using the `LSConditionalDensityEstimation` function in the `cde` package.

Hyperparameter	Value
center sampling method	k-means
num. centers	5 50
bandwidth	0.5
regularization	1.0

Sample info	Gel composition						Synthesis conditions		Rotation	Seed	Si/Al _{ICP}		
	Si	Al	Ge	B	Na ⁺	K ⁺	H ₂ O	F ⁻	OSDA1 (K222)	OSDA2 (TMA)	Cryst. temp. (°C)*	Cryst. time (h)	
UFI-1	1	0.1	0.0	0.0	0.2	0.0	15	0.0	0.15	0.05	175	168	Yes
UFI-2	1	0.07	0.0	0.0	0.2	0.0	15	0.0	0.15	0.05	175	168	Yes
UFI-3	1	0.1	0.0	0.0	0.3	0.0	15	0.0	0.35	0.05	175	168	Yes
UFI-4	1	0.1	0.0	0.0	0.2	0.0	15	0.0	0.1	0.05	175	168	Yes

Table A10 Synthesis routes of UFI zeolite. K222: Kryptofix 222 (4,7,13,16,21,24-Hexaoxa-1,10-diazabicyclo[8.8.8]hexacosane), TMA: tetramethylammonium. * Minor mode predicted by our model at 175°C