Using Flexible Docking and Reinforcement Learning to Guide Antibody Design

Ananth Shyamal, Mohamed Samb, Wilson Ho

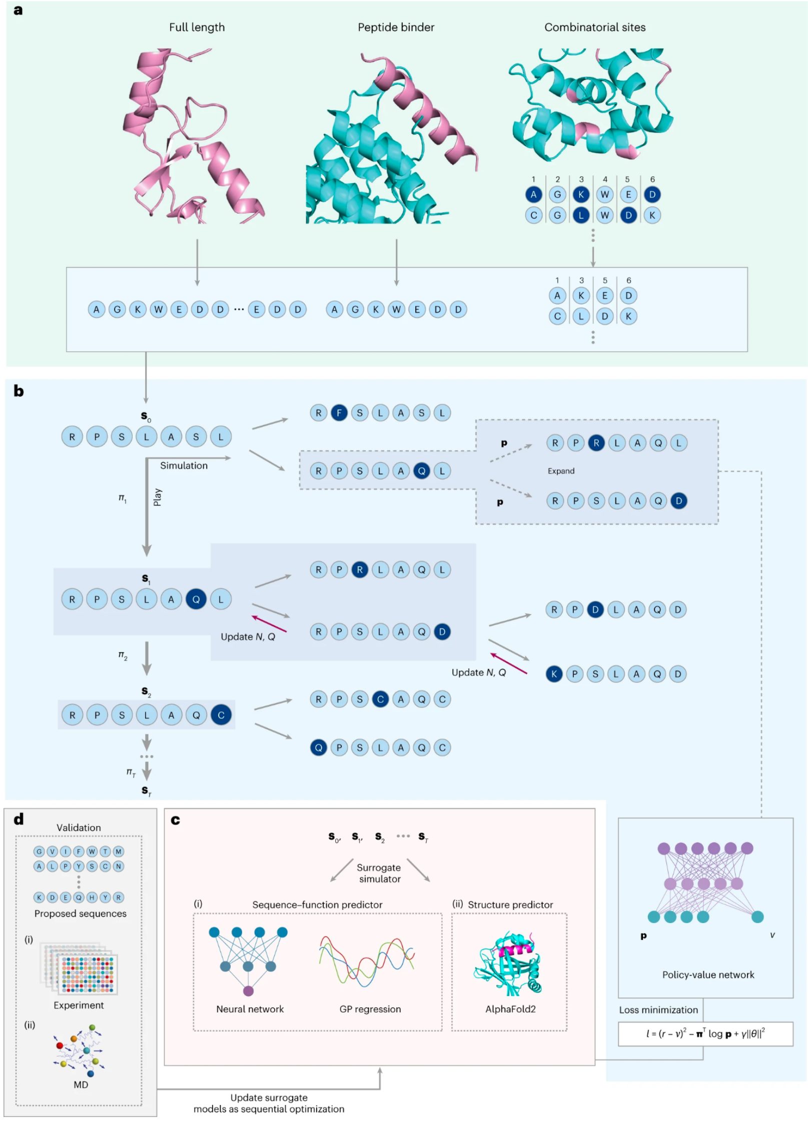
Background/Introduction:

Antibodies and antigens are central players in the body's immune system. Antibodies are produced by B cells to specifically bind to antigens and result in an immune response to the detected antigens. Modeling antigen-antibody interactions is clinically relevant due to the various therapeutic applications of antigens and antibodies (e.g., vaccine development and monoclonal antibodies). It is experimentally time consuming and expensive to model these interactions (using techniques like X-ray crystallography, NMR spectroscopy, and Cryo-EM). Therefore, in silico models that predict antigen-antibody docking are particularly useful, and such models may be used to design antibodies for specific epitopes.

Current models for protein docking employ a variety of methods. Physics-based models estimate the potential energy of configurations, which can be expressed as a function of various energy terms, including those associated with forces such as electrostatic and van der Waals. These kinds of models tend to be computationally expensive, which has led to the development of machine-learning based models that are trained on large sequence and 3D structure datasets. Existing models employ either a rigid or flexible docking scheme, where in rigid docking, both the protein and the ligand are modeled as rigid structures. While this paradigm increases computational efficiency, it often isn't a realistic model since proteins and molecules change in conformation in biological settings. One of the state-of-the-art protein docking models is DockGPT, which uses a transformer-based model that is able to perform protein docking in a flexible and site-specific manner.

High affinity binder design can be driven by many different approaches. Existing docking models are often able to perform simultaneous docking and sequence-structure co-design. A modification to DockGPT masks and predicts linear segments of a protein chain, analogous to word prediction in a language model. Jin et al. (2022) generates an antibody sequence autoregressively, iteratively predicting the next residue and performing one-step structure refinement.

Inspired by directed evolution, in silico protein design can also be directed by reinforcement learning, using iterative rounds of random mutation and candidate scoring. EvoPlay (Wang et al., 2023) implements a neural network-guided Monte Carlo tree search, based on deep learning models for chess games. The optimization process starts from a complete protein sequence and performs a series of moves representing single-site mutations. To evaluate mutants during the tree search process, EvoPlay queries a surrogate model to retrieve an environmental reward. For instance, the surrogate can be a machine learning model that maps sequences to functions. The model also supports AlphaFold, which allows incorporation of rewards based on the protein-peptide interface.



Overview of EvoPlay protein design

EvoPlay's reward function uses residue distance and contact features predicted by AlphaFold. Although these are easy to obtain from a given structure, such features provide a limited description of the interactions involved in peptide binding. Other docking models have experimented with physics-based scoring functions to evaluate protein complexes. Esquivel-Rodríguez et al. (2012), for example, considers a combination of van der Waals, electrostatic, hydrogen bonding, and solvation terms. Compared to other scoring schemes, the rigorous physics-based score provides a higher prediction accuracy. However, the score is considerably more expensive to compute, which makes it a potential bottleneck for model training.

Innovation (we do not expect you to make significant innovations but want to know what you feel is new about the work you are proposing**):**

EvoPlay currently only supports AlphaFold as a surrogate reward model. Although AlphaFold has been able to predict protein-peptide complexes with high-resolution, it is not specialized for flexible protein docking, as it does not explicitly make use of binding site information. Using the same reinforcement learning framework as EvoPlay, we plan to incorporate DockGPT as an alternative reward model. Since DockGPT is fine-tuned on flexible complexes, its predicted structures can be used to specifically design better antibodies.

We also plan to extend EvoPlay's structure-based reward function by incorporating estimates of binding energy. Libraries such as MM/PBSA can be used to computationally approximate interaction energies between two structures. By incorporating energies, the model may become better at specifically designing proteins with higher binding affinities.

Significance:

Updating EvoPlay's peptide binder structure prediction model to use binding energies predictions allows us to incorporate more important information into a reinforced learning model. Antibody-antigen binding energies are significant indicators as to whether a predicted antibody structure is correct, so their integration into training the model will yield improved results. Through incorporating binding energies into a reinforcement learning reward function, we will be able to produce more accurate antibody structure predictions, which have significant applications in a variety of fields. These improvements are extremely important for immunology, allowing us to more accurately model antigen-antibody interactions within immune systems. The significance of these innovations can also be found in therapeutics, with highly accurate antibody structure predictions supporting the development of vaccines and other antibody based therapeutics.

Specific Aims (including approaches):

Update EvoPlay to support DockGPT as a reward model. We plan to train with the same reinforcement learning framework, iteratively mutating antibody CDR sequences to achieve the best reward. Some parameter tuning in DockGPT may be necessary in order to efficiently generate structure predictions.

Update reward function to incorporate binding energies. We plan to run MM/PBSA on structures predicted by DockGPT. Although the library provides streamlined energy calculations, efficiency may still be a concern, especially for reinforcement learning. It may be necessary to build or adapt a cheap regression model that predicts binding energy.

Resources:

- Data availability:
    - Antibody-antigen complexes: SAbDab (https://opig.stats.ox.ac.uk/webapps/sabdab-sabpred/sabdab)
    - Protein 3D structures: PDB (https://www.rcsb.org/)
    - Unpaired/paired antibody sequences: OAS (https://opig.stats.ox.ac.uk/webapps/oas/)

- Methods/Code availability:

- DockGPT: https://github.com/MattMcPartlon/protein-docking
- EvoPlay: https://github.com/melobio/EvoPlay
- MM/PBSA: https://ambermd.org/AmberTools.php

Division of labor:

- Wilson: adapting DockGPT to support iterative mutations
- Mohamed: designing reward function, exploring options for efficient binding energy prediction
- Ananth: adapting EvoPlay framework, overseeing the training process

References:

Esquivel-Rodríguez, J., Yang, Y. D., and Kihara, D. (2012). Multi-LZerD: Multiple protein docking for asymmetric complexes. Proteins 80 (7), 1818–1833. doi:10.1002/prot.24079

Jin, W., Barzilay, R., Jaakkola, T. Antibody-Antigen Docking and Design via Hierarchical Equivariant Refinement. arXiv 2207.06616; https://arxiv.org/pdf/2207.06616

McPartlon, M., Xu, J. Deep Learning for Flexible and Site-Specific Protein Docking and Design. bioRxiv 2023.04.01.535079; https://doi.org/10.1101/2023.04.01.535079

Ruffolo, J.A., Chu, LS., Mahajan, S.P. et al. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. Nat Commun 14, 2389 (2023). https://doi.org/10.1038/s41467-023-38063-x

Wang, Y., Tang, H., Huang, L. et al. Self-play reinforcement learning guides protein engineering. Nat Mach Intell 5, 845–860 (2023). https://doi.org/10.1038/s42256-023-00691-9