GeoDirDock: Guiding Docking Along Geodesic Paths

Raúl Miñán, Javier Gallardo [*] **Álvaro Ciudad, Alexis Molina** [†]
Department of Artificial Intelligence
Nostrum Biodiscovery S.L.
Barcelona, Spain
`{alvaro.ciudad,alexis.molina}@nostrumbiodiscovery.com`

Abstract

This work introduces GeoDirDock (GDD), a novel approach to molecular docking that enhances the accuracy and physical plausibility of ligand docking predictions. GDD guides the denoising process of a diffusion model along geodesic paths within multiple spaces representing translational, rotational, and torsional degrees of freedom. Our method leverages expert knowledge to direct the generative modeling process, specifically targeting desired protein-ligand interaction regions. We demonstrate that GDD significantly outperforms existing blind docking methods in terms of RMSD accuracy and physicochemical pose realism. Our results indicate that incorporating domain expertise into the diffusion process leads to more biologically relevant docking predictions. Additionally, we explore the potential of GDD for lead optimization in drug discovery through angle transfer in maximal common substructure (MCS) docking, showcasing its capability to predict ligand orientations for chemically similar compounds accurately.

1 Introduction

In drug discovery, molecular docking is pivotal for probing interactions between molecules and specific protein cavities, also known as binding pockets (Meng et al., 2011). These regions on protein surfaces are targets for docking algorithms, aiming to establish the most stable configurations of ligands, or ligand *poses*, for assessing potential drug-receptor interactions. Docking traditionally relies on empirical data from *in vitro* experiments, such as crystal structures and molecular dynamics simulations, to locate these cavities for molecular binding assessment. Renowned methods like Glide, Vina, and rDock leverage physics-based scoring and a rich database of molecular interactions to navigate this process (Friesner et al., 2004; Trott & Olson, 2010; Ruiz-Carmona et al., 2014).

While traditional methods have long been a cornerstone in computational chemistry, the field experienced a paradigm shift with DiffDock (Corso et al., 2023). This seminal contribution postulated molecular docking as a generative task, marking a significant shift through the development of a diffusion method that operates, without a prior over the binding site location, across the various degrees of freedom inherent in molecular docking. However, limitations appear during prospective applications, due to its blind docking strategy. Directed docking, which employs expert knowledge or pocket prediction to identify binding regions, has demonstrated superior performance compared to blind docking (Ghersi & Sanchez, 2009). Additionally, the use of ligand root mean squared distance (RMSD) as the sole performance metric has been contested for not fully capturing ligand position accuracy, raising questions about the effectiveness of recent deep learning innovations in this field. (Buttenschoen et al., 2023; Yu et al., 2023).

In this work, we introduce GeoDirDock, a diffusion method that integrates expert knowledge into the diffusion process over translations, rotations, and torsions, to direct the diffusion

---

[*]Work done during an internship at Nostrum Biodiscovery.
[†]Author has an affiliation with Barcelona Supercomputing Center.

process towards a desired region of the protein structure. Our results show that this directed docking approach significantly improves upon blind docking methods, achieving close-to-ground truth RMSD conformations in self-docking scenarios and demonstrating enhanced performance in both RMSD accuracy and physicochemical pose plausibility. Moreover, we validate the robustness of our approach through a maximal common substructure docking test, highlighting its effective generalization across diverse and previously unseen ligand chemistries.

2 RELATED WORK

Traditional Prior-Informed Docking Methods. Traditional docking evaluates how likely molecules are to bind to specific areas of proteins, mainly focusing on binding pockets. These methods capitalize on data from crystal structures, literature, and molecular dynamics simulations to set pose search parameters and find the most suitable docking positions, using physics-based scoring methods to measure molecular interactions (Meng et al., 2011; Koes et al., 2013; McNutt et al., 2021). However, the effectiveness of these methods largely relies on having access to high-quality data, which can be a limitation in cases where such data is scarce.

Diffusion-Based Molecular Modeling. Recent advancements in diffusion models have relied upon the concept of blind docking, which does not rely on pre-existing receptor-ligand binding information. DiffDock, for example, replaces traditional methods with a denoising process that manipulates translations, rotations, and torsion angles in a $\mathbb{T}^3 \times SO(3) \times SO(2)^m$ diffusion space, evaluated using a confidence-scoring network. (Nakata et al., 2023) approached the diffusion problem in R^3, incorporating equivariant constraints across the entire space and employing an equivariant graph neural network for scoring. (Qiao et al., 2022) introduced contact prediction modules to direct the molecular diffusion process through the R^3 manifold, assessing sample plausibility with invariant point attention.

Informed Diffusion for Directed Docking. Most existing diffusion-based methods, including the ones discussed above, do not explicitly consider the precise location of the binding cavity. Despite the technical advancements in blind diffusion docking, these methods have not yet achieved the same level of effectiveness as traditional approaches. This gap has opened opportunities for developing methods that incorporate elements of traditional docking. To the best of our knowledge, the only method integrating the $\mathbb{T}^3 \times SO(3) \times SO(2)^m$ diffusion process with targeted guidance to specific protein surface points is DiffDock-Pocket[1] (Plainer et al., 2023). This method not only adopts a more traditional approach to docking but also extends it by allowing flexibility in side chains.

3 METHODS

Our approach builds on the concept prevalent in traditional docking strategies, where specific geometric shapes like spheres or boxes demarcate potential binding areas, excluding the remainder of the protein structure. We refine this concept in DiffDock by directing incremental updates during the denoising procedure toward a targeted binding region and conformation.

We utilize a guided diffusion strategy, drawing inspiration from Dhariwal & Nichol (2021), by introducing a guiding vector V_{guide} in place of a conventional trained classifier. This vector integrates domain expertise into the diffusion process by altering the update mechanism. The intensity of these alterations is controlled by the hyperparameter γ, allowing for dynamic adjustments based on the proximity and direct route to the target regions. This method ensures that the direction and magnitude of V_{guide} are precisely aligned with the distance and shortest path to the designated areas, optimizing the diffusion trajectory towards the desired binding site.

$$V_{update} = (1 - \gamma)V_{DiffDock} + \gamma V_{guide} \tag{1}$$

---

[1] Referred to as DD-Pocket for the remainder of the manuscript.

where

$$V_{guide} = \alpha \cdot v_{dir} \tag{2}$$

being v_{dir} the vector tangent to the shortest path and α the distance towards the selected center or boundary region. For a detailed analysis of the range of possible γ values, we refer the reader to Appendix D.

Although implementing this method in R^3 would be straightforward, DiffDock operates in a more complex space, $P = \mathbb{T}^3 \times SO(3) \times SO(2)^m$, to accurately reflect the degrees of freedom involved in molecular docking. Consequently, we calculate the shortest paths and distances within each component of this product space, effectively tailoring the diffusion process to this task.

3.1 Boundaries and Geodesics

Shortest paths in $\mathbb{T}^3$ can easily be determined by leveraging straight lines and the geometry of a hypothetical sphere to integrate binding pocket information effectively. This approach helps mitigate the risk of selecting incorrect docking pockets, a prevalent issue in blind docking strategies that often leads to out-of-distribution errors (Ghersi & Sanchez, 2009). The incorporation of this spherical model is instrumental in refining the selection process, enhancing the algorithm's ability to direct the denoising process to given docking sites.

However, the challenge of ensuring the physical realism of docking poses—highlighted by concerns over steric clashes and distorted bond angles (Buttenschoen et al., 2023)—necessitates further refinement of our model. To address these issues, our method extends to the $SO(3)$ and $SO(2)^m$ spaces, where we employ geodesics to define the shortest paths. This approach is critical for accurately guiding our diffusion updates with V_{guide} vectors, ensuring that the docking poses are not only theoretically viable but also physically plausible. For an in-depth explanation of how V_{guide} vectors are defined and integrated into the diffusion process, please refer to Appendix A.

3.2 Soft Constraints

To balance the precision and flexibility of DiffDock, we implement measures to modulate the influence of guidance vectors, recognizing their potential to alter the model's learned dynamics and impact performance within designated spatial regions. We introduce a dynamic adjustment of the guidance vector's influence through a decreasing sigmoid schedule for the γ parameter, progressively diminishing its impact through the diffusion process. This ensures the initial guidance by expert knowledge gradually cedes to the model's intrinsic optimization capabilities.

Moreover, we enable the consideration of multiple potential regions within the $SO(3)$ and $SO(2)^m$ spaces, guiding updates towards the nearest boundary. Once within a targeted region, DiffDock is allowed autonomy in refining rotation and torsion angles, unencumbered by external guidance. This dual-phase optimization strategy—global optimization guided by expert knowledge in early steps, followed by local, autonomous refinement by DiffDock—ensures a balance between adherence to biologically plausible paths and the discovery of optimal docking configurations.

3.3 Evaluation

To facilitate high-throughput evaluation without direct expert input for each sample, we employ a fuzzing strategy, adjusting true labels to simulate expert knowledge. This involves creating guidance spheres for translation with a 7Å radius around the ligand's actual center, directing vectors toward these spheres' peripheries, and ceasing guidance upon entrance. For rotation and torsion, we define regions around true angles, adjusted by a fuzzing factor η of 0.15, guiding vectors towards these adjusted regions and turning off the guidance when they enter. This fuzzing approach, robust across radius and η settings, allows for an effective approximation of expert-directed docking, as detailed in Appendix C. Moreover, we perform an ablation test of different initial gamma values in Appendix D and test the generalization capabilities of our algorithm in Appendix E.

4 EXPERIMENTS

To assess our approach, we employ the testing set from PDBBind proposed in Stärk et al. (2022), applying a three-fold evaluation strategy. Firstly, we analyze docking pose RMSD, focusing on the Top1 and Top5 poses by confidence for both Apo and Holo structures, and introduce Mean Square Error (MSE) as an additional metric to evaluate errors in rotation states and torsion angles, detailed in Appendix B. Secondly, we examine the physical plausibility of these docking poses using the Posebusters suite (Buttenschoen et al., 2023), providing insights into the realism of the predicted conformations. Lastly, we conduct an angle transfer test employing maximal common substructure docking to assess our method's ability to generalize across different molecular configurations.

4.1 EVALUATION OF DOCKING POSES

We evaluated GeoDirDock (GDD) against DiffDock across various setups by comparing the RMSD of generated poses to crystal structures and examining the convergence based on the number of denoising steps taken.

Initially, we tested GDD with translation guidance only (GDD-TR), mimicking conventional docking's focus on specific geometric volumes. Results showed GDD-TR surpassing DiffDock across all metrics, achieving lower RMSD values in fewer steps and maintaining performance across both Apo and Holo receptor configurations, suggesting that targeted guidance enhances docking accuracy and efficiency (Table 1).

Table 1: RMSD performance comparison for Holo and Apo settings across docking methods, showing Top-1 and Top-5 prediction accuracies. Best performances are highlighted in **bold**. Best performances with only translation guidance are stated in *italic*. The number of denoising steps and the number of samples generated are stated as (steps-samples). Results marked with an asterisk (*) were obtained from Plainer et al. (2023)

	Holo				Apo			
	Top-1 RMSD		Top-5 RMSD		Top-1 RMSD		Top-5 RMSD	
	%<2	Med	%<2	Med	%<2	Med	%<2	Med
SMINA* (rigid)	32.5	4.5	46.4	2.2	6.6	7.7	15.7	5.6
GNINA* (rigid)	42.7	2.5	55.3	1.8	9.7	7.5	19.1	5.2
DiffDock (10-10)	34.19	3.53	40.17	2.44	27.14	4.62	36.28	3.09
DiffDock (20-10)	37.08	3.50	44.66	2.60	27.51	4.56	38.40	2.85
DiffDock (20-40)	38.27	3.12	46.65	2.15	27.01	4.85	37.93	3.22
DD-Pocket* (20-10)	47.7	2.1	56.3	1.8	41.0	2.6	47.6	2.2
DD-Pocket* (20-40)	49.8	2.0	59.3	1.7	41.7	2.6	47.8	2.1
GDD-TR (10-10)	41.62	2.69	48.88	2.03	31.71	3.62	42.00	2.50
GDD-TR (20-10)	44.97	2.37	50.56	1.94	34.29	*3.21*	44.00	2.27
GDD-TR (20-40)	*48.88*	*2.05*	*57.82*	*1.70*	*34.38*	3.33	*47.28*	*2.11*
GDD-Full (10-10)	63.97	1.52	67.60	1.31	51.29	1.95	58.74	1.60
GDD-Full (20-10)	68.44	1.24	70.11	1.16	**59.43**	1.60	**65.14**	1.43
GDD-Full (20-40)	**68.72**	**1.22**	**71.51**	**1.14**	58.86	**1.59**	63.71	**1.38**

Upon extending guidance to include torsions and rotations, imposing a more substantial prior than translation alone, we anticipated improvements in the reconstruction of crystallographic structures. GDD-Full, informed across all three spaces, markedly outperformed both GDD-TR and DiffDock, recovering approximately 68% of poses within 2Å of the reference, significantly higher than with translation-only guidance. This comprehensive approach not only improved pose accuracy but also maintained the trend of achieving optimal results in fewer steps, highlighting the efficacy of our method in enhancing docking precision and computational efficiency. Detailed MSE analysis for torsion and rotation are available in Appendix B.

Table 2: Assessment of Posebusters scores for Holo and Apo scenarios, comparing docking structure and re-docking accuracies across DiffDock and GDD methods. Best performances are highlighted in **bold**. The number of denoising steps and the number of samples generated are stated as (steps-samples). Results marked with an asterisk (*) were obtained from Plainer et al. (2023)

	Holo		Apo	
	Docking Structure	Re-docking	Docking Structure	Re-docking
DiffDock (10-10)	23.13	15.31	7.06	4.12
DiffDock (20-10)	25.84	16.11	7.43	4.0
DiffDock (20-40)	26.67	16.0	14.45	4.05
DD-Pocket (20-40)*	29.4	17.4	**21.6**	**10.9**
GDD-TR (10-10)	23.00	15.33	3.43	1.71
GDD-TR (20-10)	20.00	15.33	10.29	6.29
GDD-TR (20-40)	22.00	14.33	9.14	5.14
GDD-Full (10-10)	29.33	24.33	8.62	6.90
GDD-Full (20-10)	28.00	24.00	12.0	10.86
GDD-Full (20-40)	**30.67**	**26.67**	11.43	9.71

4.2 Physical plausibility of informed diffusion poses

The Posebusters evaluation (Table 2) reveals that GDD-Full outperforms GDD-TR and DiffDock in Holo structures, demonstrating its superior accuracy in docking and re-docking tasks. This success highlights GDD-Full's ability to align with the physical aspects of molecular docking, attributed to its comprehensive guidance across multiple spaces.

In contrast, GDD-TR suffers a performance degradation in both Holo and Apo structures. This aligns with our hypothesis of full guidance being needed to reproduce physically plausible poses, as the main objective of translational guidance is the correct selection of the binding pocket, not the increment of physical plausibility of poses.

4.3 Angle transfer as maximal common substructure docking

In the lead optimization phase of drug discovery, template-based modeling is crucial for assessing binding affinity changes among chemically similar compounds (Raman, 2019). This approach involves using a single compound as a reference template to maintain consistent a shared topology while optimizing the distinct elements of each molecule.

To evaluate our method's capabilities beyond self-docking and informed directed docking, we decided to benchmark GDD as a potential tool for template-based modeling. For this we selected crystallized BACE structures from the D3R Grand Challenge 4 (Parks et al., 2020), employing previously identified templates for angle transfer.

We conducted an MCS search to transfer torsion angles associated with heavy atoms common between the template and target molecules, leaving other angles uninformed. This focused evaluation, termed GDD-Tor, exclusively examines the impact of torsion angle transfer, deliberately omitting guidance in other spatial dimensions to isolate the effects of this specific mechanism.

Table 3: Comparison of MCS Docking performance, showing RMSD and MSE metrics for Top-1 and Top-5 predictions between DiffDock and GDD-Tor methods. The number of denoising steps and the number of samples generated are stated as (steps-samples).

	Top-1 RMSD		Top-5 RMSD		Top-1 MSE		Top-5 MSE	
	%<2	Med	%<2	Med	Avg	Med	Avg	Med
DiffDock (20-40)	50.0	1.93	55.0	1.71	2.33	2.30	1.37	1.45
GDD-Tor (20-40)	63.2	1.76	63.2	1.69	1.79	1.35	1.31	1.21

As we can determine by the results in Table 3, this angle transfer already marks an improvement both in RMSD and torsion MSE. These preliminary results outline great potential for our method and we look to further develop this application and corresponding benchmarks.

5   Conclusions and Future Work

In this study, we introduce GeoDirDock, a novel framework that enhances generative docking models through geodesic guidance. GeoDirDock significantly improves docking precision and boosts the physical plausibility of results compared to traditional blind docking methods like DiffDock. By incorporating guidance across translation, torsion, and rotation dimensions, GeoDirDock outperforms both DiffDock and its translation-only variant, achieving precise docking poses with fewer computational steps. These results pose guided docking as a promising avenue for AI-enabled molecular docking, ensuring both efficiency and accuracy in molecular docking.

Future work would focus on improving in generalizability to completely unseen ligand and protein chemistries, thus, postulating GeoDirDock as a valuable tool for prospective docking campaigns. We are also keen on extending this algorithm towards including protein flexibility in the docking procedure, aiming for a more realistic setting of the docking protocol. The inclusion of more realistic priors, in both backbone and side chain angles, would solve possible steric clashes between ligands and proteins, performing implicit induced fit docking.

References

Martin Buttenschoen, Garrett M. Morris, and Charlotte M. Deane. Posebusters: Ai-based docking methods fail to generate physically valid poses or generalise to novel sequences. *arXiv preprint arXiv:2308.05777*, 2023.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and T. Jaakkola. Diffdock: Diffusion steps, twists, and turns for molecular docking. In *The Eleventh International Conference on Learning Representations*, 2023.

Gabriele Corso, Arthur Deng, Nicholas Polizzi, Regina Barzilay, and Tommi Jaakkola. Deep confident steps to new pockets: Strategies for docking generalization. In *International Conference on Learning Representations (ICLR)*, 2024.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. In *Advances in Neural Information Processing Systems*, volume 34, pp. 8780–8794. Curran Associates, Inc., 2021.

Richard A. Friesner, Jay L. Banks, Robert B. Murphy, Thomas A. Halgren, Jasna J. Klicic, Daniel T. Mainz, Matthew P. Repasky, Eric H. Knoll, Mee Shelley, Jason K. Perry, David E. Shaw, Perry Francis, and Peter S. Shenkin. Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7):1739–1749, 2004.

Dario Ghersi and Roberto Sanchez. Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites. *Proteins*, 74(2):417–424, February 2009.

David Ryan Koes, Matthew P Baumgartner, and Carlos J Camacho. Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. *Journal of chemical information and modeling*, 53(8):1893–1904, 2013.

Andrew T McNutt, Paul Francoeur, Rishal Aggarwal, Tomohide Masuda, Rocco Meli, Matthew Ragoza, Jocelyn Sunseri, and David Ryan Koes. Gnina 1.0: molecular docking with deep learning. *Journal of cheminformatics*, 13(1):43, 2021.

Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei, and Meng Cui. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.*, 7(2):146–157, June 2011.

Shuya Nakata, Yoshiharu Mori, and Shigenori Tanaka. End-to-end protein–ligand complex structure generation with diffusion-based generative models. *BMC bioinformatics*, 24(1): 1–18, 2023.

Conor D. Parks, Zied Gaieb, Michael Chiu, Huanwang Yang, Chenghua Shao, W. Patrick Walters, Johanna M. Jansen, Georgia McGaughey, Richard A. Lewis, Scott D. Bembenek, Michael K. Ameriks, Tara Mirzadegan, Stephen K. Burley, Rommie E. Amaro, and Michael K. Gilson. D3r grand challenge 4: blind prediction of protein–ligand poses, affinity rankings, and relative binding free energies. *Journal of Computer-Aided Molecular Design*, 34(2):99–119, Feb 2020.

Michael Plainer, Marcella Toth, Simon Dobers, Hannes Stärk, Gabriele Corso, Céline Marquet, and Regina Barzilay. Diffdock-pocket: Diffusion for pocket-level docking with sidechain flexibility. In *NeurIPS 2023 Workshop on New Frontiers of AI for Drug Discovery and Development*, 2023.

Zhuoran Qiao, Weili Nie, Arash Vahdat, Thomas F Miller III, and Anima Anandkumar. Dynamic-backbone protein-ligand structure prediction with multiscale generative diffusion models. *arXiv preprint arXiv:2209.15171*, 2022.

E. Prabhu Raman. Template-based method for conformation generation and scoring for congeneric series of ligands. *Journal of Chemical Information and Modeling*, 59(6):2690–2701, 2019.

Sergio Ruiz-Carmona, Daniel Alvarez-Garcia, Nicolas Foloppe, A Beatriz Garmendia-Doval, Szilveszter Juhos, Peter Schmidtke, Xavier Barril, Roderick E Hubbard, and S David Morley. rdock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput. Biol.*, 10(4):e1003571, April 2014.

Hannes Stärk, Octavian Ganea, Lagnajit Pattanaik, Regina Barzilay, and Tommi Jaakkola. Equibind: Geometric deep learning for drug binding structure prediction. In *International conference on machine learning*, pp. 20503–20521. PMLR, 2022.

Oleg Trott and Arthur J. Olson. Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2):455–461, 2010.

Yuejiang Yu, Shuqi Lu, Zhifeng Gao, Hang Zheng, and Guolin Ke. Do deep learning models really outperform traditional approaches in molecular docking? In *The Eleventh International Conference on Learning Representations Workshop on Machine Learning for Drug Discovery (MLDD)*, 2023.

A ALGORITHMS AND COMPUTATIONS

A.1 GEODESICS IN $\mathbb{T}^3$

Geodesics in $\mathbb{T}^3$ can be represented as straight lines in R^3. We parameterize a binding sphere B with a center ϵ and a radius r and p_t is defined as the ligand's center of mass. Then, the direction vector of translation guidance can be computed as:

$$v_{dir}^{tr} = \begin{cases} 0 & \text{if } p_t \in B(\epsilon; r) \\ \frac{\epsilon - p_t}{\|\epsilon - p_t\|_2} & \text{if not} \end{cases} \tag{3}$$

We choose the scaling factor α^{tr}, due to empirical performance, to be:

$$\alpha^{tr} = \|\epsilon - p_t\|_2 \tag{4}$$

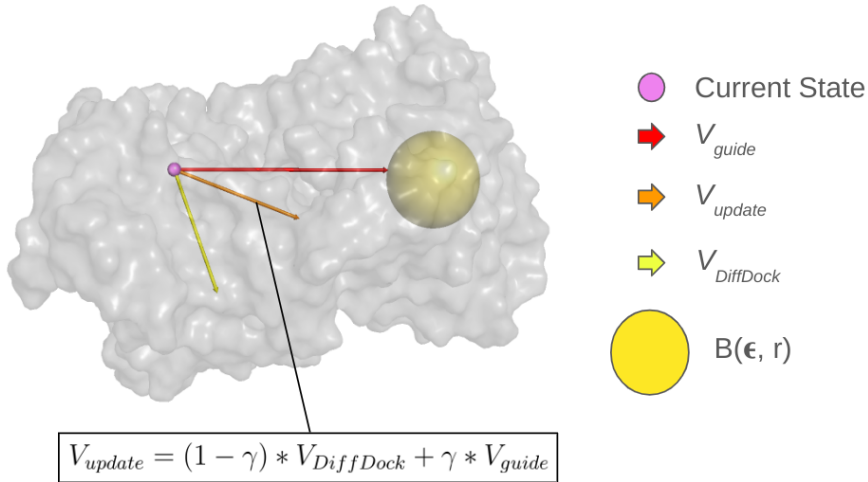$$V_{update} = (1 - \gamma) * V_{DiffDock} + \gamma * V_{guide}$$

Figure 1: Guidance in $\mathbb{T}^3$ for a sample sphere in *1a0q*.

A.2 GEODESICS IN $SO(3)$

Rigid rotations of the ligand are defined as rotations around its center of mass , corresponding to the its orientation within the pocket (Corso et al., 2023). As this is extremely complicated to define a-priori by an expert, we choose to cast this problem onto finding the optimal rotation states of said center with respect to its atoms. For this, we utilize the axis formed between the center of masses and the first atom as our prior, and to account for possible reflections of the plane formed by this vector and the DiffDock update, we will also include the axis formed by the orthogonal vector of said axis and the one formed between the center of mass and the second atom 2. Those two vectors will be the axes of rotation for which we aim to optimize the rotation states.

A unit 2-sphere embedded in R^3 can be denoted as $S^2 \subset R^3$ and in this sphere, we can encode the rotation states of our target using polar coordinates:

$$r_t = \phi \in [0, 2\pi), \theta \in [0, \pi) \tag{5}$$

We can therefore define the selected regions of interest for the rotation, Ω, as:

$$\Omega = (\bigsqcup_{i=1}^{p} [\phi_1^i, \phi_2^i]) \times (\bigsqcup_{i=1}^{th} [\theta_1^i, \theta_2^i]) \tag{6}$$
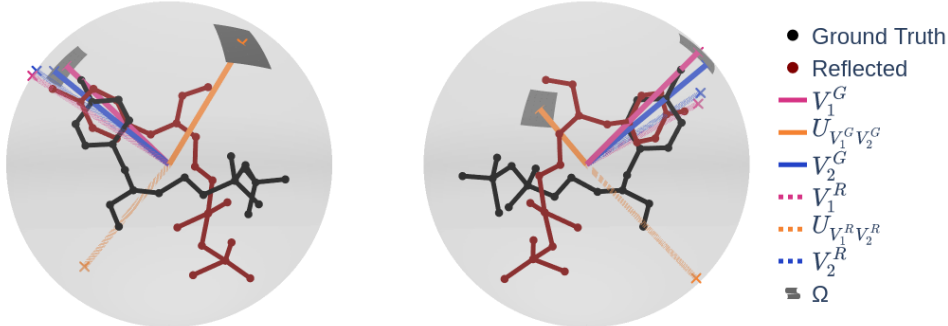
Figure 2: Rotation state vectors in S^2 for a reflected molecule versus its ground truth.

For the sake of simplicity, although our implementation supports this more general definition with multiple regions, we will define Ω as a unique region for each of the axes:

$$\Omega = [\phi_1, \phi_2] \times [\theta_1, \theta_2] \tag{7}$$

We parameterize both rotations as points on this unit sphere and because we need to compute the guiding vector with the condition of being tangent to the geodesic, we map our points as vectors $v \in R^3$, according to their position in the sphere, as:

$$v = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta) \tag{8}$$

For the corresponding ϕ and θ of each point, obtaining v_α and v_β.

We are interested in both of the geodesics that cover the shortest distance between each pair of two points. Geodesics on the sphere can be defined as the great circle centered in the origin, with the plane containing it having the orthogonal basis v_α and u, defined as:

$$g(t) = \cos(t)v_\alpha + \sin(t)u \tag{9}$$

With u defined as:

$$u = \frac{v_\beta - (v_\beta^t \cdot v_\alpha)v_\alpha}{\|v_\beta - (v_\beta^t \cdot v_\alpha)v_\alpha\|} \tag{10}$$

Point β can then be reached with:

$$t^* = \arccos(v_\beta^t \cdot v_\alpha) \tag{11}$$

With one t^* defined for each selected axis, as t_1^* and t_2^*.

We cannot directly add these vectors to the DiffDock update, v_{r_t}, so we choose to sequentially apply the rotation updates one by one.

$$v_{dir}^{rot} = \begin{cases} 0 & \text{if } r_t \in [\phi_1, \phi_2] \times [\theta_1, \theta_2] \\ t_2^* \cdot ((t_1^* \cdot (v_{r_t} \times u_1)) \times u_2) & \text{if not} \end{cases} \tag{12}$$

The scaling factor was set as 1, as we found poor empirical performance when varying it.
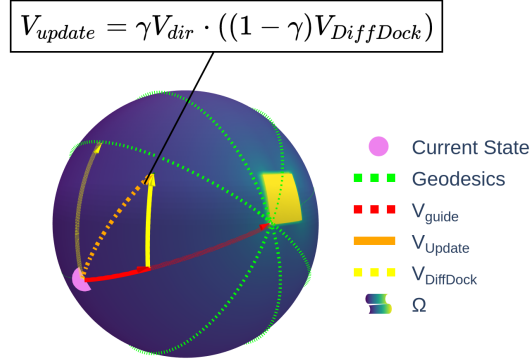
$$\alpha^{rot} = 1 \tag{13}$$

$$\boxed{V_{update} = \gamma V_{dir} \cdot ((1 - \gamma)V_{DiffDock})}$$

Figure 3: $SO(3)$ guidance in S^2 for a sample rotation state.

A.3   GEODESICS IN $SO(2)^m$

As mentioned in Corso et al. (2023), $SO(2)^m$ is diffeomorphic to the hypertorus $\mathcal{T}^m$, therefore the guidance calculations will be performed in this m-dimensional hypertorus. Recalling the definition of the torus being the product space of two 1-spheres or circles:

$$\mathcal{T}^2 = S^1 \times S^1 \tag{14}$$

and one-dimensional torus being equal to a circle $\mathcal{T}^1 = S^1$, we can define the hypertorus $\mathcal{T}^m$ as the product space of m 1-spheres:

$$\mathcal{T}^m = \underbrace{\mathcal{S}^1 \times \ldots \times \mathcal{S}^1}_{m} \tag{15}$$

With $\mathcal{T}^m$ being defined as the following quotient space:

$$\mathcal{T}^m \cong \mathcal{R}^m / \mathcal{Z}^m \cong \mathcal{R}^m / 2\pi\mathcal{Z}^m \tag{16}$$

With this information, we can determine geodesics in the hypertorus to be straight lines in the hypercube that reappear in the diametrically opposite point when they collide with a boundary. Leveraging this knowledge, we now can compute the desired shortest path geodesic, its length and the tangent vector.

Given two points $z_1, z_2 \in \mathcal{T}^m$, let $\mathcal{G}(z_1, z_2)$ be the set of all possible geodesics that join z_1 and z_2. We are interested in finding the shortest-path geodesic between z_1 and z_2:

$$\eth_{z_1, z_2} = argmin_{g \in \mathcal{G}(z_1, z_2)} l(g(z_1, z_2)) \tag{17}$$
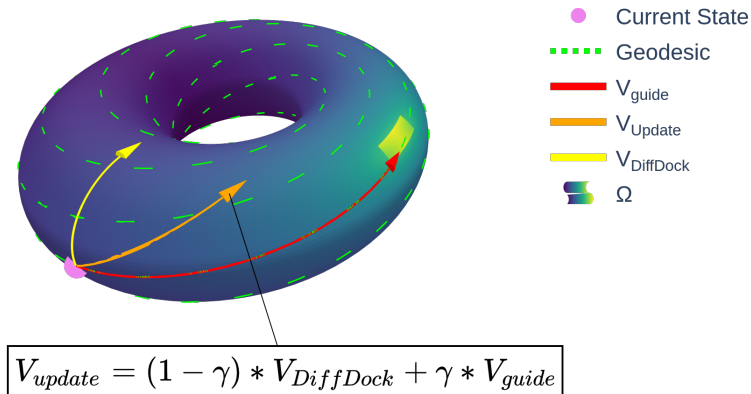
Figure 4: Guidance in $SO(2)^m$ for a 2 torsion angle compound.

With its length being defined as $l(\eth_{z_1,z_2})$, and the tangent vector being defined as $v_{dir}(\eth_{z_1,z_2})$. We will define the selected regions, Ω as:

$$\Omega = (\bigsqcup_{k=1}^{r_1}[\theta_{2k-1}^1, \theta_{2k}^1]) \times \ldots \times (\bigsqcup_{k=1}^{r_m}[\theta_{2k-1}^m, \theta_{2k}^m]). \tag{18}$$

Being each set of $\theta_{2k-1}, \theta_{2k}$ the delimiters of the region in each of the 1-spheres composing the torus.

Although we build our algorithms to be able to handle multiple regions per angle, we will omit this from the algorithms for simplicity. Ω is therefore defined as the unique region of interest in the hypertorus. As expressed in 3.2, we need to determine if our current state is within the bounds of the selected region to turn off the guidance:

Algorithm 1 Is $z \in \mathcal{T}^k$ in $Proj_{\mathcal{T}^k}\Omega$?

Input: $0 < n \le m, \{j_1, \ldots j_n\} \subset \{1, \ldots, m\}, z \in \mathcal{T}^n, \Omega$
Output: True or False
$count \leftarrow 1$
for i in $\{j_1, \ldots, j_n\}$ **do**
    if $\theta_2^i \ge \theta_1^i$ **then**
        if not $\theta_1^i \le z_{count}$ or not $z_{count} \le \theta_2^i$ **then**
            return False
        end if
    else
        if not $z_{count} \ge \theta_1^i$ and not $z_{count} \le \theta_2^i$ **then**
            return False
        end if
    end if
    $count \leftarrow count + 1$
end for
return True

We will also need to compute the vector tangent to the geodesic between the points and the length of said geodesic:

Algorithm 2 Computation of tangent vector $v_{dir}(\eth_{z_1, z_2})$ and $l(\eth_{z_1, z_2})$

Input: $m > 0, z_1, z_2 \in \mathcal{T}^m$
Output: $l(\eth_{z_1, z_2}) \in \mathcal{R}_{\ge 0}$, $v_{dir}(\eth_{z_1, z_2}) \in \partial B(0; 1) \subset \mathcal{R}^m$
$l \leftarrow 0$
for i in $\{1, \ldots, m\}$ **do**
    if $z_2^i \ge z_1^i$ **then**
        if $z_2^i - z_1^i \le z_1^i + (2\pi - z_2^i)$ **then**
            $l \leftarrow l + (z_2^i - z_1^i)^2$
            $v_{dir}^i \leftarrow z_2^i - z_1^i$
        else
            $l \leftarrow l + (z_1^i + (2\pi - z_2^i))^2$
            $v_{dir}^i \leftarrow z_2^i - (2\pi + z_1^i)$
        end if
    else
        if $z_1^i - z_2^i \le (2\pi - z_1^i) + z_2^i$ **then**
            $l \leftarrow l + (z_1^i - z_2^i)^2$
            $v_{dir}^i = z_2^i - z_1^i$
        else
            $l \leftarrow l + ((2\pi - z_1^i) + z_2^i)^2$
            $v_{dir}^i = (z_2^i + 2\pi) - z_1^i$
        end if
    end if
end for
$l \leftarrow \sqrt{l}$
$v_{dir} \leftarrow \frac{1}{l}(v_{dir}^1, \ldots, v_{dir}^m)$

In this computation, we have to take into account which of the boundaries is closer to our current state, and which direction is closer in the torus geometry. Therefore our final boundary vector τ can be computed as:

Algorithm 3 $argmin_{\tau \in \partial\Omega} \min l(\eth_{x_t, \tau_t})$

    Input: $m \geq 0, x_t, \Omega$
    Output: τ
  for i in $\{1, \ldots, m\}$ **do**
    if $x_t^i \in [\theta_1^i, \theta_2^i]$ **then** $\triangleright$ $x_t^i \in [\theta_1^i, \theta_2^i] = \text{Alg1}(n = 1, m, \{i\}, x_t^i, \Omega)$
        $\tau^i \leftarrow x_t^i$
    else
        if $\min l(\eth_{x_t^i, \theta_1^i}) \leq \min l(\eth_{x_t^i, \theta_2^i})$ **then**
            $\tau^i \leftarrow \theta_1^i$
        else
            $\tau^i \leftarrow \theta_2^i$
        end if $\triangleright$ $l(\eth_{x_t^i, \theta_k^i})$ from $\text{Alg2}(1, x_t^i, \theta_k^i)$
    end if
  end for
  $\tau \leftarrow (\tau^1, \ldots, \tau^m)$

This vector indicates the closest boundaries of our selected region to our current state, and our subsequent guiding vector can be computed using Algorithm 2.

$$v_{dir}^{tor} = \begin{cases} 0 & \text{if } x_t \in \Omega \\ v_{dir}(\eth_{x_t, \tau}) & \text{if not} \end{cases} \tag{19}$$

We choose the α scaling factor to be the length of the shortest path geodesic, also computed with Algorithm 2:

$$\alpha^{tor} = l(\eth_{x_t, \tau}) \tag{20}$$

B EVALUATION VIA ROTATIONAL AND TORSIONAL MSE

In this section, we introduce the usage of Mean Square Error in torsion and rotation as a measure that decouples distance to the binding site and bond and torsion errors.

We believe this separation to be beneficial in a high granularity analysis, as errors in binding pocket detection can easily affect the ligand RMSD measurements while hiding a correct assignment of torsion angles. So, we further analyze the previously obtained poses from Section 4.1 in rotation and torsion MSE.

Table 4: Comparison of MSE in rotation and torsion across docking methods.

	Holo				Apo			
	Top-1 MSE		Top-5 MSE		Top-1 MSE		Top-5 MSE	
	Rot	Tor	Rot	Tor	Rot	Tor	Rot	Tor
DiffDock (10-10)	1.82	2.35	0.32	1.43	1.77	2.37	0.40	1.30
DiffDock (20-10)	1.66	2.22	0.26	*1.25*	2.13	2.38	0.47	1.33
DiffDock (20-40)	1.58	2.17	*0.20*	1.40	*1.32*	2.31	0.40	1.40
GDD-TR (10-10)	1.64	2.28	0.27	1.30	1.82	2.48	0.35	1.37
GDD-TR (20-10)	*1.38*	2.28	0.22	1.27	1.97	*2.29*	*0.24*	*1.32*
GDD-TR (20-40)	1.58	*2.17*	*0.20*	1.40	1.33	2.39	0.26	1.40
GDD-Full (10-10)	0.04	0.08	0.02	0.05	0.05	0.08	0.02	0.05
GDD-Full (20-10)	**0.02**	**0.03**	**0.01**	**0.02**	**0.03**	**0.03**	**0.01**	**0.02**
GDD-Full (20-40)	**0.02**	**0.03**	**0.01**	**0.02**	**0.03**	**0.03**	**0.01**	**0.02**

In these results we can clearly determine a positive effect of the full guide on both Holo and Apo structures, supporting its hypothesized usage for correcting bond and rotation state errors. Furthermore, we also detect mild improvements with only translational guides, which leads us to believe that faster convergence upon correct binding sites allows DiffDock to focus in better bond and rotation states, increasing RMSD performance.

C   Region Fuzzing Ablation

In this section we seek to evaluate the effect of altering the η parameter and radius parameter of our algorithm, weakening the prior that we impose upon the translation, rotation and torsion states. We perform this ablation test in a smaller set of the PDBBind testing dataset, a set of 30 complexes randomly sampled.

In each of the sections, DiffDock will be compared against GDD-full with variations in the fuzzing hyperparameter corresponding to the respective space where the ablation is performed, maintaining default values for the rest of the spaces. The comparison will be performed as the mean and median step-wise error of the 40 samples and 20 steps protocol for both algorithms.

C.1   Translation

For translation we will consider DiffDock base against GDD full with spheres of radius 5,7,10, and 12 Å, 7 Å being the default of our algorithm.
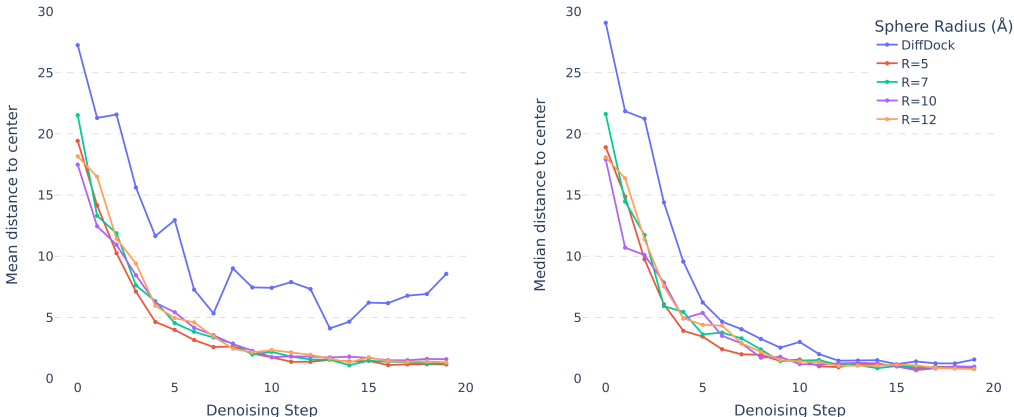

Figure 5: Mean and median distance to center of sphere versus denoising step.

As we are able to determine from Figure 5, our algorithm is very robust to the radius parameter, with a rough estimate of a binding pocket being enough for discarding outlier values that heavily affect the mean distance to the correct binding site.

We are also able to observe a faster convergence across the denoising steps, highlighting the potential of our method for reducing the computational cost of DiffDock in high-throughput virtual screening campaigns.

C.2   Rotation

In rotation, we set the same default value for both values of η for angles ϕ and θ at 0.15. We further evaluated values of 0.3, 0.45 and 0.6 for both η.
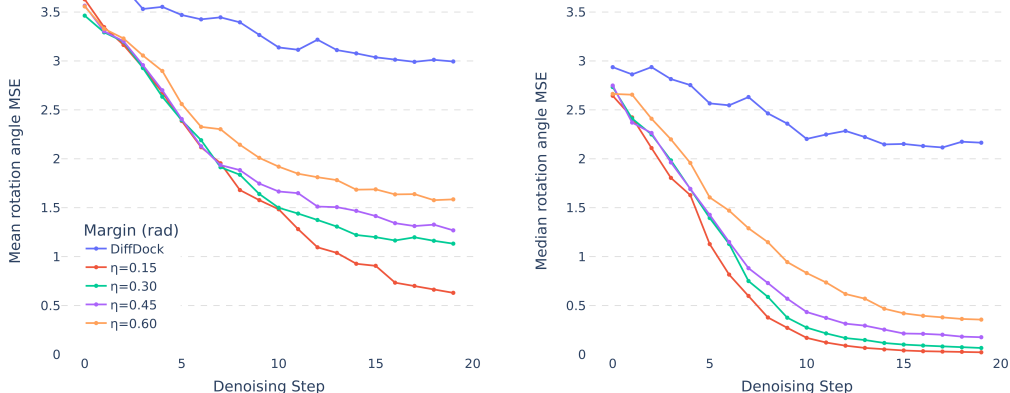
Figure 6: Mean and median MSE of rotation states versus denoising step.

In this ablation we are able to determine an important reduction in rotation MSE, both in mean and median, with high robustness to the hyperparameter η (measured in radians).

We are able to see a clear degradation in performance with an increase in the region, however, even the worst-performing ablation has a much better performance than DiffDock in rotation MSE.

C.3 TORSION

In torsion guidance we also set the same default η value for all the θ angles at 0.15. We evaluated values of 0.3, 0.45 and 0.6 for all the angles.
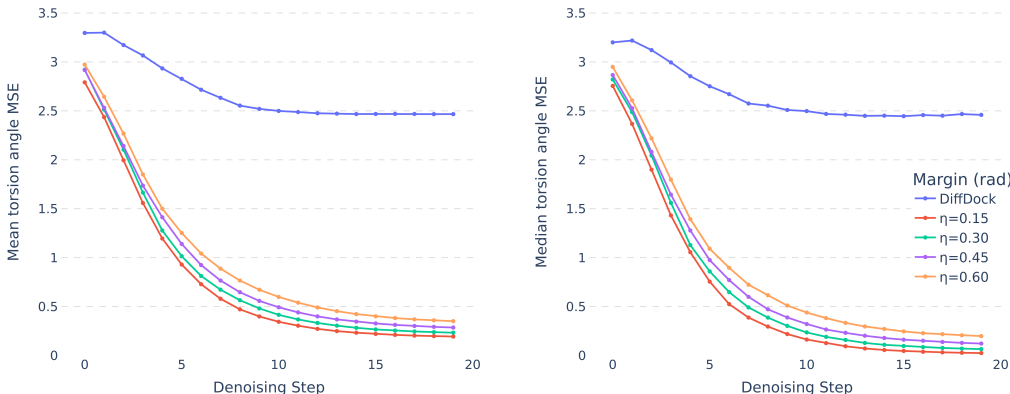
Figure 7: Mean and median MSE of torsion angles versus denoising step.

This ablation follows the trend set in the rotation analysis, with a clear increase in performance for MSE even with higher values of η, alongside an accelerated convergence rate.

We are also able to determine this η to be more robust than in rotation, with increasingly fuzzed regions performing very consistently with our base algorithm, and much better than

DiffDock. This robustness analysis allowed us to begin extending this algorithm from self-docking to the MCS domain in 4.3.

D   Guidance Strength Ablation

In this section we evaluate the effect of altering the initial γ hyperparameter, corresponding to the strength of the guidance that we impose upon the translation, rotation and torsion states. Lower values of gamma impose a weaker expert prior in the diffusion process and vice-versa.

We test possible values in the range 0 to 0.5 with 0 being DiffDock and 0.2 being our default value for GDD. Following Appendix C, the comparison will be performed as the mean and median step-wise error of the 40 samples and 20 steps protocol, with gamma values of different subspaces varying at the same time.
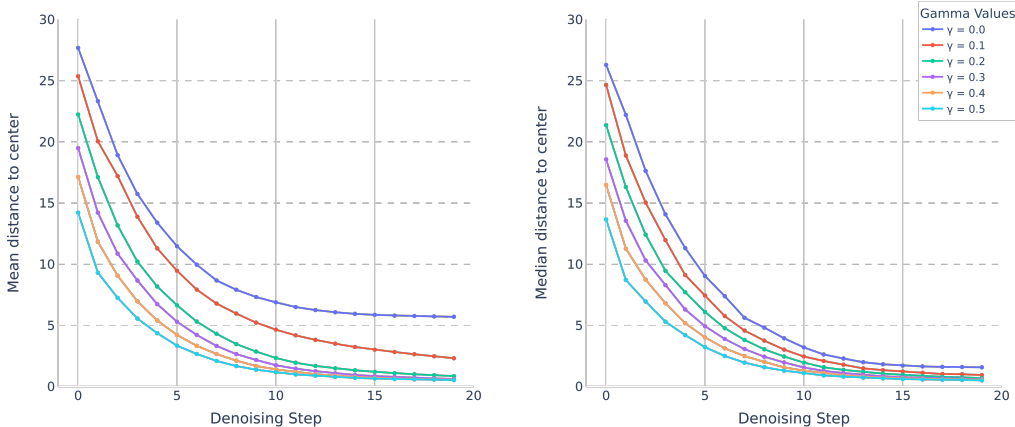
D.1   Translation

Figure 8: Mean and median distance to center of sphere versus denoising step.

As we are able to determine from Figure 8, an increase of the guidance strength allows for a faster convergence in steps compared to DiffDock. At the same time, a mild guidance of 0.1 already achieves the goal of reducing the amount of outliers in distance to the true center, as determined by the reduction in mean distance, with 0.2 achieving results very close to the ground truth. The rest of the range of values, however, do not significantly improve the previous result.
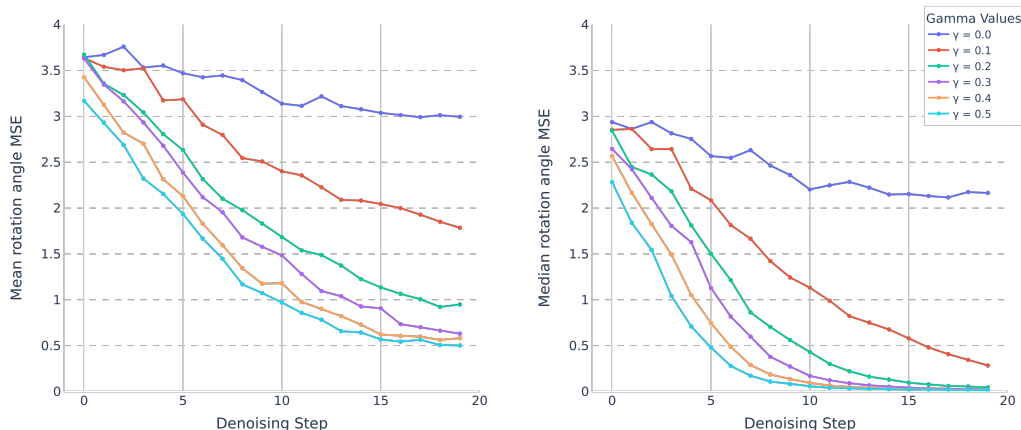
D.2 ROTATION

Figure 9: Mean and median MSE of rotation states versus denoising step.

In rotation, we are able to observe a very similar trend (Figure 9) to translation, with gamma values above 0.3 providing mostly faster convergence rather than an increase of performance, and with mild guidance values already providing an increase in performance compared to DiffDock.
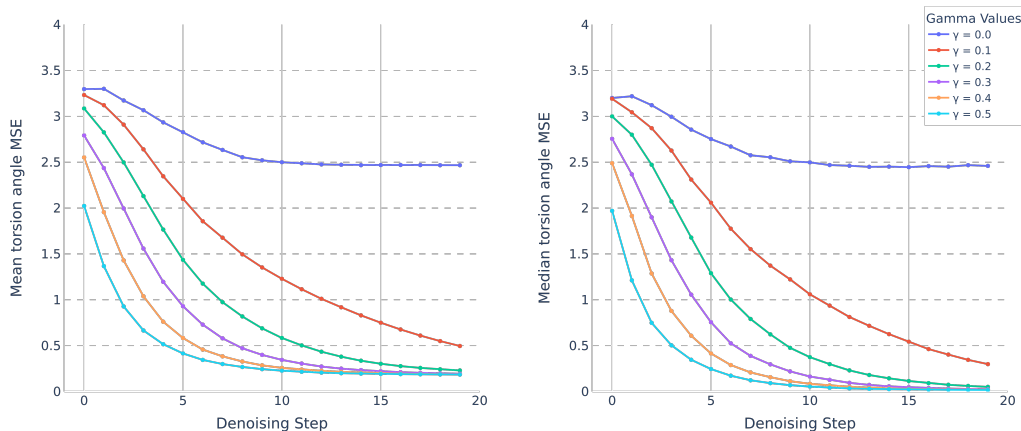
D.3 TORSION

Figure 10: Mean and median MSE of torsion angles versus denoising step.

Lastly, a very similar trend can be observed in torsion (Figure 10) with values over 0.2 providing only convergence benefits, rather than performance increase.

Overall, GeoDirDock maintains a robust performance across gamma values, with lower values of guidance already being enough for a increase of performance and higher values leading to equivalent performance and accelerated convergence.

E   Generalization Across Protein Domains

Generalization across different parts of the proteome is a desired property for deep learning docking methods that has been questioned in Corso et al. (2024). In this work, the authors develop a novel benchmark, DockGen, which contains a more chemically diverse set of binding pockets, compared to the ones contained in the original PDBBind test dataset.

Table 5: RMSD and MSE performance comparison for the DockGen benchmark, showing Top-1 and Top-5 prediction accuracies. Best performances are highlighted in **bold**. Best performances with only translation guidance are stated in *italic*. The number of denoising steps and the number of samples generated are stated as (steps-samples).

| | Top-1 RMSD | | Top-5 RMSD | | Top-1 MSE | | Top-5 MSE | |
	%<2	Med	%<2	Med	Rot	Tor	Rot	Tor
DiffDock (10-10)	2.27	9.48	7.39	5.82	3.70	2.56	1.13	1.39
DiffDock (20-10)	4.49	10.02	8.43	5.79	3.14	2.69	1.13	1.32
DiffDock (20-40)	3.93	9.61	8.99	6.08	3.20	2.69	1.02	1.42
GDD-TR (10-10)	5.06	6.62	6.18	5.38	3.35	2.82	1.10	1.57
GDD-TR (20-10)	4.52	*6.23*	8.47	4.82	*2.80*	*2.66*	0.93	1.41
GDD-TR (20-40)	*5.62*	6.29	*8.99*	*4.63*	2.83	2.72	*0.89*	*1.58*
GDD-Full (10-10)	10.11	5.81	13.48	4.91	0.83	0.61	0.12	0.28
GDD-Full (20-10)	16.29	4.91	20.22	4.27	**0.17**	**0.10**	0.05	**0.06**
GDD-Full (20-40)	**17.98**	**4.75**	**24.16**	**3.87**	0.20	**0.10**	**0.04**	**0.06**

In these results we are able to determine an increase in performance of both GDD-TR and GDD-Full, following the trend of section 4.1. This highlights the capabilities of our method to generalize across a diverse set of binding modes.