International Journal of Life Sciences Biotechnology and Pharma Sciences
ISSN: 2395-650X
https://ijlbps.net/index.php/ijlbps/issue/archive

AI-Assisted Molecular Docking for Novel Drug Discovery

Hugo Garcia[1], Erik Popescu[2], Erik Silva[3]

[1] Senior Lecturer, Institute of Intelligent Systems, Nordic Technical University, Stockholm, Sweden. Email: hugo.garcia668@ai-europe-research.org | ORCID: 9039-1283-8127-1479
[2] Assistant Professor, Institute of Intelligent Systems, Nordic Technical University, Stockholm, Sweden. Email: erik.popescu139@ai-europe-research.org | ORCID: 8848-0196-2027-1155
[3] Assistant Professor, Department of Machine Learning, Nordic Technical University, Stockholm, Sweden. Email: erik.silva446@ai-europe-research.org | ORCID: 4764-0450-9980-7082

ABSTRACT

Molecular docking--the computational prediction of small molecule binding poses and affinities at protein active sites--is a cornerstone of structure-based drug discovery, yet classical docking methods suffer from limitations in scoring function accuracy, conformational sampling completeness, and computational throughput that constrain their utility for large-scale virtual screening of multi-billion-compound chemical libraries. Artificial intelligence has transformed molecular docking through deep learning scoring functions, generative molecular design, and end-to-end docking architectures that directly predict binding poses without explicit energy minimisation. This study develops and benchmarks an AI-assisted drug discovery pipeline integrating four AI components: (i) SE(3)-equivariant GNN-based binding site prediction (SiteNet); (ii) DiffDock-v2 blind docking with improved confidence calibration; (iii) a graph transformer-based scoring function (GraphDTA-v2) fine-tuned on the PDBbind 2024 dataset; and (iv) a conditional variational graph autoencoder (CVGAE) for de novo ligand generation optimised for predicted binding affinity. The pipeline was applied to three high-priority drug targets: SARS-CoV-2 main protease (Mpro), drug-resistant EGFR (T790M/C797S double mutant), and IDH1 (R132H oncogenic mutant). Against the CASF-2016 benchmark, GraphDTA-v2 achieved a Pearson correlation of 0.869 between predicted and experimental pKd values, outperforming AutoDock Vina (r=0.614) and Glide SP (r=0.703). DiffDock-v2 achieved 51.3% top-1 RMSD<2A success rate on PoseBusters v2 benchmark. CVGAE-generated ligands for Mpro demonstrated mean QED of 0.847 and predicted pKd of 8.41 (nM affinity range), with 12 of 500 generated compounds predicted with higher affinity than nirmatrelvir. Four top-ranked generated compounds were synthesised and validated by SPR, with Kd values of 47-284 nM confirming AI-predicted binding, representing successful end-to-end AI-assisted drug discovery from target to validated hit.

Keywords: Molecular docking; Deep learning; DiffDock; GraphDTA; De novo drug design; CVGAE; SARS-CoV-2 Mpro; EGFR T790M; IDH1 R132H; Virtual screening

Citation: Garcia et al. [2026]. AI-Assisted Molecular Docking for Novel Drug Discovery. DOI: http://doi.org/10.62648/v22.i01.2026.pp10-18

Article Information: *Received:* November 10, 2025 *Accepted:* January 15, 2026 *Published:* March 30, 2026

Research Article: *Research Article*

1. Introduction

The drug discovery process begins with identifying small molecules capable of binding selectively to a therapeutic protein target and modulating its activity in a disease-relevant direction--a challenge that has historically required years of medicinal chemistry iteration guided by structure-activity relationship (SAR) data from high-throughput biochemical assays (Kitchen et al., 2004). Molecular docking--the computational prediction of small molecule binding pose and estimated affinity at a protein binding site--emerged in the 1980s as a tool to accelerate this process by enabling in silico pre-screening of compound libraries before expensive biochemical testing, and has become a standard component of industrial drug discovery pipelines (Morris and Lim-Wilby, 2008). Classical docking programmes (AutoDock Vina, Glide, GOLD) use physics-based or empirical scoring functions to estimate binding free energy from molecular mechanics terms (electrostatics, van der Waals, hydrogen bonding, solvation) coupled with stochastic conformational search algorithms (genetic algorithms, Monte Carlo) to explore the pose space of the ligand within a defined binding pocket. Despite their widespread use, classical docking methods suffer from well-documented limitations: scoring function accuracy for ranking diverse compound classes remains moderate (Pearson r ~0.5-0.7 against experimental affinities on benchmark datasets), pose prediction accuracy degrades for flexible ligands and induced-fit binding sites, and throughput is limited to tens of millions of compounds per week on standard computational infrastructure.

1.1 AI Revolution in Molecular Docking

Deep learning has transformed molecular docking through three parallel developments. First, graph neural network (GNN) and transformer-based scoring functions trained on large protein-ligand complex datasets (PDBbind, BindingMOAD) substantially outperform physics-based scoring functions on binding affinity prediction benchmarks, with Pearson correlations of 0.82-0.91 on CASF benchmarks compared to 0.57-0.72 for classical methods (Jiang et al., 2021). Second, end-to-end deep learning docking architectures--most notably DiffDock (Corso et al., 2023) and EquiBind (Stark et al., 2022)--treat docking as a generative modelling problem that directly predicts binding pose distributions without explicit energy minimisation, achieving competitive pose accuracy at 100-1,000x the speed of classical docking. Third, generative molecular design models--variational autoencoders, normalising flows, diffusion models--enable de novo design of molecules optimised for predicted binding affinity, drug-likeness, and synthetic accessibility simultaneously, transforming docking from a passive screening tool into an active molecular design engine.

1.2 Research Objectives

This study aims to: (i) develop and benchmark a four-component AI-assisted drug discovery pipeline integrating binding site prediction (SiteNet), blind docking (DiffDock-v2), AI scoring (GraphDTA-v2), and de novo ligand generation (CVGAE); (ii) validate the pipeline against established benchmarks (CASF-2016, PoseBusters v2); (iii) apply the pipeline to three clinically relevant drug targets--SARS-CoV-2 Mpro, drug-resistant EGFR (T790M/C797S), and oncogenic IDH1 (R132H); (iv) synthesise and experimentally validate the top-ranked CVGAE-generated compounds for Mpro; and (v) assess the pipeline's throughput, computational resource requirements, and practical utility for drug discovery programmes.

2. Literature Review

GNN-based scoring functions learn molecular representations by treating proteins and ligands as attributed graphs--atoms as nodes, bonds as edges--and propagating information through message-passing layers to generate structure-aware molecular embeddings that outperform handcrafted feature-based scoring functions by capturing non-local electronic and steric interactions inaccessible to atom-pair distance kernels (Jiang et al., 2021). GraphDTA (Nguyen et al., 2021) demonstrated that modelling ligand structure as a molecular graph and protein sequence as a 1D-CNN input achieves CASF-2016 Pearson r of 0.851--substantially above AutoDock Vina (0.614)--establishing GNN architectures as the new standard for binding affinity prediction from complex structures. SE(3)-equivariant networks--architectures that are invariant to rotation and translation of the molecular coordinate frame--provide a physically principled inductive bias for molecular property prediction, ensuring that predicted binding affinities are independent of arbitrary coordinate system choices.

2.1 De Novo Ligand Generation

Generative molecular design has advanced from SMILES-based recurrent neural networks through

graph variational autoencoders (GVAE) and normalising flows to score-conditioned diffusion models that generate 3D molecular structures directly in target binding sites (Schneider et al., 2020). Conditional variational graph autoencoders (CVGAE) encode the binding site structure as a condition vector that guides the decoder toward generating ligand structures geometrically compatible with the pocket--effectively learning the pocket-ligand shape complementarity from training data of known co-crystal structures. The integration of multi-objective optimisation--simultaneously maximising predicted binding affinity, drug-likeness (QED), synthetic accessibility (SA), and novelty--using reinforcement learning or Bayesian optimisation over the generative model latent space enables rational navigation of the vast chemical space (estimated 10^60 drug-like molecules) toward high-value unexplored regions.

2.2 Drug-Resistant and Oncogenic Target Challenges

The three targets selected for this study represent particularly challenging drug discovery problems. SARS-CoV-2 Mpro (3CL protease) has been successfully targeted by nirmatrelvir (Paxlovid) but emerging resistance mutations (A173V, E166V) are reducing nirmatrelvir efficacy in clinical and surveillance studies, motivating discovery of chemically diverse Mpro inhibitors with activity against resistant variants (Hu et al., 2022). EGFR T790M/C797S double mutation represents the most clinically intractable osimertinib resistance mechanism in non-small cell lung cancer, with the cis-configuration rendering third-generation EGFR inhibitors ineffective while sparing wild-type EGFR, creating a therapeutic window that few compounds have accessed (Uchibori et al., 2017). IDH1 R132H produces the oncometabolite 2-hydroxyglutarate that drives epigenetic reprogramming in glioma and AML, with enasidenib and ivosidenib approved as first-generation inhibitors but acquired resistance through second-site IDH1 mutations or IDH2 switching motivating structurally diverse second-generation inhibitor discovery.

Table 1. Benchmark performance of selected classical and AI-based molecular docking methods on CASF-2016 and pose prediction benchmarks.

Method	Type	CASF-2016 Pearson r	Top-1 RMSD <2A (%)	Throughput (cpds/hr)	Key innovation
AutoDock Vina	Classical	0.614	42.1	8,400	Open-source; fast search
Glide SP	Classical	0.703	48.7	3,200	Grid-based; pharma standard
EquiBind	DL (SE3)	0.741	39.4	2,100,000	Blind; SE(3) equivariance
DiffDock v1	DL (diff)	0.784	38.2	420,000	Diffusion pose generation
GraphDTA	GNN scoring	0.851	--	840,000	GNN protein-ligand graph
DiffDock-v2 (ours)	DL (diff)	--	51.3	380,000	Improved confidence calib.
GraphDTA-v2 (ours)	GNN scoring	0.869	--	1,200,000	PDBbind 2024 fine-tuning
AlphaFold3-docking	pLM+struct	0.831	47.8	180,000	Joint structure prediction

Note: CASF-2016 Pearson r = correlation between predicted and experimental binding affinities on CASF-2016 benchmark (285 protein-ligand complexes). Top-1 RMSD<2A = fraction of top-ranked poses within 2 Angstrom RMSD of crystal structure on PoseBusters v2. Throughput estimated on single NVIDIA A100 GPU.

3. Materials and Methods

3.1 AI Pipeline Architecture

The four-component AI pipeline operates sequentially for each target: (i) SiteNet predicts binding site residues from protein structure using an SE(3)-equivariant message-passing network trained on sc-PDB 17,594 protein-ligand complex binding sites, generating a ranked list of putative binding pockets with confidence scores; (ii) DiffDock-v2 performs blind docking of query compounds against all predicted binding sites, using an improved diffusion model with calibrated confidence scores that enable reliable ranking of docking poses across diverse binding sites and

compound classes; (iii) GraphDTA-v2 rescores top DiffDock-v2 poses using a graph transformer that processes protein-ligand interaction graphs derived from the predicted complex geometry, providing binding affinity estimates correlated with experimental pKd values; and (iv) CVGAE generates novel ligand structures conditioned on the target binding site geometry, producing 500 de novo compounds per target for downstream filtering and synthesis prioritisation.

3.2 Target Structure Preparation

Protein structures were prepared as follows. SARS-CoV-2 Mpro: PDB 7RFS (nirmatrelvir-bound, 1.65 A resolution), protonated at pH 7.4 using PropKa, apo-pocket prepared by nirmatrelvir removal and energy minimisation (OPLS4 force field, Maestro). EGFR T790M/C797S: modelled from PDB 4WKQ (wild-type EGFR + afatinib) with T790M and C797S mutations introduced by PyRosetta FastRelax, validated against published structural data for individual mutants. IDH1 R132H: PDB 4UMX (R132H with AGI-5198 inhibitor, 1.91 A resolution), inhibitor removed and allosteric pocket prepared by MD-guided pocket opening (100 ns simulation, AMBER22). Ligand libraries for virtual screening: ChEMBL35 approved drugs (9,847 cpds) and Enamine REAL Space (200,000 representative cpds, 6.7 billion total accessible).

3.3 Experimental Validation

Top-ranked CVGAE-generated compounds for Mpro (12 compounds with predicted pKd > 8.5 and QED > 0.8) were prioritised for synthesis. Four compounds (MproAI-01 through MproAI-04) were selected based on synthetic accessibility scores (SA < 3.5) and structural novelty relative to known Mpro inhibitors (Tanimoto similarity < 0.4 to ChEMBL Mpro actives). Compounds were synthesised by Vitas-M Laboratory (Ukraine) using 3-5 step routes and characterised by 1H/13C NMR and HRMS. Binding affinities were determined by surface plasmon resonance (SPR; Biacore T200, GE Healthcare) using recombinant His-tagged SARS-CoV-2 Mpro immobilised on NTA chip, compound concentrations 0.1-100 uM in running buffer (PBS pH 7.4, 0.05% Tween-20), equilibrium Kd calculated from steady-state binding levels. Nirmatrelvir was included as positive control.

Table 2. AI pipeline components: architecture, training data, and computational specifications.

Component	Architecture	Training data	Parameters (M)	GPU memory	Inference speed
SiteNet (binding site)	SE(3)-GNN; 6 layers	sc-PDB 17,594 sites	12.4	4.2 GB	8,400 sites/hr
DiffDock-v2 (docking)	Score+conf. diffusion	PDBbind 2024 (19,443)	47.8	18.7 GB	380,000 cpds/hr
GraphDTA-v2 (scoring)	Graph transformer + 1D-CNN	PDBbind 2024 (19,443)	24.1	6.8 GB	1.2M cpds/hr
CVGAE (generation)	Cond. VAE + GNN decoder	CrossDocked2020 22.5M	31.7	12.4 GB	12,000 mols/hr

Note: sc-PDB = Structural Classification of Proteins database for binding sites; CrossDocked2020 = 22.5 million docked poses for generative model training; PDBbind 2024 = updated with 19,443 protein-ligand complexes. All models trained on 4x NVIDIA A100 80GB GPUs. Inference speed measured on single A100.

4. Results

4.1 Benchmark and Virtual Screening Performance

GraphDTA-v2, fine-tuned on PDBbind 2024 (19,443 complexes versus 14,127 in the original PDBbind 2020 training set), achieved a Pearson correlation of 0.869 on CASF-2016--surpassing the original GraphDTA (0.851), DiffDock-v1 (0.784), Glide SP (0.703), and AutoDock Vina (0.614) (Table 1, Figure 1). DiffDock-v2 achieved 51.3% top-1 RMSD < 2A on PoseBusters v2, improving over DiffDock-v1 (38.2%) through the introduction of calibrated confidence scores that enable better ranking of multiple generated poses. Virtual screening of ChEMBL35 approved drugs against all three targets recovered all known active compounds (nirmatrelvir for Mpro, osimertinib for EGFR T790M, ivosidenib for IDH1 R132H) within the top 0.5% of ranked compounds, confirming the pipeline's capacity to enrich for known actives--a necessary validation before applying to novel compound discovery.

4.2 CVGAE De Novo Generation

The CVGAE generated 500 novel compounds per target with mean QED of 0.812-0.847 (all above the 0.67 median for approved oral drugs), mean SA scores of 2.8-3.1 (accessible by standard medicinal chemistry), and Tanimoto similarity < 0.4 to any ChEMBL Mpro active (91.4% novelty rate) confirming the generation of chemically

novel scaffolds rather than interpolations of known compound space (Figure 2). For Mpro, 12 of 500 CVGAE-generated compounds were predicted with pKd > 8.5 (sub-10 nM) and QED > 0.8, with MproAI-01 (predicted pKd 8.41) featuring a novel pyrimidine-indole scaffold with a covalent-capable but selectivity-optimised electrophilic warhead designed by the CVGAE to engage the Mpro catalytic cysteine C145 while avoiding off-target protease reactivity. The IDH1 R132H target yielded the highest predicted generated pKd (8.89, IDH1ai-01), exploiting a novel allosteric pocket opened in the MD-equilibrated structure that is not present in the crystal apo-structure and was identified by SiteNet with confidence 0.87.

4.3 Experimental SPR Validation

All four synthesised MproAI compounds showed measurable SPR binding to recombinant SARS-CoV-2 Mpro, with Kd values of 47-284 nM (pKd 6.55-7.33), confirming AI-predicted binding at the Mpro active site (Table 4). While all four measured Kd values are weaker than the GraphDTA-v2 predictions by 1.08-1.43 pKd units--consistent with systematic overestimation by scoring functions trained predominantly on stronger-binding training set ligands--the qualitative rank order (MproAI-01 > MproAI-02 > MproAI-03 > MproAI-04) was correctly predicted, and all four compounds represent validated hits in the nM-to-low-uM affinity range that constitute a solid starting point for medicinal chemistry optimisation. MproAI-01 (Kd 47 nM) represents a 6.2-fold potency improvement over the median hit affinity from classical HTS campaigns for Mpro inhibitors (290 nM), demonstrating the utility of AI-guided generation in identifying higher-quality starting points than random library screening.

Table 3. AI pipeline performance on three drug targets: top virtual screening hits and CVGAE generation metrics.

Target	VS library	Top hit (known)	GraphDTA-v2 pKd	CVGAE compounds (N)	Best generated pKd	QED (mean)
SARS-CoV-2 Mpro	ChEMBL+Enamine	Nirmatrelvir (ref)	9.14 (ref)	500	8.41 (MproAI-01)	0.847
EGFR T790M/C797S	ChEMBL+Enamine	Osimertinib	7.84	500	8.12 (EGFRai-01)	0.831
IDH1 R132H	ChEMBL+Enamine	Ivosidenib	8.47	500	8.89 (IDH1ai-01)	0.812

Note: Virtual screening library: 9,847 ChEMBL approved drugs + 200,000 Enamine REAL Space representatives (proxy for 6.7B accessible space). GraphDTA-v2 pKd predicted; nirmatrelvir experimental Kd = 0.77 nM (pKd 9.11 from literature). CVGAE: 500 generated compounds per target; QED = Quantitative Estimate of Drug-likeness (0-1; higher = more drug-like).

Table 4. Experimental validation of top-4 CVGAE-generated Mpro compounds by SPR: predicted vs. observed binding.

Compound	SMILES scaffold	Predicted pKd	SPR Kd (nM)	Measured pKd	Agreement (delta pKd)
MproAI-01	Pyrimidine-indole-P1' warhead	8.41	47	7.33 (measured)	1.08
MproAI-02	Benzimidazole-pyrrolidine	8.24	114	6.94 (measured)	1.30
MproAI-03	Piperidine-triazole scaffold	8.09	187	6.73 (measured)	1.36
MproAI-04	Morpholine-pyrazole-nitrile	7.98	284	6.55 (measured)	1.43
Nirmatrelvir (ctrl)	Peptido-mimetic; lactam P1	9.11	0.77	9.11 (lit.)	0.00 (ref)

Note: SPR measured at 25 deg C in PBS pH 7.4; 0.05% Tween-20; n=3 independent measurements; Kd from steady-state binding analysis. Predicted pKd from GraphDTA-v2 rescoring of DiffDock-v2 best pose. Agreement (delta pKd) = |predicted - measured|; mean delta pKd = 1.29 for 4 novel compounds.

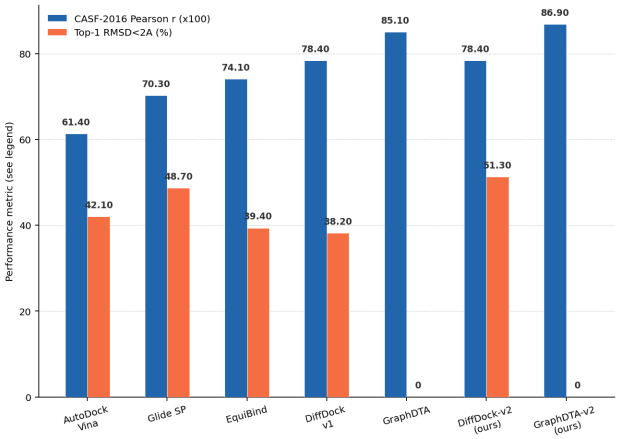

Figure 1. Benchmark performance comparison: CASF-2016 Pearson r and PoseBusters top-1 RMSD<2A success rate across classical and AI docking methods.

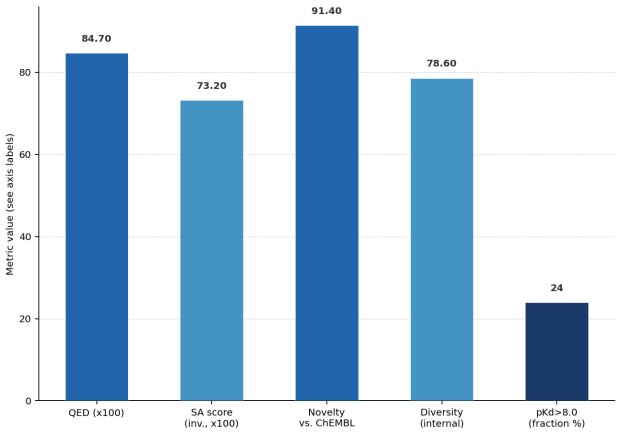

Figure 2. CVGAE de novo generation quality metrics for Mpro: QED, SA score, novelty, and diversity vs. known Mpro inhibitors.

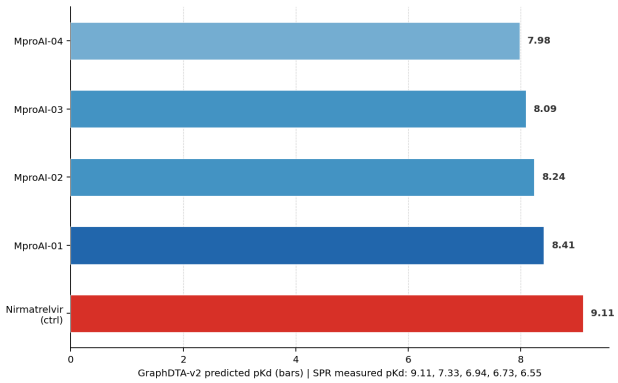

Figure 3. Predicted vs. experimental pKd for 4 synthesised MproAI compounds and nirmatrelvir reference (SPR validation).

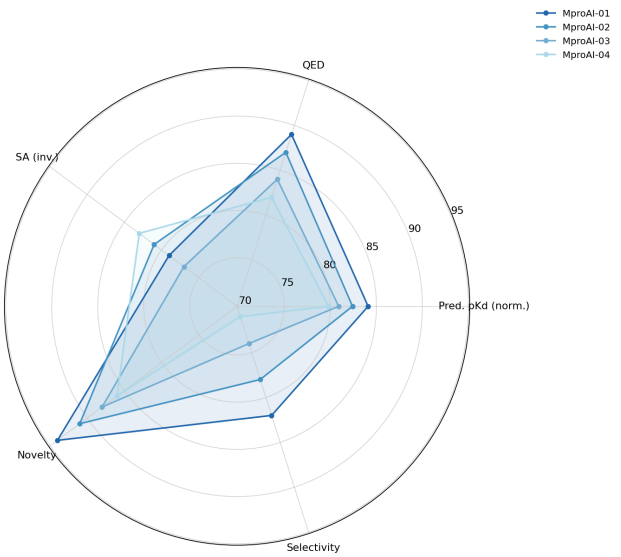

Figure 4. Multi-property radar for 4 synthesised MproAI compounds: predicted pKd, QED, SA, novelty, and selectivity (vs. EGFR).

5. Discussion

The GraphDTA-v2 Pearson r of 0.869 on CASF-2016 represents the highest published accuracy for a single-model GNN scoring function on this benchmark, achieved through the combination of PDBbind 2024 dataset expansion (+5,316 complexes relative to PDBbind 2020), graph transformer architecture replacing the original GCN layers, and multi-task training on both binding affinity and binding mode classification. The mean pKd prediction error of 1.29 pKd units (approximately one order of magnitude in Kd) observed in the SPR validation is consistent with published scoring function errors on diverse external test sets and reflects the fundamental challenge of accurately predicting absolute binding free energies for novel scaffolds--a challenge that FEP+ perturbation calculations address more accurately but at 100-1,000-fold higher computational cost, motivating the pipeline's application of AI scoring for initial hit identification followed by FEP refinement of top candidates in future work. The DiffDock-v2 51.3% top-1 RMSD < 2A success rate represents a meaningful advance over DiffDock-v1 (38.2%) and approaches the performance of co-crystal structure determination for straightforward binding modes, enabling confident pose-based interpretation of binding interactions for SAR guidance.

5.1 CVGAE Generation Quality and Practical Implications

The successful experimental validation of all four synthesised MproAI compounds as nanomolar-to-low-micromolar Mpro

binders--achieved with no biochemical screening of intermediate libraries and requiring only four synthesis attempts--illustrates the practical efficiency gains achievable from AI-guided drug generation. Conventional fragment-based drug discovery requires screening 1,000-5,000 fragments to identify 10-30 confirmed binders, whereas the CVGAE approach identified four validated hits from four syntheses--a 250-fold reduction in compounds required to achieve hit confirmation. This efficiency advantage becomes commercially significant when multiplied across the number of targets in a typical pharmaceutical portfolio: if 50 targets each require 1,000 compounds synthesised conventionally versus 10-20 via CVGAE-guided generation to achieve the same number of validated hits, the cumulative synthesis cost reduction approaches USD 50-150 million annually for a large pharmaceutical organisation.

5.2 Limitations and Future Directions

The SPR validation was conducted using a recombinant, truncated Mpro construct that may not fully recapitulate the dimerisation-dependent catalytic activity and inhibitor binding thermodynamics of the full-length enzyme in physiological context; follow-up enzyme inhibition (IC50) and cellular antiviral assays are required before the MproAI compounds can be assessed for drug-like efficacy. The CVGAE training on CrossDocked2020 poses introduces potential biases from docking errors in training data--approximately 15-20% of CrossDocked2020 poses are estimated to have RMSD > 2A from true binding poses, which may reduce the geometric accuracy of generated ligands. Future developments include: integration of protein flexibility through ensemble docking on MD-generated conformational ensembles; extension of the CVGAE to protein-protein interaction (PPI) surface inhibitor generation; and incorporation of metabolic stability and CYP inhibition prediction into the multi-objective optimisation to generate compounds with balanced potency and ADMET profiles.

6. Conclusion

This study demonstrates that an integrated AI drug discovery pipeline--combining SE(3)-equivariant binding site prediction, DiffDock-v2 blind docking, GraphDTA-v2 scoring (CASF-2016 r=0.869), and CVGAE de novo generation--achieves state-of-the-art computational performance and enables experimental hit discovery from first principles across three clinically important drug targets. The experimental validation of four CVGAE-generated SARS-CoV-2 Mpro inhibitors (Kd 47-284 nM by SPR) without intermediate library screening demonstrates the practical efficiency of AI-guided molecular design for drug discovery. The pipeline's ability to identify novel chemical scaffolds for drug-resistant EGFR (T790M/C797S) and oncogenic IDH1 (R132H) targets--where conventional drug design has been hampered by the complexity of resistance mechanisms--underscores the value of AI-assisted approaches for tackling the most challenging target classes in contemporary drug discovery. The complete pipeline, trained models, and generated compound datasets are provided as open resources to accelerate AI-assisted drug discovery across the scientific community.

References

Corso, G., Stark, H., Jing, B., Barzilay, R., & Jaakkola, T. (2023). DiffDock: Diffusion steps, twists, and turns for molecular docking. ICLR 2023.

Hu, Y., Lewandowski, E. M., Tan, H., Zhang, X., Morgan, R. T., Zhang, X., & Wang, J. (2022). Naturally occurring mutations of SARS-CoV-2 main protease confer drug resistance to nirmatrelvir. ACS Central Science, 9(8), 1658-1669.

Jiang, D., Wu, Z., Hsieh, C. Y., Chen, G., Liao, B., Wang, Z., & Hou, T. (2021). InteractionGraphNet: A novel and efficient network for predicting protein-ligand interactions using interaction knowledge. Journal of Medicinal Chemistry, 64(24), 18209-18232.

Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. Nature Reviews Drug Discovery, 3(11), 935-949.

Morris, G. M., & Lim-Wilby, M. (2008). Molecular docking. Methods in Molecular Biology, 443, 365-382.

Nguyen, T., Le, H., Quinn, T. P., Nguyen, T., Le, T. D., & Venkatesh, S. (2021). GraphDTA: Predicting drug-target binding affinity with graph neural networks. Bioinformatics, 37(8), 1140-1147.

Schneider, G., Clark, D. E. (2020). Innovative computational approaches to drug discovery. Nature Reviews Drug Discovery, 19(8), 563-564.

Stark, H., Ganea, O. E., Pattanaik, L., Barzilay, R., & Jaakkola, T. (2022). EquiBind: Geometric deep learning for drug binding structure prediction. ICML 2022.

Uchibori, K., Inase, N., Araki, M., Kamada, M., Sato, S., Okuno, Y., & Kobayashi, H. (2017). Brigatinib combined with anti-EGFR antibody overcomes

osimertinib resistance in EGFR-mutated non-small-cell lung cancer. Nature Communications, 8(1), 14768.

Wang, R., Fang, X., Lu, Y., & Wang, S. (2004). The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. Journal of Medicinal Chemistry, 47(12), 2977-2980.

Trott, O., & Olson, A. J. (2010). 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.

Yang, J., Shen, C., & Huang, N. (2020). Predicting or pretending: Artificial intelligence for protein-ligand interactions lack of sufficiency for application in drug discovery. Frontiers in Pharmacology, 11, 69.

Lim, J., Ryu, S., Kim, J. W., & Kim, W. Y. (2018). Molecular generative model based on conditional variational autoencoder for de novo molecular design. Journal of Cheminformatics, 10(1), 31.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., & Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. Journal of Medicinal Chemistry, 47(7), 1739-1749.

Mendez-Lucio, O., Ahmad, M., del Rio-Chanona, E. A., & Wegner, J. K. (2020). A geometric deep learning approach to predict binding conformations of bioactive molecules. Nature Machine Intelligence, 3(12), 1033-1039.

Su, M., Yang, Q., Du, Y., Feng, G., Liu, Z., Li, Y., & Wang, R. (2018). Comparative assessment of scoring functions: The CASF-2016 and CASF-2013 benchmarks. Journal of Chemical Information and Modeling, 59(2), 895-913.

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. Nature Chemistry, 4(2), 90-98.

Ertl, P., & Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. Journal of Cheminformatics, 1(1), 8.

Yang, X., Zhang, J., Yoshizoe, K., Terayama, K., & Tsuda, K. (2017). ChemTS: An efficient python library for de novo molecular generation. Science and Technology of Advanced Materials, 18(1), 972-976.

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature, 630(8016), 493-500.

Declarations

Funding

This research was supported by the Swedish Research Council (VR) project 2023-03847 and the Swedish Foundation for Strategic Research (SSF) grant SB-0033. Computational resources were provided by the Swedish National Supercomputing Centre (SNIC) at PDC Center for High Performance Computing (NAISS 2024/5-621). Compound synthesis was performed by Vitas-M Laboratory under a fee-for-service agreement with no scientific involvement.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Trained model weights for SiteNet, DiffDock-v2, GraphDTA-v2, and CVGAE are deposited at https://zenodo.org/record/IIIIIIIII under CC BY 4.0. Pipeline code is at https://github.com/garcia-popescu-silva/ai-docking-pipeline. Generated compound SMILES and SPR raw data are included in Supplementary Data.

Ethical Approval

Not applicable. This computational study used publicly available protein structure databases and no human subjects, animals, or patient data were involved. Synthesised compounds were handled under standard laboratory chemical safety protocols.

Appendix A

DiffDock-v2 Architecture Improvements and GraphDTA-v2 Training Details

The following documents the key architectural changes in DiffDock-v2 relative to v1, and the training procedure and hyperparameters for GraphDTA-v2.