MATDOCK: MULTI-MOLECULE DOCKING IN POROUS MATERIALS WITH FLOW MATCHING

Malte Franke*
Chair of Process Systems Engineering
RWTH Aachen University
malte.franke@rwth-aachen.de

Mingrou Xie*
Department of Chemical Engineering
Massachusetts Institute of Technology
mrx@mit.edu

Akshay Subramanian, Juno Nam, Rafael Gómez-Bombarelli†
Department of Materials Science and Engineering
Massachusetts Institute of Technology
{akshay_s,junonam,rafagb}@mit.edu

ABSTRACT

Molecular docking in materials is important for creating geometries for downstream computations such as structure optimization and transition-state finding. In this work, we present the first use of generative models for multi-molecule docking in periodic materials. MatDock uses flow matching to dock multiple molecules of the same identity in periodic materials. We illustrate its use in docking molecules in porous materials (zeolites) and compare between uniform sampling and Voronoi-based sampling methods. MatDock can be extended beyond just docking to generating energy-optimized docked structures, thus bypassing the key computational bottleneck in creating material-molecule complexes.

1 INTRODUCTION

Heterogenous catalysis and adsorption drives much of industrial reactions and separations to produce essential chemicals for society. To investigate these processes, the determination of an accurate material-molecule structure is critical. Computationally, molecular docking in periodic materials is used to create initial structures that are then optimized structurally or used in transition-state finding to generate accurate structures for analysis. Good initial structures speed up convergence and also converge to more accurate structures. Furthermore, for both molecular adsorption on material surfaces and occlusion within porous material pores, a distribution of possible docked poses can exist for a given crystal-molecule pair (Hoffman et al. (2020)). Automated enumeration of geometries risks both excluding poses overlooked by the docking algorithm and including redundant or irrelevant poses. Computational cost also limits the number of poses for downstream computations. Generative models tackle these problems by drawing multiple, representative samples from the learned distribution during inference simultaneously.

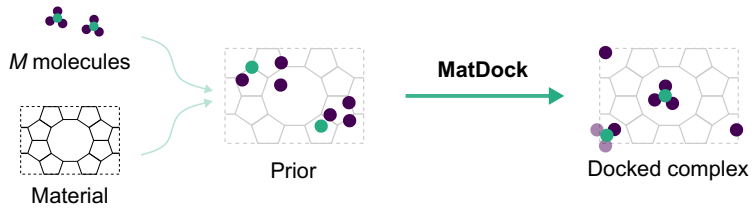

Figure 1: Multi-molecule docking in nanoporous materials with MatDock. The dotted lines indicate the periodic boundaries of the material.

*Equal contribution.
†Corresponding author.

While generative models for docking are numerous in the protein-ligand binding literature, to our knowledge no work exists for generative molecular docking in periodic materials, much less multi-ligand docking. In this work, we showcase MatDock, which uses flow matching (FM) Lipman et al. (2023) for blind docking of multiple molecules *simultaneously* within periodic materials. We illustrate its use in multi-molecule docking in porous materials called zeolites and investigate the effect of different sampling strategies. Formally, using fractional coordinates $\boldsymbol{f_{mols}} = [\boldsymbol{f}^1, \ldots, \boldsymbol{f^M}] = [f^1, \ldots, f^{MN}] \in \mathcal{F} = [0,1)^{MN\times 3}$ with a loading of M docked molecules with N heavy atoms each, lattice parameters $\boldsymbol{l}$, molecular graph $\boldsymbol{G_{mol}}$ (constructed using SMILES strings), crystal elements $\boldsymbol{a_{crys}}$ and crystal fractional coordinates $\boldsymbol{f_{crys}}$, we fit a generative model $p(\boldsymbol{f_{mols}}|\boldsymbol{l}, \boldsymbol{G_{mol}}, \boldsymbol{a_{crys}}, \boldsymbol{f_{crys}})$. For the rest of this paper, assume $\boldsymbol{f} \equiv \boldsymbol{f_{mols}}$ for brevity.

1.1 Related work

Traditional high-throughput methods for multi-molecule docking in zeolites *sequentially* place molecule conformers in the unit cell, check for validity and repeat this process until a stopping criterion is reached. Schwalbe-Koda & Gómez-Bombarelli (2021b;a) integrate prior knowledge about the pore space to speed up the sampling of valid poses, and we use their method to generate the datasets used in this work.

While many data-driven molecular docking models exist in the protein-ligand docking space, most of them focus on single molecule docking Corso et al. (2023); Lu et al. (2022); Cao et al. (2024) with the notable exception of FlowSite (Stärk et al. (2024)). FlowSite docks multiple ligands of differing identities in a single, predefined protein pocket. However, all protein-ligand docking models, FlowSite included, do not account for periodic boundary conditions. Furthermore, multiple docking sites can exist in a periodic crystal. On the other hand, recent generative models for crystals such as DiffCSP(Jiao et al. (2024)), MatterGen (Zeni et al. (2024)), and FlowMM (Miller et al. (2024)) successfully model periodicity by leveraging the topology of the torus, but have not been used for molecular docking which breaks the symmetries of the crystalline structure.

We combine ideas from both generative models for materials and multi-ligand docking in proteins to build a generative model for docking in periodic structures. MatDock adapts the Riemannian FM model FlowMM Miller et al. (2024) for blind docking of multiple ligands with the same molecular graph $\boldsymbol{G_{mol}}$ *simultaneously* in a periodic crystal with multiple available binding sites.

2 Methods

2.1 Riemannian flow matching

Materials modeling relies on incorporating periodic boundary conditions: particles on one side of the bounding cell are neighbors of particles on the opposite side. We leverage the structure of the torus - a Riemannian manifold - to naturally model periodicity. Riemannian manifolds $\mathcal{M}$ are defined by local coordinate systems that continuously transform into each other, and their Riemannian metric g (choice of inner product). Geometric variables on Riemannian manifolds like distances, volumes, and minimum length curves (*geodesics*) can then be defined via the inner product $\langle u, v\rangle$ for $u, v \in \mathcal{T}_p\mathcal{M}$, where $\mathcal{T}_p\mathcal{M}$ is the tangent space at point $p \in \mathcal{M}$. In this work, we consider $\mathcal{F}$, a set of $MN \times 3$ flat tori equipped with the (local) Euclidean inner product.

Let $i = i(m,n)$ be the global index in $\boldsymbol{f}$ for atom n in molecule m. Given samples from the true distribution $\boldsymbol{f_1} \sim q(\boldsymbol{f}; \boldsymbol{l}, \boldsymbol{G_{mol}}, \boldsymbol{a_{crys}}, \boldsymbol{f_{crys}})$ we aim to define a conditional vector field (VF) $u_t(\cdot|\boldsymbol{f_1})$ that can cross the periodic boundary for FM (Chen & Lipman (2024)). Using the atom-wise wrapping function $w(f^i) \coloneqq f^i - \lfloor f^i \rfloor \in [0,1)^3$ which returns only the fractional part of its argument, the exponential and logarithmic maps on tori are defined as

$$\exp_{f^i}(\dot{f}^i) \coloneqq w(f^i + \dot{f}^i), \tag{1}$$

$$\log_{f_0^i}(f_1^i) \coloneqq \frac{1}{2\pi}\text{atan2}\left[\sin(2\pi(f_1^i - f_0^i)), \cos(2\pi(f_1^i - f_0^i))\right] \tag{2}$$

with $\dot{f}^i \in \mathcal{T}_{f^i}\mathcal{F}^i$ for $i = 1, \ldots, MN$. Intuitively, the logarithmic map $\log_{f_0^i}(f_1^i)$ is the vector on the tangent space $\mathcal{T}_{f_0^i}\mathcal{F}^i$ at f_0^i pointing towards f_1^i, whereas applying the exponential map results in

walks on the torus respecting periodicity. We then define our conditional VF as

$$u_t(\boldsymbol{f}|\boldsymbol{f}_1) = \frac{\log_{\boldsymbol{f}}(\boldsymbol{f}_1)}{1-t}. \tag{3}$$

which can point across the periodic boundary and is translation equivariant. Translating the molecule positions $\boldsymbol{f}$ (relative to $\boldsymbol{f}_{crys}$) hence results in a different VF - a desirable property for docking. The VF should handle atoms close to pore spaces differently than atoms close to material atoms. A conceptualization of this property and the difference to FlowMM can be found in C. The flow $\psi_t(\boldsymbol{f}|\boldsymbol{f_1})$ is generated by u_t and pushes $\boldsymbol{f}$ towards $\boldsymbol{f_1}$. During training, we will use ψ_t to obtain $\boldsymbol{f_t} = \psi_t(\boldsymbol{f_0}|\boldsymbol{f_1})$, which for simple geometries like tori have the closed form

$$f_t^i \coloneqq \exp_{f_0^i}(t \log_{f_0^i}(f_1^i)). \tag{4}$$

Finally, we train a graph neural network v_t^θ as described in B minimizing the following loss

$$\mathcal{L}_{RFM}(\theta) = \mathbb{E}_{t,q(\boldsymbol{f}_1),p(\boldsymbol{f_0})} \left\| v_t^\theta(\boldsymbol{f_t}) - \frac{\log_{\boldsymbol{f_t}}(\boldsymbol{f}_1)}{1-t} \right\|^2. \tag{5}$$

2.2 Base distributions

We choose to independently draw one uniformly distributed reference point $\mu_m \sim \text{Uniform}(0,1) \in [0,1)^3$ for each molecule M. Afterwards, we sample N corresponding atoms from a Gaussian centered at μ_m, and wrap everything back to the unit cell (henceforth called "Uniform-Gaussian" or "U-G"). Then the base density is the product of molecule densities $p(\boldsymbol{f}) = \prod_{m=1}^{M} p(\boldsymbol{f}^m)$, where $p(\boldsymbol{f}^m) \coloneqq \prod_{n=1}^{N} w\left(\mathcal{N}\left(f^i|\mu_m, \Sigma\right)\right)$. We set $\tilde{\sigma}^2 = (3\,\text{Å})^2 I \in \mathbb{R}^{3\times3}$ in Cartesian coordinates and transform back to fractional coordinates with the lattice matrix $\tilde{\boldsymbol{l}} \in \mathbb{R}^{3\times3}$, i.e. $\Sigma = \tilde{\boldsymbol{l}}^{-1}\tilde{\sigma}^2(\tilde{\boldsymbol{l}}^{-1})^\top$.

Similar to Schwalbe-Koda & Gómez-Bombarelli (2021b), we further leverage prior information about the topology of the material by computing the set of Voronoi nodes $V = \{v_1, \ldots, v_K\}$. Intuitively, Voronoi nodes v_k lie within zeolite pores and could serve as an informative prior for the task of docking molecules within the same void spaces. They are computed using Voronoi tessellation based on $\boldsymbol{f_{crys}}$, i.e. $v_k = f(\boldsymbol{f_{crys}}) \in [0,1)^3$. We can express distributions using Voronoi nodes like Uniform-Gaussian distributions by simply exchanging μ_m with $\tilde{\mu_m} \sim \text{Uniform}(V)$ (henceforth called "Voronoi-Gaussian" or "V-G").

2.3 Evaluation metrics

In protein-ligand docking, performance is typically evaluated by the top-5 % RMSD < 2 Å, where RMSD is the root mean squared deviation of atomic positions. Given 5 samples generated at inference for a particular data point, generation is considered successful if at least one sample has a RMSD < 2 Å. Performance is then evaluated by the percentage of successful generations. However, zeolite-molecule complexes exhibit both pore symmetry (docking one molecule within either of two equivalent pores constitutes a valid pose) and molecule permutation symmetry (swapping the order of two docked molecules remains valid since the molecules have the same identity). Thus, our focus is on 1) reasonable reconstruction of the molecule and physical clashes 2) between molecules or 3) between molecule(s) and zeolite.

We assess molecule reconstruction by calculating the RMSD between each generated molecule and each target conformer after rotational and translational alignment with the Kabsch algorithm, and take the lowest RMSD for each molecule. The average of RMSDs in a multi-molecule pose is taken as the final reconstruction error. As MatDock generates only the heavy atoms of the molecules, a pose is deemed to have a molecule-molecule clash if any atoms from the first molecule are closer than 2.25 Å (roughly twice the C-H bond length) to the second molecule, and to have a zeolite-molecule clash if any molecule atoms are closer than 2.0 Å (approximately C-H bond length + O-H bond length) to the zeolite. We note that the distance cutoffs are chosen conservatively and that performance could vary with the cutoffs.

2.4 Data

We train MatDock on 5,948 datapoints including zeolite-molecule pairs of different loadings, generated with the VOID library (Schwalbe-Koda & Gómez-Bombarelli (2021b)) (details in Appendix

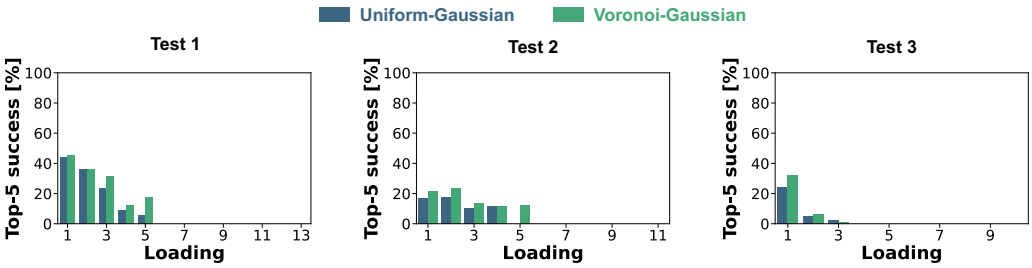

Figure 2: Percentage of successfully docked poses across top 5 samples for different molecular loadings/ zeolite unit cell, for held out datasets Tests 1 - 3 (left to right), comparing between Uniform-Gaussian and Voronoi-Gaussian sampling.

A). Performance is measured on three held out datasets of increasing difficulty: Test 1 - seen zeolites and unseen molecules, Test 2 - unseen zeolites and seen molecules, and Test 3 - unseen zeolites and unseen molecules.

3 RESULTS

	Test 1		Test 2		Test 3	
Evaluation	**U-G**	**V-G**	**U-G**	**V-G**	**U-G**	**V-G**
% No mol-mol clash (↑)	72.95 ± 1.2	$\mathbf{76.26 \pm 1.3}$	$\mathbf{83.63 \pm 2.0}$	82.89 ± 2.6	$\mathbf{79.25 \pm 0.5}$	78.06 ± 1.7
% No zeolite-mol clash (↑)	37.56 ± 1.4	$\mathbf{39.17 \pm 1.2}$	33.78 ± 3.6	$\mathbf{38.15 \pm 2.4}$	13.28 ± 1.4	$\mathbf{19.69 \pm 0.9}$
% RMSD < 2 Å(↑)	$\mathbf{63.86 \pm 1.3}$	62.64 ± 1.4	$\mathbf{42.22 \pm 2.1}$	40.67 ± 0.7	49.31 ± 1.0	$\mathbf{50.78 \pm 1.0}$
% Success (↑)	21.22 ± 1.1	$\mathbf{21.77 \pm 1.4}$	10.74 ± 1.7	$\mathbf{14.15 \pm 1.8}$	7.75 ± 0.9	$\mathbf{10.31 \pm 1.1}$
% Top-5 success (↑)	28.15	**30.51**	14.41	**19.26**	10.31	**13.44**

Test 1: seen zeolites, unseen mols; **Test 2**: unseen zeolites, seen mols; **Test 3**: unseen zeolites, unseen mols
U-G: Uniform-Gaussian, **V-G**: Voronoi-Gaussian.

Table 1: Performance metrics for Uniform-Gaussian sampling and Voronoi-Gaussian sampling. "Success" is defined as fulfillment of all three criteria from Section 2.3: reasonable reconstruction of molecule and no molecule-molecule or zeolite-molecule clashes.

Across performance metrics listed in Table 1, V-G sampling generally outperforms U-G sampling. Especially for unseen zeolites in held out datasets Tests 2 and 3, V-G sampling provides helpful priors about molecule-occupiable void spaces in the zeolite and reduces the percentage of zeolite-molecule clashes, with an improvement of 4 - 6% compared to uniform sampling. However, molecule-molecule clashes are more frequent with V-G sampling; we hypothesize that this might be due to the possibility of sampling Voronoi nodes too close to each other.

When broken down by loading, V-G sampling generally outperforms U-G sampling across all loadings, with the difference increasing as the test sets become more out-of-distribution (Fig. 2). We note that the cutoffs for determining a successful docked pose can affect the relative performance of the two sampling methods, but the improved performance of V-G sampling in the more difficult Test 3 dataset can be seen with more lenient cutoffs as well (Appendix D).

4 DISCUSSION

Our current method does not penalize overly high loadings as the number of molecules is provided as user input. However, it is straightforward to compute the maximum loading from the molecular volume and the molecule-occupiable pore volume (Willems et al. (2012)). Studies for optimizing the loading will be covered in future work.

We also observed that atoms sampled far away from zeolite pores sometimes struggle to reach the pores. This could be a limitation of the current architecture, which captures strictly local information

through edge creation with only neighboring atoms within a radius cutoff. However, fully connected networks are prohibitively expensive due to the hundreds of atoms in the zeolite-molecule complex. Encoding of long-range interactions remains an active area of research.

MatDock can be extended to generation of optimized structures from the same inputs with the appropriate datasets. We envision the use of MatDock to produce an ensemble of good initial guesses for transition states, intermediates adsorbed on or in other materials, and other systems of interest that can be further refined with simulations.

ACKNOWLEDGMENTS

M.F. acknowledges funding support from the Prof. Dr.-Ing. Erich Müller foundation, the Alexa Melcher and Thomas Schiffer foundation and the RWTH Aachen research ambassador scholarship. M.X. acknowledges funding support from the Agency for Science, Technology and Research (A*STAR) NSS-PhD scholarship. Computations were performed on the RWTH HPC under project rwth1232 and thes1904 and on MIT Supercloud.

REFERENCES

Duanhua Cao, Mingan Chen, Runze Zhang, Zhaokun Wang, Manlin Huang, Jie Yu, Xinyu Jiang, Zhehuan Fan, Wei Zhang, Hao Zhou, Xutong Li, Zunyun Fu, Sulin Zhang, and Mingyue Zheng. SurfDock is a surface-informed diffusion generative model for reliable and accurate protein–ligand complex prediction. *Nature Methods*, November 2024. ISSN 1548-7091, 1548-7105. doi: 10.1038/s41592-024-02516-y. URL https://www.nature.com/articles/s41592-024-02516-y.

Ricky T. Q. Chen and Yaron Lipman. Flow Matching on General Geometries, February 2024. URL http://arxiv.org/abs/2302.03660. arXiv:2302.03660 [cs, stat].

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking, February 2023. URL http://arxiv.org/abs/2210.01776. arXiv:2210.01776 [physics, q-bio].

Alexander J. Hoffman, Jason S. Bates, John R. Di Iorio, Steven V. Nystrom, Claire T. Nimlos, Rajamani Gounder, and David Hibbitts. Rigid Arrangements of Ionic Charge in Zeolite Frameworks Conferred by Specific Aluminum Distributions Preferentially Stabilize Alkanol Dehydration Transition States. *Angewandte Chemie*, 132(42):18845–18853, October 2020. ISSN 0044-8249, 1521-3757. doi: 10.1002/ange.202007790. URL https://onlinelibrary.wiley.com/doi/10.1002/ange.202007790.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal Structure Prediction by Joint Equivariant Diffusion, March 2024. URL http://arxiv.org/abs/2309.04475. arXiv:2309.04475 [cond-mat].

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow Matching for Generative Modeling, February 2023. URL http://arxiv.org/abs/2210.02747. arXiv:2210.02747 [cs, stat].

Wei Lu, Qifeng Wu, Jixian Zhang, Jiahua Rao, Chengtao Li, and Shuangjia Zheng. TANKBind: Trigonometry-Aware Neural NetworKs for Drug-Protein Binding Structure Prediction. preprint, Biophysics, June 2022. URL http://biorxiv.org/lookup/doi/10.1101/2022.06.06.495043.

Benjamin Kurt Miller, Ricky T. Q. Chen, Anuroop Sriram, and Brandon M. Wood. FlowMM: Generating Materials with Riemannian Flow Matching, June 2024. URL http://arxiv.org/abs/2406.04713. arXiv:2406.04713 [cond-mat, physics:physics, stat].

Daniel Schwalbe-Koda and Rafael Gómez-Bombarelli. Benchmarking binding energy calculations for organic structure-directing agents in pure-silica zeolites. *The Journal of Chemical Physics*, 154(17):174109, May 2021a. ISSN 0021-9606, 1089-7690. doi: 10.1063/5.0044927. URL https://pubs.aip.org/jcp/article/154/17/174109/200069/Benchmarking-binding-energy-calculations-for.

Daniel Schwalbe-Koda and Rafael Gómez-Bombarelli. Supramolecular Recognition in Crystalline Nanocavities through Monte Carlo and Voronoi Network Algorithms. *The Journal of Physical Chemistry C*, 125(5):3009–3017, February 2021b. ISSN 1932-7447, 1932-7455. doi: 10.1021/acs.jpcc.0c10108. URL https://pubs.acs.org/doi/10.1021/acs.jpcc.0c10108.

Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. Harmonic Self-Conditioned Flow Matching for Multi-Ligand Docking and Binding Site Design, March 2024. URL http://arxiv.org/abs/2310.05764. arXiv:2310.05764 [cs].

Thomas F. Willems, Chris H. Rycroft, Michaeel Kazi, Juan C. Meza, and Maciej Haranczyk. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous and Mesoporous Materials*, 149(1):134–141, February 2012. ISSN 13871811. doi: 10.1016/j.micromeso.2011.08.020. URL https://linkinghub.elsevier.com/retrieve/pii/S1387181111003738.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, Bichlien Nguyen, Hannes Schulz, Sarah Lewis, Chin-Wei Huang, Ziheng Lu, Yichi Zhou, Han Yang, Hongxia Hao, Jielan Li, Ryota Tomioka, and Tian Xie. MatterGen: a generative model for inorganic materials design, January 2024. URL http://arxiv.org/abs/2312.03687. arXiv:2312.03687 [cond-mat].

A DATA

For each zeolite-molecule pair, molecules are docked using either the Voronoi or the Monte Carlo algorithm with default settings through the VOID library (Schwalbe-Koda & Gómez-Bombarelli (2021b)). For each docking, molecules are added to the unit cell of the zeolite sequentially until no further molecules can be added without physical clashes. Two poses for each of the top 5 loadings are saved to create the dataset. Hence, each zeolite-molecule pair can have up to 20 distinct poses. We curated the set of zeolites and molecules from experimentally synthesized zeolites and molecules that are known to act as structure-directing templates for synthesizing zeolites. We also selected frameworks with the following characteristics to cut down on computational cost due to large system sizes but ensure sufficiently large zeolite pores are represented: 1) unit cell only; 2) number of atoms less than 200; 3) unit cell volume less than 6000 Å^3; 4) largest included sphere diameter greater than 4 Å; and 5) at least 1-D pore dimensionality.

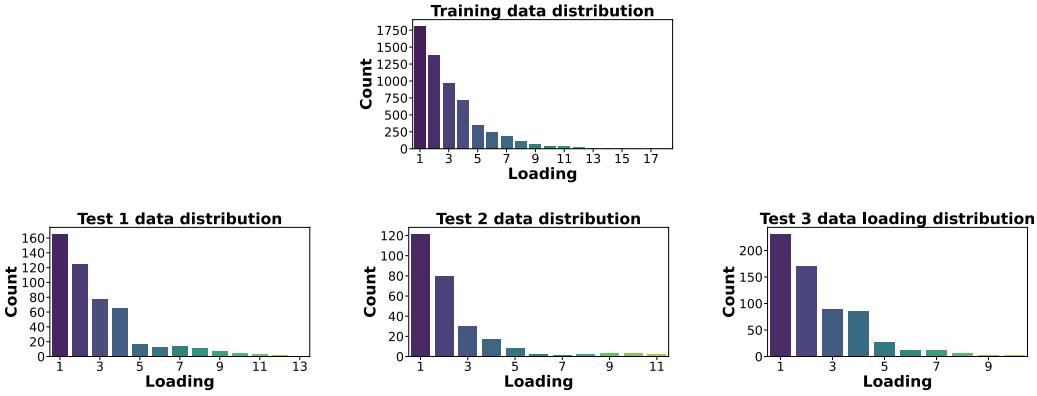

Figure 3: Distributions of loading (number of molecules per zeolite) for each dataset.

B MODEL ARCHITECTURE

Similar to FlowMM, we use an adapted version of EGNN by Jiao et al. (2024) to parametrize the vector field v_t^θ:

$$\mathbf{h}^i_{(0)} = \phi_{\mathbf{h}_{(0)}}(a^i), \tag{6}$$

$$\mathbf{m}^{ij}_{(s)} = \varphi_m\left(\mathbf{h}^i_{(s-1)}, \mathbf{h}^j_{(s-1)}, \boldsymbol{l}, \mathbf{e}^{i,j}, \text{SinusoidalEmbedding}\left(f^j - f^i\right)\right), \tag{7}$$

$$\mathbf{m}^i_{(s)} = \sum_{j=1}^{N} \mathbf{m}^{ij}_{(s)}, \tag{8}$$

$$\mathbf{h}^i_{(s)} = \mathbf{h}^i_{(s-1)} + \varphi_h\left(\mathbf{h}^i_{(s-1)}, \mathbf{m}^i_{(s)}\right), \tag{9}$$

$$\dot{f}^i = \varphi_{\dot{f}}\left(\mathbf{h}^i_{(\max s)}\right) \tag{10}$$

where $\mathbf{h}^i$ are hidden node representations, $\mathbf{m}^{ij}$ denotes the message between nodes i and j, and ϕ, φ_i are neural networks. Furthermore, the SinusoidialEmbedding is periodic with respect to global translation and is defined by

$$\text{SinusoidalEmbedding}(x) := (\sin(2\pi k x), \cos(2\pi k x))^T_{k=0,\ldots,n_{freq}}, \tag{11}$$

with hyperparameter n_{freq}. We train all models for 3500 epochs with the hyperparameters shown in Tab. 2.

We use Euler's method with 50 integration steps during inference.

Parameter	Value
Hidden Dimension	128
Time Embedding Dimension	64
Number of Layers	6
Activation Function	silu
n_{freq}	64

Table 2: Model hyperparameters.

C TRANSLATION EQUIVARIANT VECTOR FIELD

We highlight a key difference to FlowMM in the construction of the target VF. In FlowMM, to obtain a solution in the equivalence class of lattice translations, the target VF is made translation invariant by removing the mean velocity $\frac{1}{N}\sum_{i=1}^{n} \log_{f^i} f_1^i$. In Fig. 4, we show conceptually that a translation invariant VF is undesirable for the molecular docking task defined in this work, and motivate keeping the VF translation equivariant.

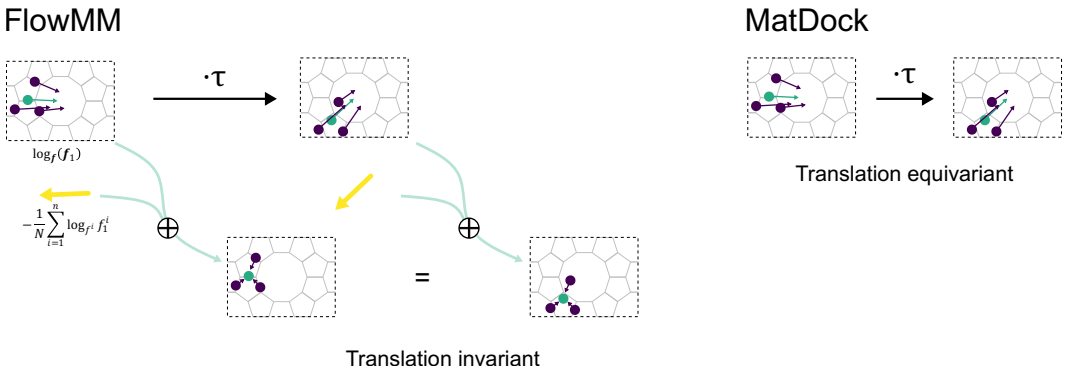

Figure 4: Difference between FlowMM and MatDock in translation symmetry. FlowMM removes the mean target VF, leading to a translationally invariant target VF. MatDock leaves u_t translation equivariant.

D SUCCESSFUL POSES BY LOADING

We show the performance of MatDock when relaxing the docking success definition to more permissible values.

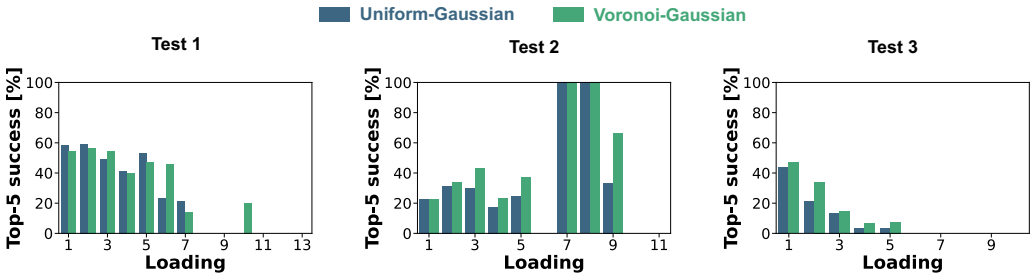

Figure 5: Percentage of successfully docked poses across top 5 samples for different molecular loadings/ zeolite, for test sets Tests 1 - 3 (left to right), comparing between Uniform-Gaussian and Voronoi-Gaussian sampling. In contrast to Fig. 2, this plot uses a molecule-molecule clash threshold of 1.75 Å, a zeolite-molecule clash threshold of 1.50 Å, and an RMSD threshold of 2 Åfor the definition of docking success.

Comparison between Fig. 2 and Fig. 5 reveals that loosening the physical clash thresholds boosts both sampling methods' performance and can change their relative performance. However, V-G sampling's relative performance to U-G sampling still improves with increasing difficulty of the tests sets, supporting the hypothesis that explicitly providing information about empty pore space through Voronoi nodes is useful for the docking task.

E SAMPLING CONFORMERS

Instead of sampling from a Gaussian distribution, we can utilize prior knowledge of the molecular conformation to bias the distribution towards more molecule-like structures. We can model such distribution as $p(\boldsymbol{f}^m) := w\left((\tilde{\boldsymbol{x}^i}\boldsymbol{R}^T)\tilde{\boldsymbol{l}}^{-1} + \mu_m\right)$ with random rotation matrix $\boldsymbol{R} \in \mathbb{R}^{3\times3}$, and mean-free random conformer in Cartesian coordinates $\tilde{\boldsymbol{x}^i} \in \mathbb{R}^{N\times3}$. We show preliminary results for both Uniform-Conformer and Voronoi-Conformer sampling below.

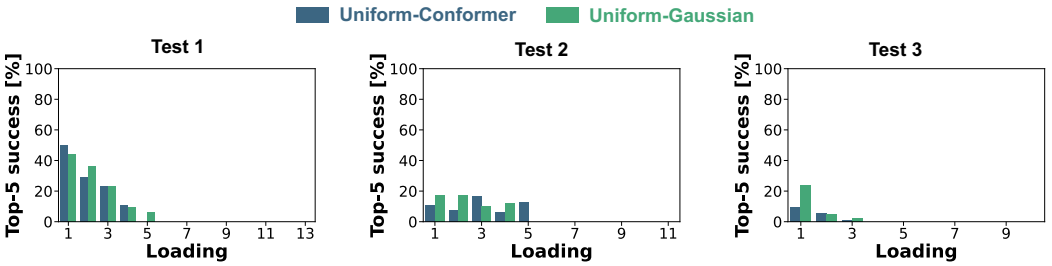

Figure 6: Percentage of successfully docked poses across top 5 samples for different molecular loadings/ zeolite, for test sets Tests 1 - 3 (left to right), comparing between Uniform-Gaussian and Uniform-Conformer sampling.

	Test 1		Test 2		Test 3	
Evaluation	**U-G**	**U-C**	**U-G**	**U-C**	**U-G**	**U-C**
% No mol-mol clash (↑)	**72.95 ± 1.2**	69.29 ± 1.3	**83.63 ± 2.0**	78.81 ± 2.3	**79.25** ± 0.4	73.72 ± 2.5
% No zeolite-mol clash (↑)	**37.56 ± 1.4**	30.94 ± 2.1	**33.78 ± 3.6**	25.78 ± 1.2	**13.28 ± 1.4**	8.84 ± 0.8
% < 2 Å(↑)	63.86 ± 1.3	**69.92 ± 1.1**	**42.22 ± 2.1**	42.07 ± 1.6	**49.31 ± 1.0**	46.81 ± 2.0
% Success (↑)	**21.22 ± 1.1**	19.88 ± 2.3	**10.74 ± 1.7**	7.56 ± 0.7	**7.75 ± 0.9**	4.66 ± 0.9
% Top-5 success (↑)	28.15	**28.35**	**14.81**	9.63	**10.31**	5.0

U-G: Uniform-Gaussian, **U-C**: Uniform-Conformer.

Table 3: Performance metrics for Uniform-Gaussian sampling and Uniform-Conformer sampling.

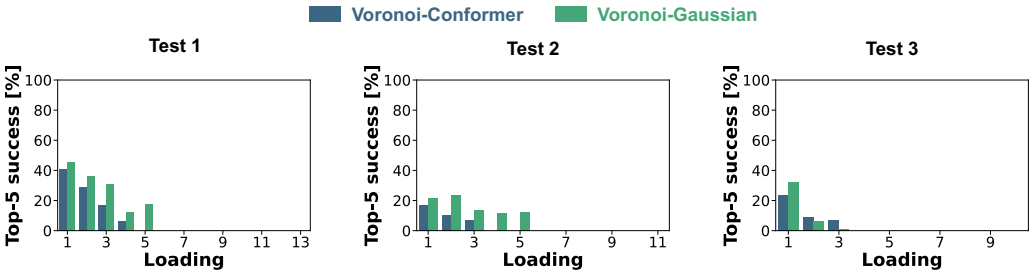

Figure 7: Percentage of successfully docked poses across top 5 samples for different molecular loadings/ zeolite, for test sets Tests 1 - 3 (left to right), comparing between Voronoi-Gaussian and Voronoi-Conformer sampling.

The results in Tables 3 and 4 show superior performance of Gaussian sampling over sampling random conformers in almost all categories. Inspecting individual generated trajectories, we find that

Evaluation	Test 1		Test 2		Test 3	
	V-G	**V-C**	**V-G**	**V-C**	**V-G**	**V-C**
% No mol-mol clash (↑)	**76.26 ± 1.3**	58.82 ± 6.1	**82.89 ± 2.6**	63.48 ± 1.0	**78.06 ± 1.7**	55.06 ± 0.8
% No zeolite-mol clash (↑)	**39.17 ± 1.2**	35.43 ± 1.1	**38.15 ± 2.4**	26.07 ± 2.9	**19.69 ± 0.9**	17.59 ± 1.0
% < 2 Å(↑)	**62.64 ± 1.4**	59.41 ± 5.9	**40.67 ± 0.7**	35.56 ± 1.5	**50.78 ± 1.0**	34.62 ± 1.3
% Success (↑)	**21.77 ± 1.4**	17.64 ± 2.5	**14.15 ± 1.8**	8.37 ± 0.6	**10.31 ± 1.1**	8.66 ± 1.0
% Top-5 success (↑)	**30.51**	23.82	**19.26**	11.48	**13.44**	11.72

V-G: Voronoi-Gaussian, **V-C**: Voronoi-Conformer.

Table 4: Performance metrics for Voronoi-Gaussian sampling and Voronoi-Conformer sampling.

conformers are often split during integration and hence do not stay a coherent molecule. We hypothesize that, in comparison to Gaussian sampling, randomly sampling a conformer leads to lower variance during training , which combined with our (local) message-passing model leads to a smaller receptive field. Therefore, the model perceives less empty pore space and has a harder time finding a valid pose, resulting in worse performance. Further studies could investigate this hypothesis by employing fully-connected graphs or other long-distance message-passing schemes.

F EXAMPLES OF GENERATED SAMPLES

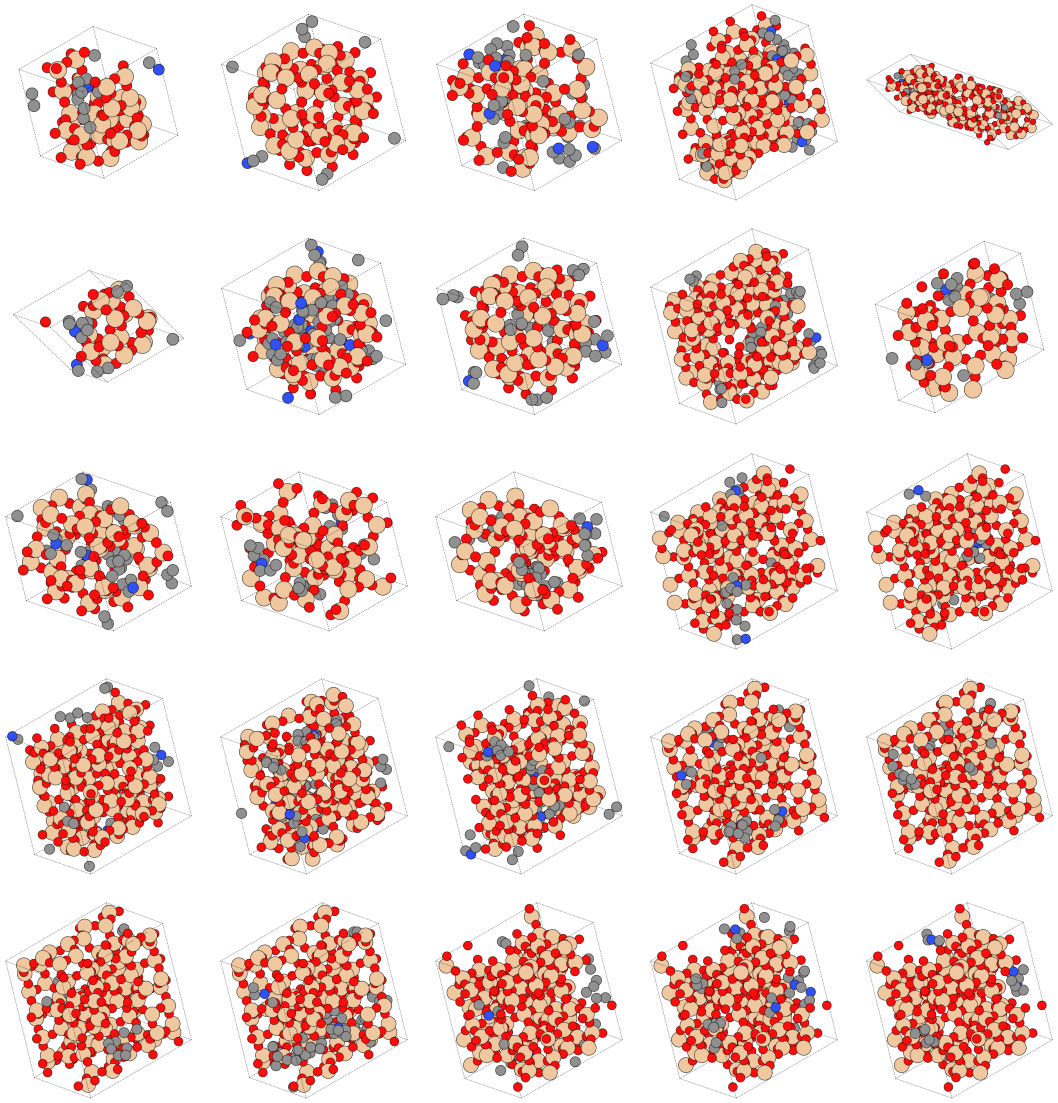

Figure 8: Docked structures created with MatDock with Uniform-Gaussian sampling on test set 1. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.

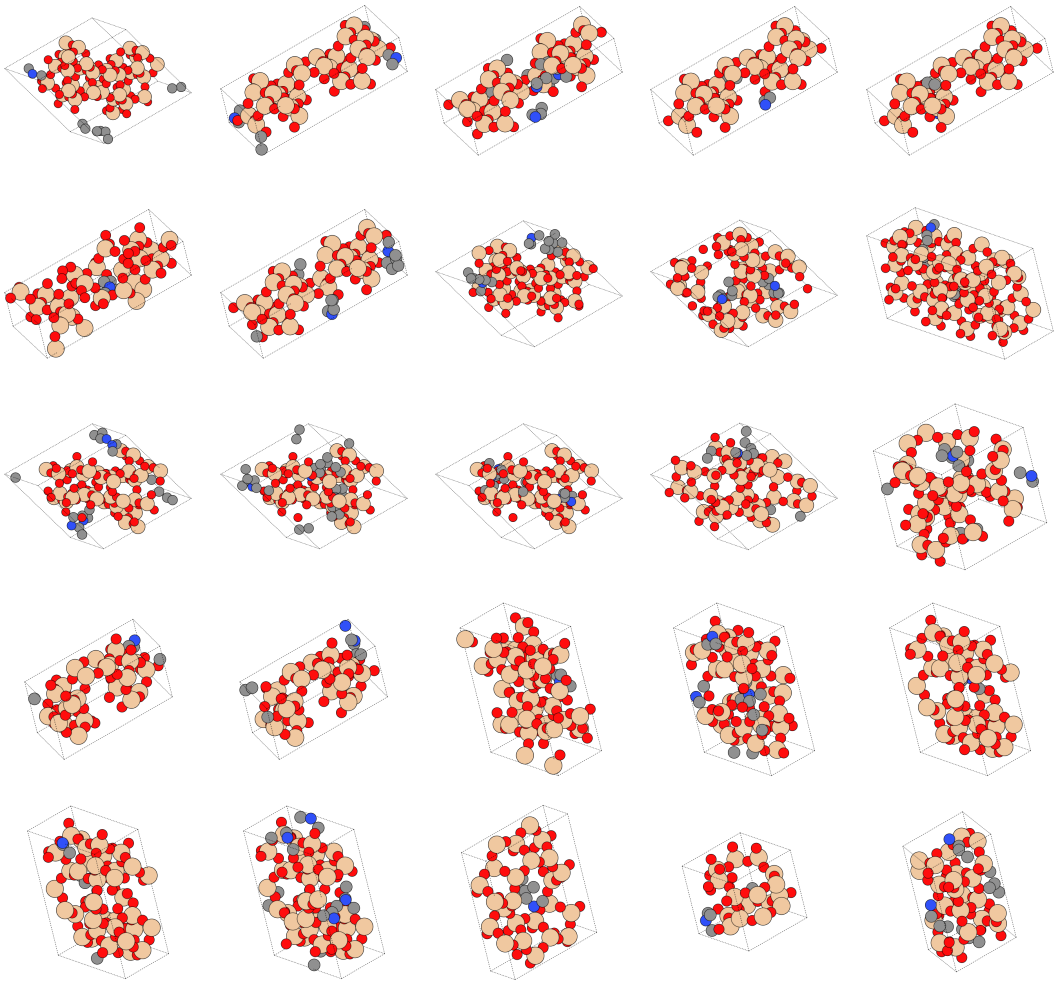

Figure 9: Docked structures created with MatDock with Uniform-Gaussian sampling on test set 2. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.

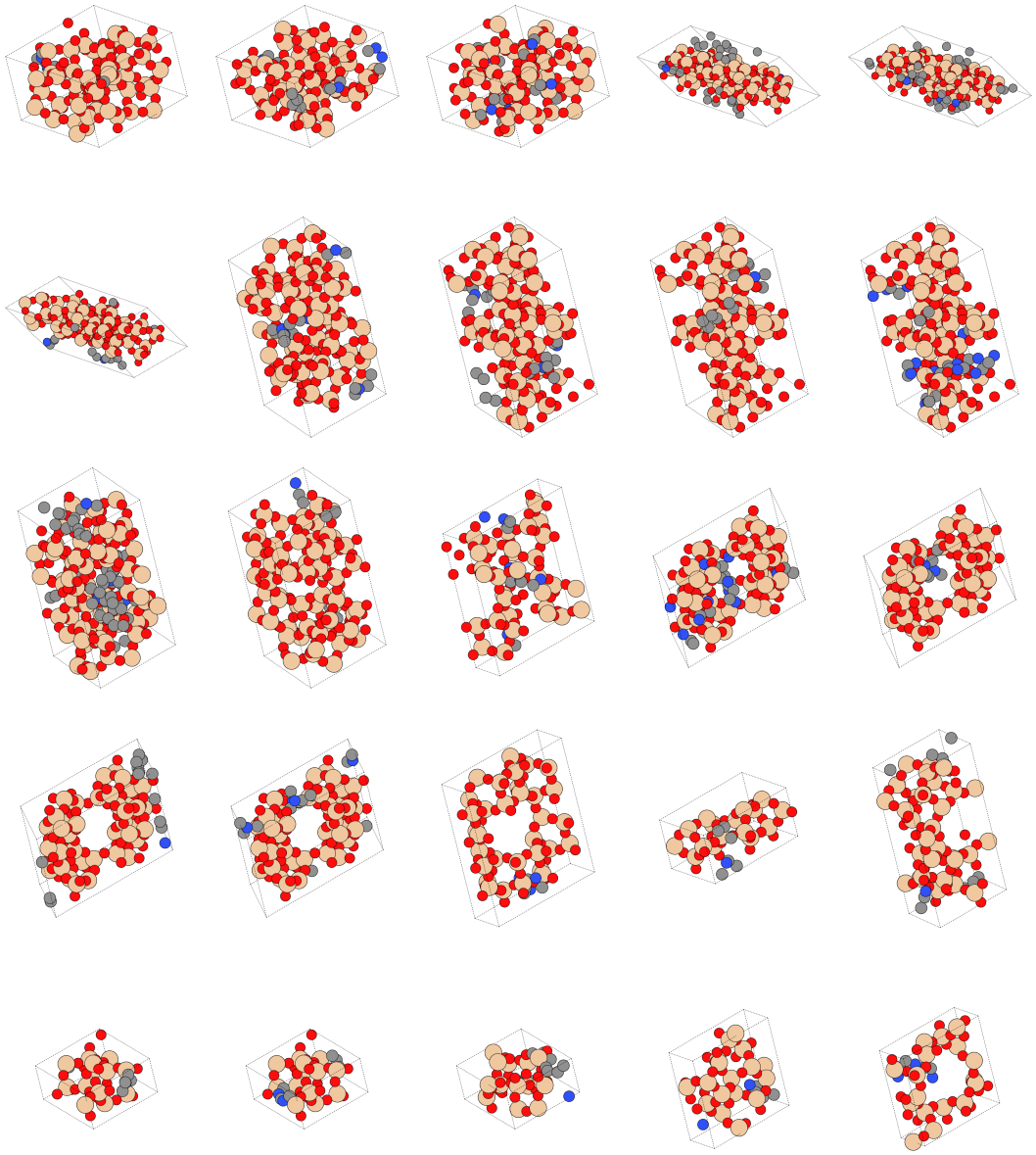

Figure 10: Docked structures created with MatDock with Uniform-Gaussian sampling on test set 3. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.

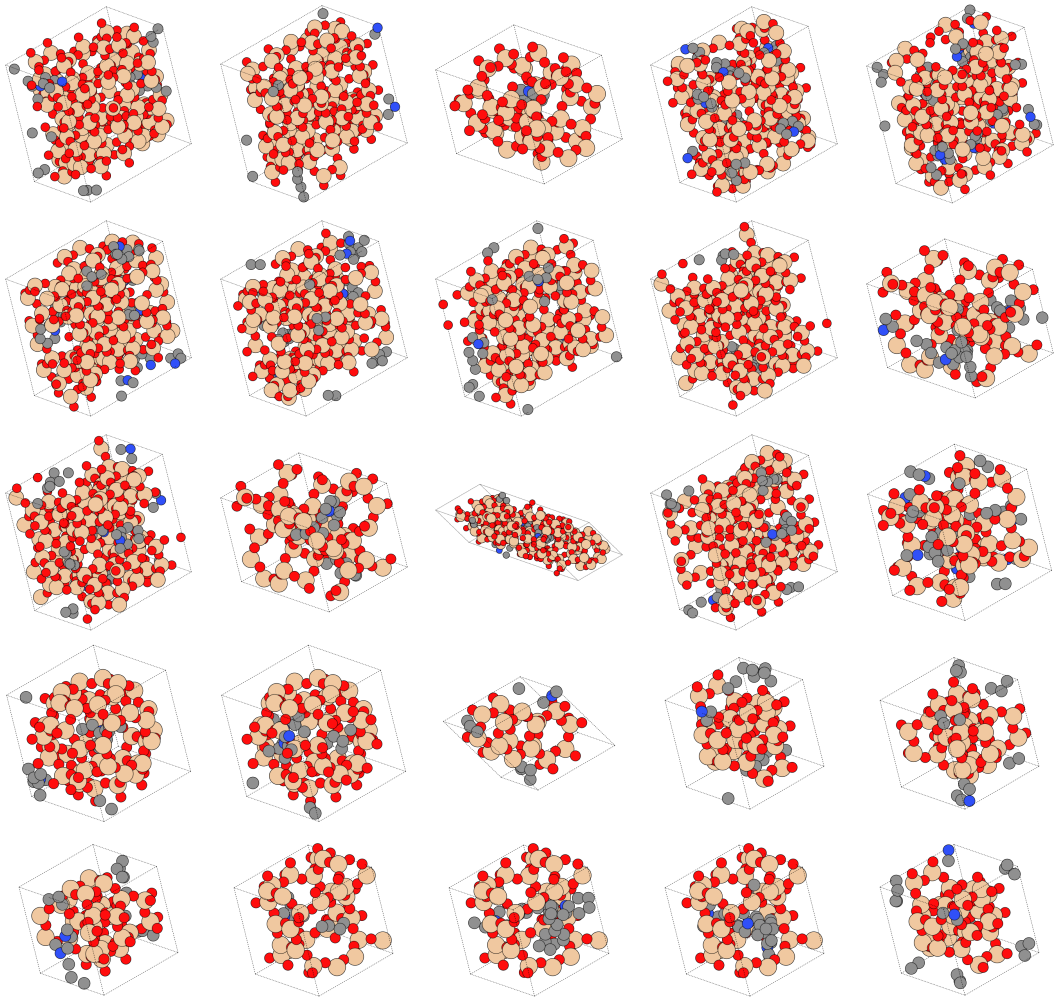

Figure 11: Docked structures created with MatDock with Voronoi-Gaussian sampling on test set 1. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.

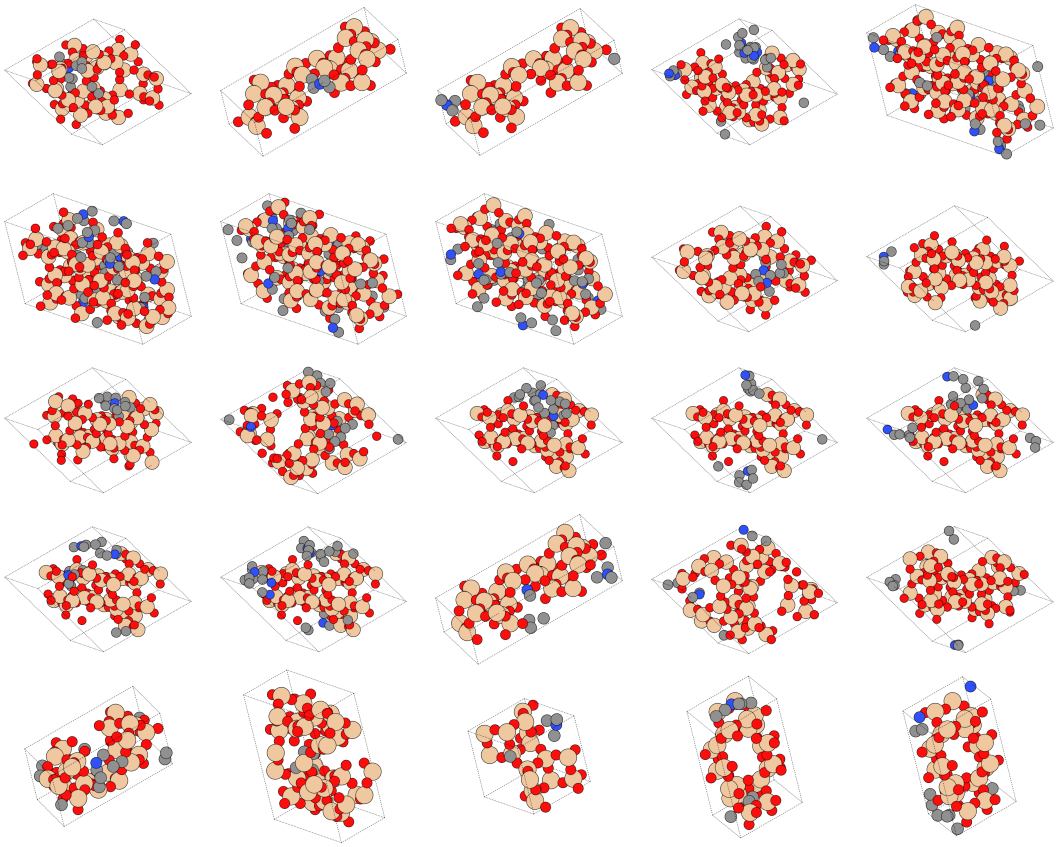

Figure 12: Docked structures created with MatDock with Voronoi-Gaussian sampling on test set 2. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.

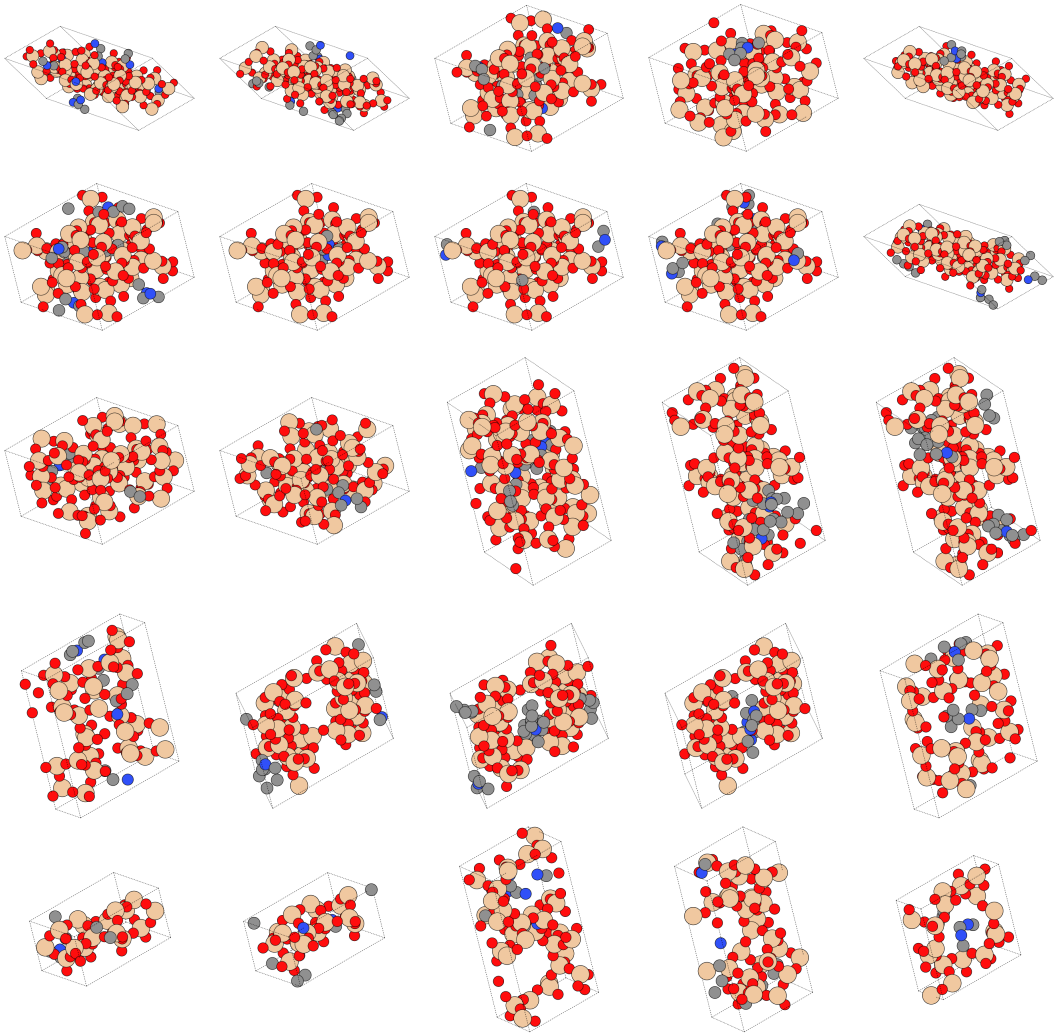

Figure 13: Docked structures created with MatDock with Voronoi-Gaussian sampling on test set 3. For each of the 5 docking runs over the test set, we randomly draw 5 examples. Light orange = Si; dark red = O; grey = C; blue = N. Hydrogens are not shown.