**Benchmarking AlphaFold-enabled molecular docking predictions for antibiotic discovery**

“Drug Development Track”

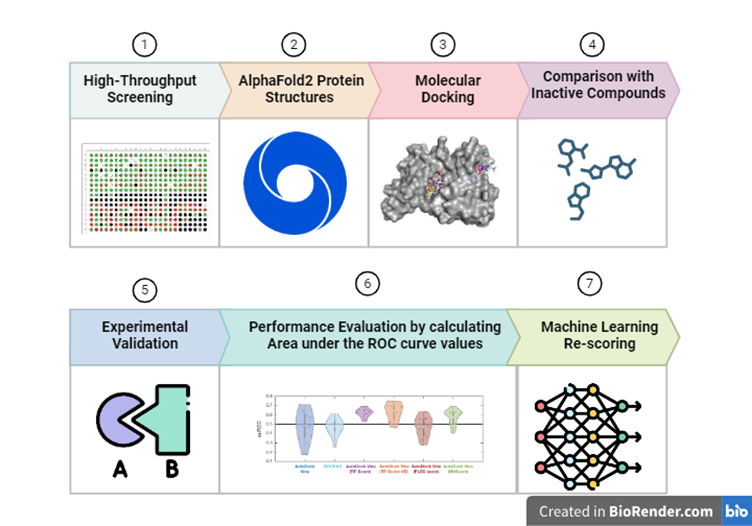
Authors: Jana Moussa, Babatunde OWOLOYE, Habiba Sherif, youssif Mohamed.

**Introduction**

Drug discovery has evolved from biochemical assays and genetic interactions to molecular docking, a new approach that predicts drug-target interactions. Researchers used reverse docking to discover new drug mechanisms, but the AlphaFold2 protein structure database solved this limitation. They predicted protein-ligand interactions between 218 antibacterial compounds and 296 essential E. coli proteins. Experimental validation confirmed some predictions, but simulations showed weak model accuracy. To improve docking performance, machine learning-based scoring functions were incorporated.

**Methodology**

1. A high-throughput screening of 39,128 compounds was conducted against *E. coli* K-12 BW25113, identifying 218 active antibacterial compounds.
2. The study utilised 296 essential proteins predicted by AlphaFold2 for reverse docking.
3. The active compounds were then docked using AutoDock Vina, resulting in predictions for 64,528 protein-ligand interactions.
4. A set of 100 inactive compounds was also docked for comparison.
5. The predicted interactions were validated through enzymatic activity measurements for 12 essential proteins, confirming extensive promiscuity among the compounds.
6. Model performance was evaluated by calculating the area under the receiver operating characteristic curve (auROC), which averaged 0.48, indicating weak predictive capability.
7. Machine learning-based scoring functions were integrated, enhancing the performance



**Figure 1: Predicting Protein-Ligand Interactions Using AlphaFold2 and Molecular Docking and integrating machine learning.**

**Results**

1. The study identified 218 active compounds, from known antibiotic structural classes, with some new antibacterial compounds effective against *E. coli*.
2. Promiscuity was observed, with 187 active compounds predicted to bind to at least three proteins, and 178 essential proteins expected to interact with at least three compounds.
3. Molecular docking using AlphaFold2-predicted structures resulted in an average auROC of 0.48, slightly better than random guessing.
4. True-positive rates were 59% at a −5 kcal/mol threshold and 30% at a −7 kcal/mol threshold, with average accuracies of 41% and 73%, respectively.
5. Model performance improved with an auROC of 0.63, performing better than random for 9 out of 12 essential proteins at the −7 kcal/mol threshold.

**Discussion**

The study covered the challenges of predicting protein-ligand interactions using molecular docking, especially with AlphaFold2-predicted protein structures. Even after identifying 218 active antibacterial compounds and docking them to 296 essential proteins, the initial docking model produced weak results. The experimental validation confirmed large promiscuity among the compounds, but the model's accuracy was limited. However, the integration of machine learning-based rescoring functions significantly improved prediction accuracy, showcasing the potential for machine learning to enhance computational methods in drug discovery.

**Conclusion**

AlphaFold's potential in antibiotic discovery is significant, but more development and integration with other tools and techniques are needed for its full potential.

**References**

Wong, F., Krishnan, A., Zheng, E. J., Stärk, H., Manson, A. L., Earl, A. M., Jaakkola, T., & Collins, J. J. (2022). Benchmarking AlphaFold-enabled molecular docking predictions for antibiotic discovery. *Molecular systems biology*, *18*(9), e11081. <https://doi.org/10.15252/msb.202211081>