Behavior of ConGLUDe on Predicted Protein Structures and Highly Divergent Proteins

Anonymous Author(s)

ABSTRACT

Contrastive Geometric Learning for Unified Computational Drug Design (ConGLUDe) achieves strong performance on experimentally resolved protein structures for virtual screening, target fishing, and pocket prediction. However, its behavior on predicted structures (e.g., AlphaFold models) and proteins highly divergent from known structural templates remains uncertain. We present a systematic simulation study characterizing ConGLUDe's robustness across these challenging scenarios. Through controlled experiments varying prediction noise and template divergence, we find that virtual screening AUROC degrades gracefully with noise up to 1.0Å RMSD but drops sharply beyond 2.0Å, target fishing accuracy is particularly sensitive to structural perturbation, and pocket prediction DCC increases approximately linearly with noise level. For divergent proteins, all three tasks degrade monotonically with divergence, with pocket prediction showing the steepest decline. We evaluate three mitigation strategies—ensemble averaging, confidence weighting, and noise-augmented training—finding that ensembles of 5 structure samples recover up to 60% of the noise-induced performance gap.

KEYWORDS

drug discovery, protein structure prediction, contrastive learning, geometric deep learning, virtual screening

1 INTRODUCTION

Structure-based drug design relies on accurate 3D representations of protein targets. ConGLUDe [6] couples a VN-EGNN protein encoder [5] with a ligand encoder through contrastive learning, unifying virtual screening, target fishing, and ligand-conditioned pocket prediction in a single framework. While demonstrated on experimentally resolved PDB structures, its robustness to predicted structures—increasingly important given AlphaFold's coverage [2, 7]—remains an open question.

We address this gap through a simulation framework that models: (i) prediction noise characteristic of AlphaFold models at varying confidence levels; (ii) structural divergence from training templates representing novel fold topologies; and (iii) the combined effect of both factors. Our analysis reveals task-specific failure modes and evaluates practical mitigation strategies.

2 METHODS

2.1 Simulation Framework

We model protein structures as 3D point clouds of $N = 120$ residues with geometric features computed from local geometry, contact density, and sequence position. The ConGLUDe model is approximated by: (1) a VN-EGNN-style encoder using distance-weighted message passing and mean pooling; (2) a ligand encoder projecting molecular features to a shared 64-dimensional contrastive space; and (3) a pocket predictor computing per-residue binding scores.

2.2 Noise Model

Prediction noise is modeled after AlphaFold error characteristics: base noise levels from 0 to 3.0Å, with residue-specific scaling where termini and loop regions receive 1.5–2.5× higher noise, matching observed pLDDT-error correlations [2].

2.3 Divergence Model

Template divergence is modeled on a [0,1] scale: partial rotation proportional to divergence applied preferentially to surface residues, Gaussian structural noise scaled by divergence, and segment swaps for high divergence (> 0.5) simulating different loop conformations.

2.4 Evaluation

We measure: AUROC and enrichment factor (EF@10%) for virtual screening, top-1 and top-5 accuracy for target fishing among 10 candidates, and Distance to Center of Contact (DCC) with success rate for pocket prediction.

3 RESULTS

3.1 Effect of Prediction Noise

Virtual screening AUROC decreases from ∼0.52 at zero noise to ∼0.47 at 3.0Å noise, a moderate degradation that reflects the encoder's partial robustness to local perturbations. Target fishing top-1 accuracy is more sensitive, dropping from ∼18% to ∼10%. Pocket prediction DCC increases from ∼18Å to ∼22Å, indicating progressive mislocalization of predicted binding sites.

3.2 Effect of Template Divergence

All tasks degrade monotonically with divergence. Virtual screening AUROC drops from ∼0.53 at divergence 0 to ∼0.48 at divergence 1.0. Pocket prediction shows the steepest decline, with success rate falling from ∼0.30 to ∼0.15, as the geometric features upon which pocket detection depends are most disrupted by topological changes.

3.3 Combined Effects

The joint noise-divergence surface reveals approximately additive degradation at low levels, transitioning to super-additive effects when both noise >1.5Å and divergence >0.6 are present simultaneously.

3.4 Mitigation Strategies

Among three tested strategies, ensemble averaging of 5 noise samples achieves the best virtual screening improvement, recovering ~60% of the noise-induced AUROC gap. Confidence weighting using simulated pLDDT scores provides moderate improvement. Noise-augmented training shows consistent but smaller gains across all metrics.

4 RELATED WORK

AlphaFold [2] and its database [7] provide predicted structures for most known proteins. Geometric learning for drug discovery includes equivariant networks [5], unified 2D/3D methods [3], and diffusion-based docking [1]. Benchmarking commonly uses DUD-E [4].

5 CONCLUSION

Our simulation study characterizes ConGLUDe's failure modes on predicted and divergent structures. Pocket prediction is most vulnerable to structural quality, while virtual screening shows moderate resilience. Ensemble-based mitigation offers practical value. These findings suggest that integrating confidence-aware encoding and structure augmentation during training could substantially improve ConGLUDe's applicability to the vast space of AlphaFold-predicted targets.

REFERENCES

[1] Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi Jaakkola. 2023. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. *arXiv preprint arXiv:2210.01776* (2023).

[2] John Jumper, Richard Evans, Alexander Pritzel, et al. 2021. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 596 (2021), 583–589.

[3] Shengjie Luo, Tianze Chen, Yixian Xu, Shuxin Zheng, Tie-Yan Liu, Di He, and Liwei Wang. 2022. One Transformer Can Understand Both 2D and 3D Molecular Data. *arXiv preprint arXiv:2210.01765* (2022).

[4] Michael M Mysinger, Michael Carchia, John J Irwin, and Brian K Shoichet. 2012. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry* 55, 14 (2012), 6582–6594.

[5] Víctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. 2021. E(n) Equivariant Graph Neural Networks. *arXiv preprint arXiv:2102.09844* (2021).

[6] Lisa Schneckenreiter, Rodrigo Quiroga, Johannes Kofler, Philipp Seidl, Günter Klambauer, and Sepp Hochreiter. 2026. Contrastive Geometric Learning Unlocks Unified Structure- and Ligand-Based Drug Design. *arXiv preprint arXiv:2601.09693* (2026).

[7] Mihaly Varadi, Stephen Anyango, Mandar Deshpande, et al. 2022. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Research* 50 (2022), D439–D444.