Systems-Structure-Based Drug Design

Vincent D. Zaballa *
Department of Biomedical Engineering
University of California, Irvine
Irvine, CA

Elliot E. Hui
Department of Biomedical Engineering
University of California, Irvine
Irvine, CA

Abstract

Recent advances in generative deep learning have transformed small molecule design, but most methods lack biological systems context, focusing narrowly on specific protein pockets. We introduce a non-differentiable diffusion guidance method that integrates systems biology models, enhancing small molecule generation with pathway context. Using the Bone Morphogenetic Protein (BMP) pathway, we generate small molecules specific to one protein over competing proteins. This method enhances the precision and efficiency of small molecule drug discovery by incorporating systems biology insights into generative models.

1   Introduction

Structure-based drug discovery (SBDD) is a valuable tool for discovering potent small molecule drugs Anderson (2003). SBDD operates off the principle that treating a disease can be reduced to targeting a protein relevant in the disease process, whether directly causal or simply correlated with the disease process. Using SBDD to develop new small molecule therapies provides promise for treating a wide range of diseases, such as more targeted cancer therapeutics Zhong et al. (2021). Enzymes are an example of an excellent target class for SBDD where the active site is typically known and evolutionarily conserved, thus a drug can be created to inhibit this active site to prevent the operation of the protein involved in the disease process. Notable examples are the creation of a small molecule inhibitor of the SARS-CoV-2 (COVID-19) protease Liu et al. (2022) and HIV protease Lam et al. (1994). Recent deep learning methods have naturally fit into the paradigm of SBDD by generating small molecules often conditioned on some type of relevant physical information. These methods are a great fit for inhibiting proteins with known active sites.

On the other hand, there are many proteins that do not have a clear active site, such as those involved in protein-protein interactions (PPIs). Binding to an active site in a PPI, known as orthosteric modulation, depends on finding "hot spots" via mutageneis analysis of the interface of the two proteins, and which may also depend on the plasticity of the proteins involved Ran & Gestwicki (2018); Arkin et al. (2014); Cukuroglu et al. (2014). While mutagenesis can reveal conserved residues at the interface, it assumes knowledge of how proteins may interact in the first place. Assuming knowledge of how two proteins interact, designing a small molecule that is specific to the PPI of interest disregards how the small molecule may interact with similar proteins, such as protein receptor complexes in the TGF-β signaling pathway. This non-specific development likely contributes to downstream failures in clinical trials due to toxicity of the drug and related off-target effects Xie et al. (2011). This risk can be mitigated by further screening for off-target effects but this can be prohibitively expensive and introduces another small molecule screening step that lengthens the process of drug development Dawidowski et al. (2017). There is a need for developing small molecules that are *specific* to the target of interest, yet a paucity of machine learning methods to do this. A notable exception is work by Harris et al. (2023) although their method optimizes small molecules specific to one kinase over another assuming a known binding site in each kinase *a priori*.

---

*email: vzaballa@uci.edu

1. Binding analysis **2.** Pocket identification **4.** Pocket-guided diffusion molecule generation

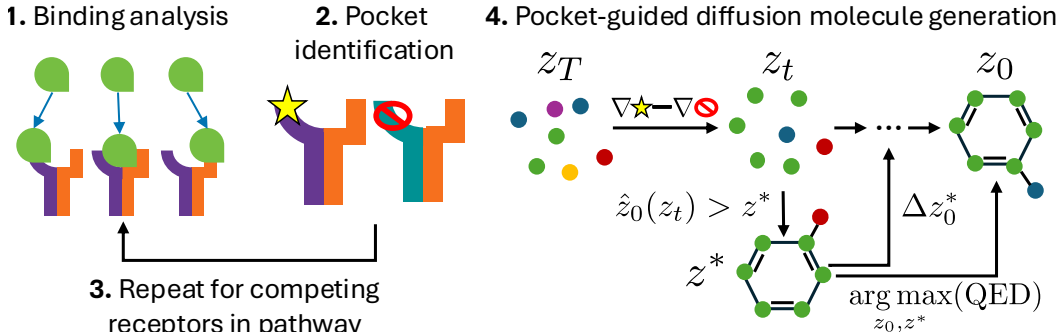**3.** Repeat for competing receptors in pathway

Figure 1: Overview of Systems-Structure-Based Drug Discovery (SSBDD) identifying orthosteric targets and creating a small molecule specific to the target of interest. **1.** Perform a binding analysis to understand which configuration of the protein ligand and the receptor complex is most likely using systems biology data. **2.** Identify the most important point on the protein surface using mutagenesis. **3.** Repeat process for competing protein complexes in the pathway. **4.** Using known orthosteric target sites, perform *forward* guided diffusion conditioning generation steps on the gradient of the desired and undesired target sites. Checkpointing is performed on denoised predictions and checkpointed molecules are used in non-differentiable *reverse* guidance. At the end of sampling, the checkpointed molecule is compared according to a metric with the final denoised molecule and the best is returned.

We demonstrate a bioinformatics pipeline and machine learning method to generate small molecules that are more specific to a protein in a signaling pathway involved in PPIs. The bioninformatics process helps select which proteins may be of interest to target and at which sites to target. The machine learning method is based on an a non-differentiable guided diffusion sampling process conditioned on how well small molecules dock specifically to a target of interest in the pathway of interest. Finally, we evaluate the combined pipeline in an example of the Bone Morphogenetic Protein (BMP) pathway Wang et al. (2014), showing how we are able to guide small molecule generation to inhibit one protein's PPI active site while minimizing the likelihood of binding to another protein in the pathway that shares a protein ligand in the pathway. Given the inclusion of systems biology information Alon (2019), we introduce Systems-Structure-Based Drug Design (SSBDD).

2 Background

Diffusion Models Diffusion models are a class of deep learning method capable of learning the gradient, or score, of a denoising process going from a noisy distribution z_T to the data distribution z_0 by a series of diffusion steps Song & Ermon (2020); Ho et al. (2020). Diffusion has been used in a wide variety of biological tasks such as docking Corso et al. (2022), protein generation Ingraham et al. (2022); Watson et al. (2023), modeling protein conformation landscapes Lu et al. (2024), and small molecule generation Schneuing et al. (2022); Hoogeboom et al. (2022); Morehead & Cheng (2024).

We follow the notation of Hoogeboom et al. (2022) in describing our diffusion process. Data samples are point clouds $z_0 = [x, h]$ which represent the 3D coordinates of N atoms $x \in \mathbb{R}^{N \times 3}$ and atom features $h \in \mathbb{R}^{N \times d}$. Given a data point x, a step in the diffusion process is defined as:

$$q(z_t|x) = \mathcal{N}(z_t|\alpha_t x_t, \sigma_t^2 \mathbf{I}), \tag{1}$$

where $\alpha_t \in \mathbb{R}^+$ controls signal retention and $\sigma_t \in \mathbb{R}^+$ controls the addition of noise. Typically α_t is modeled by a function that smoothly transitions from $\alpha_0 \approx 1$ to $\alpha_T \approx 0$. The noising process can be described by a variance preserving process Sohl-Dickstein et al. (2015); Ho et al. (2020) where $\alpha_t = \sqrt{1 - \sigma_t^2}$. We can also define the signal to noise ratio $\text{SNR}(t) = \alpha_t^2 / \sigma_t^2$ Kingma et al. (2021) to simplify the notation. Since the diffusion is Markovian, we can can write the transition from a time step t to $s < t$ as

$$q(z_s \mid z_{\text{data}}, z_t) = \mathcal{N}\left(z_s \Big| \frac{\alpha_{t|s}\sigma_s^2}{\sigma_t^2} z_t + \frac{\alpha_s \sigma_{t|s}^2}{\sigma_t^2} z_0, \frac{\sigma_{t|s}^2 \sigma_s^2}{\sigma_t^2} \mathbf{I}\right), \tag{2}$$

where $\alpha_{t|s} = \alpha_t/\alpha_s$ and $\sigma_{t|s}^2 = \sigma_t^2 - \alpha_{t|s}^2\sigma_s^2$ following notation of Hoogeboom et al. (2022). The true denoising process depends on the data z_0 which is unavailable when generating new samples. However, an approximate sample $\hat{z}_0$ is calculated using a neural network ϕ_θ. Thus, Eq. (1) can be reparameterized as $z_t = \alpha_t z_0 + \sigma_t \epsilon$ with $\epsilon \sim \mathcal{N}(0, I)$ to directly predict the Gaussian noise $\hat{\epsilon}_\theta = \phi_\theta(z_t, t)$. We can rearrange to show the denoised prediction as a function of the current noisy prediction $\hat{z}_0(z_t) = z_t/\alpha_t - \hat{\epsilon}_\theta \sigma_t/\alpha_t$. We diverge from other notation by showing the prediction as a function of the noised sample at time t. The neural network is trained by maximizing the variational lower bound of the data distribution, equivalent to the simplified training objective $\mathcal{L}_{\text{train}} = \frac{1}{2}||\hat{\epsilon} - \phi_\theta(z_t, t)||^2$ up to a scaling factor Ho et al. (2020); Kingma et al. (2021).

Diffusion Guidance There has been great interest in guiding diffusion models for image generation that range from adding explicit conditional information during the training process Ho & Salimans (2022); Nichol et al. (2022) and guiding the diffusion generation process using a separate and differentiable guidance function after training Dhariwal & Nichol (2021); Wang et al. (2022). Bansal et al. (2023) demonstrated a general-purpose method for guided diffusion simply using a differentiable guidance function. They split guidance into forward and backward guidance. Forward guidance during sample generation updates the noise prediction using a weighted gradient of the guidance function

$$\hat{\epsilon}_\theta(z_t, t) = \epsilon_\theta(z_t, t) + s(t) \cdot \nabla_{z_t}\ell(c, f(\hat{z}_0(z_t))) \tag{3}$$

where we can use the prediction of the denoised data point at time t as $\hat{z}_0(z_t)$, $s(t)$ is a hyperparameter for the guidance strength, c is contextual information such as a protein pocket, and $\ell(\cdot, \cdot)$ is a function that depends on the context and denoised prediction. Backward guidance helps to ensure the generated data point satisfies the guidance function. Instead of calculating the gradient w.r.t. the denoised prediction, we calculate the difference in clean data space Δz_0 by optimizing

$$\Delta z_0 = \arg\min_\Delta \ell(c, f(\hat{z}_0 + \Delta)) \tag{4}$$

where the Δ can be optimized using m-step gradient descent starting at $\Delta = 0$. Once we find the Δ that optimizes the guidance function, we can translate the change back to the original noised data space and calculate the guided denoising prediction $\hat{\epsilon}$ as $z_t = \sqrt{\alpha_t}(\hat{z}_0 + \Delta z_0) + \sqrt{1 - \alpha_t}\hat{\epsilon}$. Using the definition of the predicted denoised data point from the current noised data point, the augmentation to the original denoising prediction is

$$\hat{\epsilon} = \epsilon_\theta(z_t, t) - \sqrt{\alpha_t/1 - \alpha_t}\Delta z_0. \tag{5}$$

Equivariant Graph Neural Networks Atoms live in a space that is invariant to translations and rotations. To build these inductive biases into the neural network ϕ_θ, we use a Graph Neural Network (GNN) that is a function $f : \mathcal{X} \to \mathcal{Y}$ equivariant with respect to the group G if the action of the group element, $g \in G$, on the function is equivalent to the input, i.e. $f(g.x) = g.f(x)$ where $g.$ represents the group action. We follow Morehead & Cheng (2024) and use the GCPNET++ GNN model that maintains SE(3) equivariance, which is equivariant to 3D roto-translations, while maintaining chirality of molecules by using Geometry-Complete Perceptron Networks (GCPNETS) that define direction-robust local geometric reference frames Morehead & Cheng (2023).

Molecular Docking Docking of small molecule candidates is split between known-pocket and blind docking. Corso et al. (2022) demonstrated the use of diffusion methods for the blind docking setting while achieving state of the art results at the time. Related work for known-pocket diffusion, DiffDock-Pocket Plainer et al. (2023), predicts ligand poses in a specific binding pocket of interest, such as orthosteric sites, by modeling the flexibility of the side chains of a binding pocket that would influence the conformation of the docked small molecule. Recent methods have focused on the problem of adequately modeling the protein conformation when unbound, its apoprotein state, and when bound to a molecule, its holoprotein state Corso et al.; Lu et al. (2024). Regardless of the scope, most diffusion-based docking methods include a confidence model that provides a score of how confident the model is in its prediction of a ligand to the binding spot.

Systems Biology & Protein-Protein Interactions Understanding how proteins interact is critical to develop therapies for PPI inhibitors. Systems biology is the study of biological networks, including networks of proteins. Protein network response can be modeled using mass action kinetics Antebi et al. (2017) and have physically-meaningful parameters that can be linked to protein structure predictions Zaballa & Hui (2024). This link provides a way to evaluate which structural interactions are likely while cross-validating with a gene reporter assay data fit to systems biology models.

3 Systems-Structure-Based Drug Discovery

Identifying Targets We follow the method of Zaballa & Hui (2024) and identify PPI candidates by evaluating structural predictions' binding affinities with systems biology data. Briefly, we perform evolutionary analysis to identify conserved regions of the proteins, dock using Ambiguous Interaction Restraints informed by conserved residues, and cross-validate a structural binding affinity prediction with systems biology data. Furthermore, we use the conserved regions at the interface of two proteins in the pathway to identify a candidate orthosteric binding pocket that can be used to generate novel molecules. We repeat this process for competing proteins in the pathway that share a diffusible protein ligand in order to find the pocket that we want to avoid hitting. More details can be found in Appendix A.3.

Non-Differentiable Diffusion Guidance To guide molecule generation specific to one orthosteric site and not to another, we use DiffDock-Pocket Plainer et al. (2023) as the guidance function and use the confidence prediction of a generated small molecule as the guidance signal. Thus, we would like higher confidence for the pocket of interest and lower confidence for the competing protein. While the confidence model can provide a signal it is not a differentiable function, which is required for both forward and backward guidance. We overcome this by approximating the gradient of the forward diffusion guidance using the central difference form of finite difference with an offset of $\varepsilon = 1e^{-2}$ as

$$\nabla_{z_t}\ell(c, f(\hat{z}_0)) \approx \frac{\ell(c, f(\hat{z}_0(z_t + \varepsilon))) - \ell(c, f(\hat{z}_0(z_t - \varepsilon)))}{2\varepsilon}. \tag{6}$$

In our approach to reverse guidance, instead of perturbing the clean data point within the guidance function, we introduce a checkpointing mechanism that selects a molecule based on its guidance score as the basis for future reverse guidance steps. Specifically, at each reverse step, we evaluate the denoised prediction $\hat{z}_0$ for the current z_t and compare it to a checkpointed molecule z^*, which represents the molecule that has achieved the highest guidance score in previous iterations. The guidance score is evaluated using the quantitative estimate of drug-likeness drug-likeness metric (QED) Bickerton et al. (2012), where a higher score corresponds to a more desirable molecule. If the current denoised prediction $\hat{z}_0$ achieves a higher QED score than z^*, we update z^* to this new molecule. We then calculate an adjusted reverse diffusion gradient point as $\Delta z_0^* = z_0^* - \hat{z}_0(z_t)$ This mechanism allows us to use the best-performing molecule as the reference for subsequent reverse guidance steps, ensuring that the generated molecule is progressively optimized according to the guidance metric. Thus we adjust Eq. (5) to be

$$\hat{\epsilon} = \epsilon_\theta(z_t, t) - \sqrt{\alpha_t/1 - \alpha_t}\Delta z_0^*. \tag{7}$$

Since this method relies on a good checkpoint, we predict a batch of samples $\hat{z}_0(z_t)$ and checkpoint on the best sample. Curiously, we found that this crude reverse guidance was required for convergence of non-differentiable guided diffusion. This may be due to forward guidance steering sampled molecules out of distribution of a lower-dimensional manifold and the reverse guidance providing a signal that steers the sample back to the data manifold. Finally, we can compare this checkpointed molecule to the final diffusion guidance molecule and return the molecule that achieves the best QED score. We chose checkpointing on the QED score instead of an aggregate of the confidences and the QED as it helped to stabilize the sampling process. The complete algorithm can be seen in Appendix A.1.

4 Experiments

We evaluated SSBDD on a model of the BMP pathway Su et al. (2022) using previously-collected data Klumpe et al. (2022) to perform the systems biology validation. In particular, we chose to design a small molecule specific to the BMPR1A-ACVR2A receptor complex in combination with the BMP4 diffusible ligand due to the high binding affinity of the complex. We chose the ACVR1-ACVR2A receptor complex due to its predicted high binding affinity to the BMP4 protein ligand. Since both complexes bind strongly to BMP4, we would like to design a small molecule that is more specific to BMPR1A-ACVR2A-BMP4 (BA4) than ACVR1-ACVR2A-BMP4 (AA4). We found "hot spots" that are likely important to the PPI between each receptor complex and the BMP4 diffusible ligand using evolutionary analysis data. We used the coordinates of the "hot spots" as the center of the DiffDock-Pocket diffusion process. Selection of the protein complexes and evolutionary analysis is discussed in Appendix A.3.

Table 1: Metrics of small molecules generated with 60, 75, and 90 atoms using three samples each ($n = 3$) on QED score, SA score, and the confidence for both the BA4 and AA4 protein complex pockets of interest. Values are reported as mean $\pm$ standard deviation.

Number of Atoms	QED ($\uparrow$)	SA ($\downarrow$)	BA4 Conf. ($\uparrow$)	AA4 Conf. ($\downarrow$)
60	$\mathbf{0.764 \pm 0.016}$	6.805 ± 1.205	-0.732 ± 0.363	$\mathbf{-7.851 \pm 0.702}$
75	0.703 ± 0.019	$\mathbf{5.181 \pm 1.088}$	-0.482 ± 1.037	-7.782 ± 0.387
90	0.740 ± 0.065	5.533 ± 1.353	-1.641 ± 1.390	-6.480 ± 1.390

To perform the guided diffusion process, we generated 60, 75, and 90 atoms. We return the final or checkpointed molecules and then evaluated according to its QED, synthetic accessibility (SA) score, and DiffDock-Pocket confidence score in each pocket for both the pocket of interest and competing pocket. We show a list of the sampling hyperparameters in Appendix A.2.

Table 1 shows SSBDD is able to consistently generate small molecules specific to the target of interest, as demonstrated by the BA4 and AA4 confidence scores. We found that there is likely an optimal sized small molecule that specifically binds a target of interest. We see the smaller and larger molecules generally return a better QED score than the 75-atom small molecule but the 75-atom molecule is generally more confident in the target of interest than the small and large small molecules. In general, we also found that checkpointing on earlier predicted molecules generally led to smaller final molecules since we chose to keep the largest fragment of a structure prediction. This variability in size introduces more randomness in the process and treats the atom size as an upper bound of the molecule size, and suggests optimizing for the size of a small molecule is an important variable when making target-specific molecules.
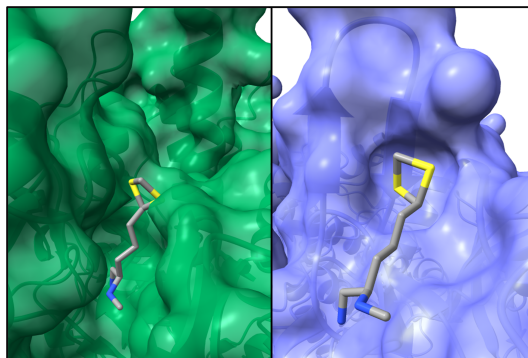

Figure 2: A 75-atom small molecule generated and fit into the competing (left) and targeted (right) pocket of interest. Vina score (lower is better) for the competing complex was $\mathbf{4.27}$ kcal/mol and $-\mathbf{1.91}$ kcal/mol for the targeted pocket.

We evaluated the Autodock-Vina score in one of the 75-atom small molecules in the target and competing protein pockets Fig. 2 and found dramatically stronger binding affinity of the molecule to the target of interest as opposed to the competing target. This seems counterintuitive as the competing protein pocket is larger and secluded from water, but the ligand's size may prevent it from fully engaging in key bending residues. By contrast, the ligand in the targeted pocket better fits the pocket's shape likely resulting in more energetically favorable interactions, such as hydrogen bonding and van der Waals contacts, leading to a stronger overall binding affinity.

5   Discussion

We demonstrated a bioinformatics pipeline for selecting a protein pocket of interest, as well as competing pockets, and a novel non-differentiable diffusion guidance method that is able to generate molecules with greater specificity to the target of interest. The non-differentiable guidance function provides greater flexibility in terms of which functions can guide the diffusion sampling process. In evaluating our method across various molecule sizes, we observed that medium-sized molecules typically exhibited superior performance in the targeted pocket; however, the optimal molecule size likely varies depending on the characteristics of both the target and competing pockets.

Future Work While our method helps overcome challenges associated with non-differentiable functions, there is a penalty in sampling efficiency by having to evaluate non-differentiable functions with a batch of molecules taking 1-2 hours to sample. Creating a differentiable alternative to DiffDock-Pocket Plainer et al. (2023), e.g. using flow matching, would greatly accelerate sampling. Additionally, adapting diffusion models that generate small molecules of variable size Campbell et al. (2023) would allow for a more in-depth study of the interdependence between size of the molecule and shape of binding pockets of interest.

Acknowledgements

References

Alon, U. *An introduction to systems biology: design principles of biological circuits.* Chapman and Hall/CRC, 2019.

Anderson, A. C. The process of structure-based drug design. *Chemistry & biology*, 10(9):787–797, 2003.

Antebi, Y. E., Linton, J. M., Klumpe, H., Bintu, B., Gong, M., Su, C., McCardell, R., and Elowitz, M. B. Combinatorial signal perception in the bmp pathway. *Cell*, 170(6):1184–1196, 2017.

Arkin, M. R., Tang, Y., and Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chemistry & biology*, 21(9):1102–1114, 2014.

Bansal, A., Chu, H.-M., Schwarzschild, A., Sengupta, S., Goldblum, M., Geiping, J., and Goldstein, T. Universal guidance for diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 843–852, 2023.

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., and Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012.

Campbell, A., Harvey, W., Weilbach, C., Bortoli, V. D., Rainforth, T., and Doucet, A. Trans-dimensional generative modeling via jump diffusion models, 2023. URL `https://arxiv.org/abs/2305.16261`.

Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., Pupko, T., and Ben-Tal, N. Consurf: using evolutionary data to raise testable hypotheses about protein function. *Israel Journal of Chemistry*, 53(3-4):199–206, 2013.

Corso, G., Somnath, V. R., Getz, N., Barzilay, R., Jaakkola, T., and Krause, A. Flexible docking via unbalanced flow matching. In *ICML'24 Workshop ML for Life and Material Science: From Theory to Industry Applications*.

Corso, G., Stärk, H., Jing, B., Barzilay, R., and Jaakkola, T. Diffdock: Diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv:2210.01776*, 2022.

Cukuroglu, E., Engin, H. B., Gursoy, A., and Keskin, O. Hot spots in protein–protein interfaces: Towards drug discovery. *Progress in biophysics and molecular biology*, 116(2-3):165–173, 2014.

Dawidowski, M., Emmanouilidis, L., Kalel, V., Tripsianes, K., Schorpp, K., Hadian, K., Kaiser, M., Mäser, P., Kolonko, M., Tanghe, S., et al. Inhibitors of pex14 disrupt protein import into glycosomes and kill trypanosoma parasites. *Science*, 355(6332):1416–1420, 2017.

Dhariwal, P. and Nichol, A. Diffusion models beat gans on image synthesis, 2021. URL `https://arxiv.org/abs/2105.05233`.

Harris, C., Didi, K., Schneuing, A., Du, Y., Jamasb, A. R., Bronstein, M. M., Correia, B., Lio, P., and Blundell, T. L. Flexible small-molecule design and optimization with equivariant diffusion models. In *ICLR 2023-Machine Learning for Drug Discovery workshop*, 2023.

Ho, J. and Salimans, T. Classifier-free diffusion guidance, 2022. URL `https://arxiv.org/abs/2207.12598`.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models, 2020. URL `https://arxiv.org/abs/2006.11239`.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3d, 2022. URL `https://arxiv.org/abs/2203.17003`.

Ingraham, J., Baranov, M., Costello, Z., Barber, K., Wang, W., Ismail, A., Frappier, V., Lord, D., Ng-Thow-Hing, C., Van Vlack, E., et al. Illuminating protein space with a programmable generative model. biorxiv 2022. *CrossRef][PubMed]*, 2022.

Ketata, M. A., Laue, C., Mammadov, R., Stärk, H., Wu, M., Corso, G., Marquet, C., Barzilay, R., and Jaakkola, T. S. Diffdock-pp: Rigid protein-protein docking with diffusion models. *arXiv preprint arXiv:2304.03889*, 2023.

Kingma, D., Salimans, T., Poole, B., and Ho, J. Variational diffusion models. *Advances in neural information processing systems*, 34:21696–21707, 2021.

Klumpe, H. E., Langley, M. A., Linton, J. M., Su, C. J., Antebi, Y. E., and Elowitz, M. B. The context-dependent, combinatorial logic of bmp signaling. *Cell systems*, 13(5):388–407, 2022.

Lam, P. Y., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N., et al. Rational design of potent, bioavailable, nonpeptide cyclic ureas as hiv protease inhibitors. *Science*, 263(5145):380–384, 1994.

Liu, H., Iketani, S., Zask, A., Khanizeman, N., Bednarova, E., Forouhar, F., Fowler, B., Hong, S. J., Mohri, H., Nair, M. S., et al. Development of optimized drug-like small molecule inhibitors of the sars-cov-2 3cl protease for treatment of covid-19. *Nature communications*, 13(1):1891, 2022.

Lu, W., Zhang, J., Huang, W., Zhang, Z., Jia, X., Wang, Z., Shi, L., Li, C., Wolynes, P. G., and Zheng, S. Dynamicbind: Predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model. *Nature Communications*, 15(1):1071, 2024.

Morehead, A. and Cheng, J. Geometry-complete perceptron networks for 3d molecular graphs, 2023. URL https://arxiv.org/abs/2211.02504.

Morehead, A. and Cheng, J. Geometry-complete diffusion for 3d molecule generation and optimization. *Communications Chemistry*, 7(1):150, 2024.

Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I., and Chen, M. Glide: Towards photorealistic image generation and editing with text-guided diffusion models, 2022. URL https://arxiv.org/abs/2112.10741.

Plainer, M., Toth, M., Dobers, S., Stark, H., Corso, G., Marquet, C., and Barzilay, R. Diffdock-pocket: Diffusion for pocket-level docking with sidechain flexibility. 2023.

Ran, X. and Gestwicki, J. E. Inhibitors of protein–protein interactions (ppis): an analysis of scaffold choices and buried surface area. *Current opinion in chemical biology*, 44:75–86, 2018.

Schneuing, A., Du, Y., Harris, C., Jamasb, A., Igashov, I., Du, W., Blundell, T., Lió, P., Gomes, C., Welling, M., et al. Structure-based drug design with equivariant diffusion models. *arXiv preprint arXiv:2210.13695*, 2022.

Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., and Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. PMLR, 2015.

Song, Y. and Ermon, S. Generative modeling by estimating gradients of the data distribution, 2020. URL https://arxiv.org/abs/1907.05600.

Su, C. J., Murugan, A., Linton, J. M., Yeluri, A., Bois, J., Klumpe, H., Langley, M. A., Antebi, Y. E., and Elowitz, M. B. Ligand-receptor promiscuity enables cellular addressing. *Cell systems*, 13(5): 408–425, 2022.

Van Zundert, G., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P., Karaca, E., Melquiond, A., van Dijk, M., De Vries, S., and Bonvin, A. The haddock2. 2 web server: user-friendly integrative modeling of biomolecular complexes. *Journal of molecular biology*, 428(4):720–725, 2016.

Wang, R. N., Green, J., Wang, Z., Deng, Y., Qiao, M., Peabody, M., Zhang, Q., Ye, J., Yan, Z., Denduluri, S., et al. Bone morphogenetic protein (bmp) signaling in development and human diseases. *Genes & diseases*, 1(1):87–105, 2014.

Wang, Y., Yu, J., and Zhang, J. Zero-shot image restoration using denoising diffusion null-space model, 2022. URL `https://arxiv.org/abs/2212.00490`.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

Xie, L., Xie, L., and Bourne, P. E. Structure-based systems biology for analyzing off-target binding. *Current opinion in structural biology*, 21(2):189–199, 2011.

Zaballa, V. D. and Hui, E. E. Reducing uncertainty through mutual information in structural and systems biology. *arXiv preprint arXiv:2407.08612*, 2024.

Zhong, L., Li, Y., Xiong, L., Wang, W., Wu, M., Yuan, T., Yang, W., Tian, C., Miao, Z., Wang, T., et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. *Signal transduction and targeted therapy*, 6(1):1–48, 2021.

A Appendix

A.1 Non-Differentiable Diffusion Guidance Algorithm

We follow the algorithm notation from Bansal et al. (2023) in Algorithm 1 and denote the prediction of a new data point in sampling as $S(z_t, \hat{\epsilon}_\theta, t)$, where the subsequent diffusion step is dependent on the current noised data point, prediction, and time point.

Algorithm 1 Non-Differentiable Universal Guidance

Parameter: Recurrent steps k, number of checkpoint samples m, and guidance strength $s(t)$
Required: z_T sampled from $\mathcal{N}(0, I)$, diffusion model ϵ_θ, noise scales $\{\alpha_t\}_{t=1}^{T}$, guidance function f, loss function ℓ, and prompt c

    $z^* \leftarrow 0$
    for $t = T, T-1, \ldots, 1$ **do**
        for $n = 1, 2, \ldots, k$ **do**
            Calculate $\hat{z}_0(z_t) = z_t/\alpha_t - \epsilon_\theta \sigma_t/\alpha_t$
            Calculate $\hat{\epsilon}_\theta$ using forward universal guidance as in Eq. (3)
            Sample m values of $z_t^{(i)} \leftarrow S(z_{t+1}, \hat{\epsilon}_\theta, t+1)$ for $i = 1, 2, ..., m$
            for each sample $z_t^{(i)}$, where $i = 1$ to m **do**
                if $\hat{z}_0(z_t^{(i)}) > z^*$ **then**
                    Update $z^* \leftarrow \hat{z}_0(z_t^{(i)})$
                end if
            end for
            $\Delta z_0^* \leftarrow z^* - \hat{z}_0(z_t)$
            Perform reverse guidance $\hat{\epsilon} \leftarrow \hat{\epsilon} - \sqrt{\alpha_t/(1-\alpha_t)}\Delta z_0^*$
        end for
        $z_{t-1} \leftarrow S(z_t, \hat{\epsilon}_\theta, t)$
        $\epsilon' \sim \mathcal{N}(0, I)$
        $z_t \leftarrow \sqrt{\alpha_t}z_{t-1} + \sqrt{1-\alpha_t}\epsilon'$
    end for

A.2 Sampling Hyperparameters & Details

The sampling hyperparameters used can be seen in Table 2. Since predicting a denoised data point is difficult early in the sampling process, we do not begin Algorithm 1 until half way through generation, when $t = 0.5$, assuming the noise distribution is at $T = 1$ and the data distribution is at $t = 0$. Additionally, many time steps are unable to predict a denoised molecule so we implement the algorithm using a try/except loop, falling back to normal diffusion sampling if predicting a denoised molecule is not possible.

Table 2: Sampling hyperparameters.

Recurrent steps, k	3
Guidance strength, $s(t)$	$0.1 * \sqrt{1 - \alpha_t}$
Finite difference, ε	$1e^{-2}$
Samples m from predicted batch $z_t \leftarrow S(z_{t+1}, \hat{\epsilon}_\theta, t+1)$	10

A.3 Selection of BMP Signaling Proteins

We used a fit model of the BMP pathway to assess which complexes might be good candidates for inhibition using a small molecule as shown in Fig. 3. Given that both histograms place more density at the higher end of binding affinities, which represents stronger binding between the complex and the BMP4 homodimer, we determined they were good candidates to perform PPI inhibition.

After determining that the two complexes were good candidates, we performed a docking analysis to determine which parts of the proteins bound with one another. We used HADDOCK Van Zundert et al. (2016) and Ketata et al. (2023) to find probable binding sites betwee the receptor complex and
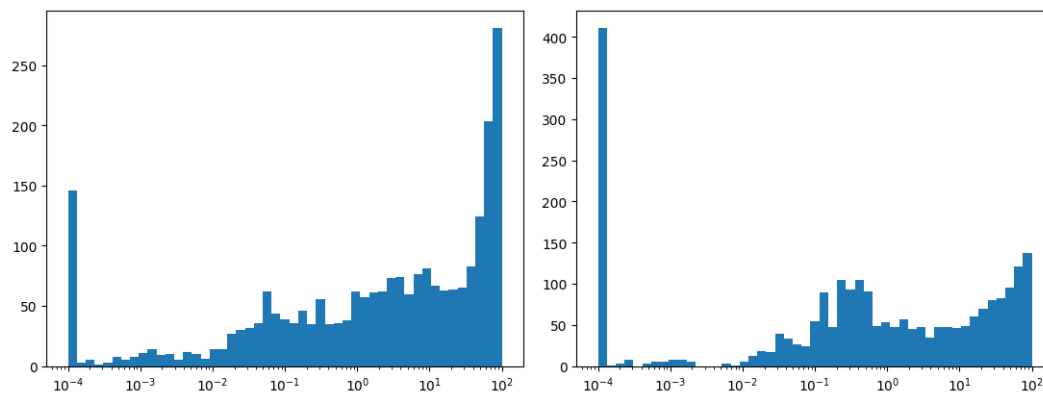
Figure 3: Histogram of the posterior parameter fit of the BMP model for the ACVR1-ACVR2A-BMP4 (left) and BMPR1A-ACVR2A-BMP4 (right) trimeric complexes. Binding affinity K_{eq}, which is $K_{eq} = -\ln k_d$, is shown on the x-axis and where higher values mean more tightly binding complexes. Both suggest a strong binding affinity between the BMP4 homodimer and receptor complex.

the BMP4 homodimer ligand. Additionally, after analyzing the BMP pathway and finding good PPI candidates, we performed an evolutionary analysis using the ConSurf server Celniker et al. (2013) to find small molecule binding "hot spots". The two points can be seen in Fig. 4.
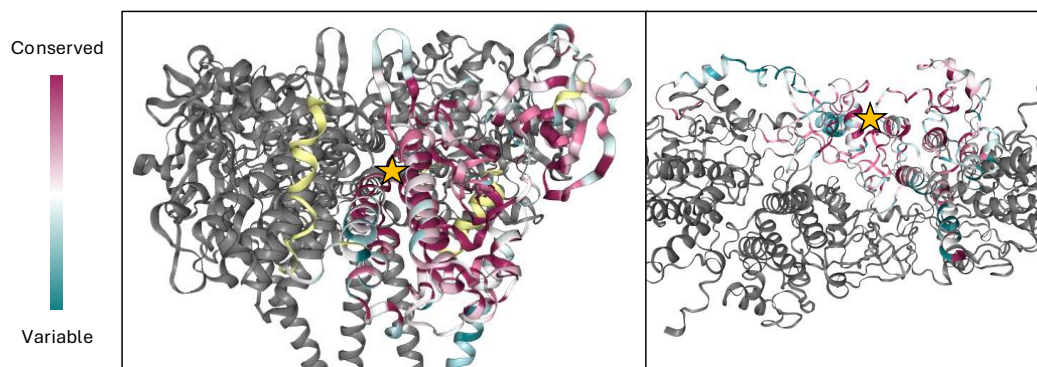

Figure 4: Evolutionary analysis demonstrating "hot spot" in both the BMPR1A-ACVR2A (left) and the ACVR1-ACVR2A (right) receptors. The yellow start denotes a conserved point that also overlaps with a probable binding domain between each receptor complex and the BMP4 homodimer ligand. The BMPR1A-ACVR2A is shown from the side and the ACVR1-ACVR2A is shown from the top.