Accurate prediction of antibody-antigen and protein-protein complex structure accelerates next-gen biologics development

ab.bd@xtalpi.com
Booth #519

Chunqiu Xia, Chuheng Liang, Yihang Zhou, Yunfei Long, Yi Li

Problem: Predicting the structure of antigen-antibody complexes is a critical challenge

While AlphaFold2 has shown remarkable success in predicting protein monomer structures, accurately predicting the structure of protein complexes remains a challenge.

AlphaFold-Multimer is currently the most renowned algorithm for predicting protein complex structures, but its performance in determining the structure of antigen-antibody complexes remains suboptimal.

Several existing methods, such as IgFold[3], OmegaFold[4], and ESMFold[5], can predict antibody structures, but not antigen-antibody complexes.

To address this limitation, we propose **XtalFold™, a novel approach capable of predicting antigen-antibody complex structures with high accuracy and enabling many downstream applications.**

Solution: XtalFold™ is an AI-driven approach for effective antigen-antibody complex structure prediction

Recent literature[6] has utilized a rigorously curated test set from the Protein Data Bank (PDB) to assess the performance of AlphaFold-Multimer in predicting antigen-antibody complex structures.

To benchmark the performance of XtalFold™ with AlphaFold-Multimer, we selected a fair test set (N=39) comprising newly released structures that are **NOT INCLUDED** in the training dataset of both AlphaFold-Multimer and XtalFold™, and are **DISTINCT** from the sequences or structures of the training set.

Additionally, we incorporated a structural confidence (SC) score as a filter into XtalFold™ to further enhance prediction quality. The effectiveness of this filter was compared with AlphaFold-Multimer's pLDDT metric.

PDB IDs and chain names of the test set (N=39)

7VNB_AB, 7E53_BA, 7ANQ_BA, 7NFQ_CA, 7NFR_BA, 7NX0_DC, 7AR0_BA, 7AQY_CB, 7DAA_HLA, 7L6V_BA, 7T5F_CA, 7T5F_ED, 7L6V_DA, 7S11_IMD, 7M1H_GA, 7L6V_CA, 7L6V_FA, 7M1H_FA, 7VUX_HLA, 7M1H_EA, 7LZP_ED, 7NA9_DA, 7PS6_HLE, 7E72_CDF, 7PS4_HLE, 7PS2_HLG, 7PS0_HLE, 7Q0G_ABE, 7MZJ_HLA, 7MZK_NMB, 7BNV_HLA, 7MZM_HLA, 7N4J_HLA, 7L0L_HLBA, 7Q0I_HLC, 7NX3_BCF, 7L7R_BAG, 7L7R_DCG, 7BBJ_HLA.

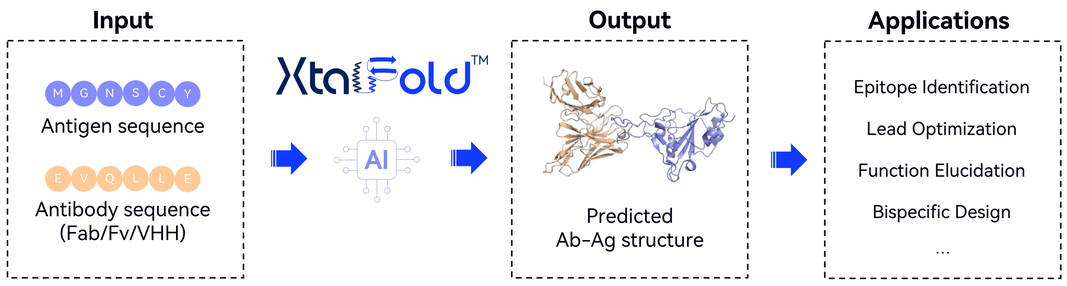


Figure 1. Schematic diagram of XtalFold™. The input for this algorithm requires only the antigen and antibody sequences (Fv/Fab/VHH), and the output generated is the antigen-antibody complex structure. Additionally, we can use SC as an effective filter to significantly improve the quality of prediction results.

Conclusion: XtalFold™ delivers a high success rate for Ab-Ag complex structure prediction

We have successfully developed an AI-driven solution capable of predicting the complex structure of antigen-antibody with higher success rate and interface accuracy, relying solely on sequence information.

XtalFold™ demonstrates significant potential for predicting both antigen-antibody and general protein-protein complex structures.

We are leveraging this groundbreaking tool in conjunction with our Generative AI XenProT™, Predictive AI Xentient™, and full-scale wet lab to unlock numerous downstream applications in next-generation biologics discovery and engineering.

Benchmark results: XtalFold™ outperforms AlphaFold-Multimer in a fair test set across multiple metrics

XtalFold™ outperforms AlphaFold-Multimer across multiple performance metrics (DockQ, TM-score, and interface RMSD). After filtering out low-confidence predictions, XtalFold™ can achieve up to 90% success rate.

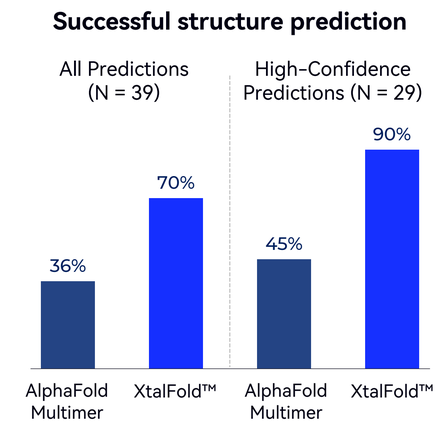


Figure 2. Prediction success rate of XtalFold™ and AlphaFold-Multimer based on DockQ. Successful prediction is defined as DockQ > 0.23. High-confidence predictions are a subset of predictions with higher confidence scores generated by each model, i.e., pLDDT for AlphaFold and SC (Structural Confidence) for XtalFold. Thresholds are set to allow comparison of the equal number (N = 29) of higher-quality predictions.

Regarding iRMSD, XtalFold achieved **4.79 Å** compared to 11.13 Å of AlphaFold. Among high-confidence predictions, interface accuracy of XtalFold structures reached **2.25 Å** (SC > 0.8) in contrast to 10.97 Å of AlphaFold (pLDDT > 0.8).

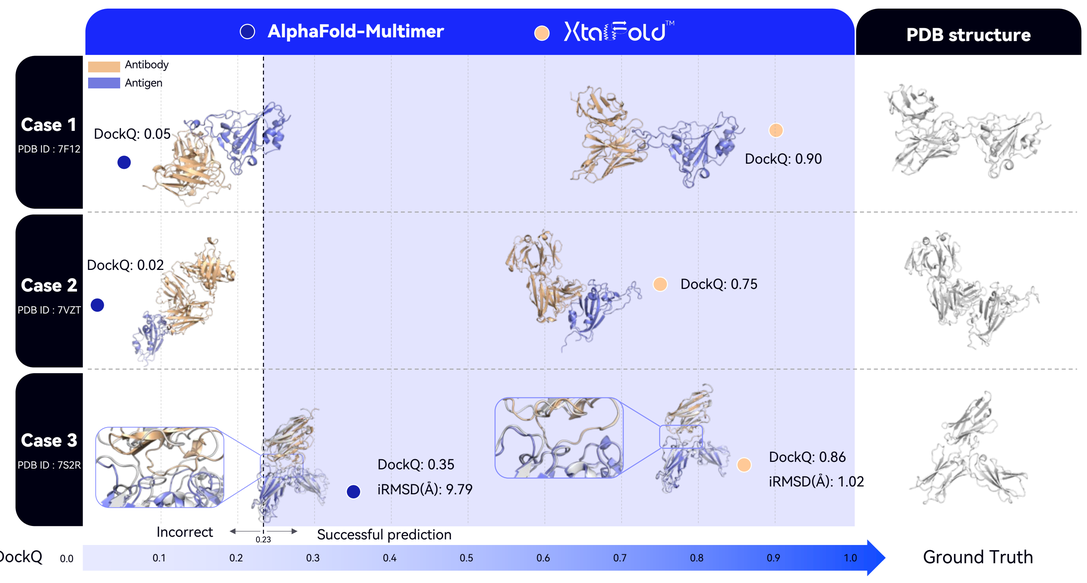
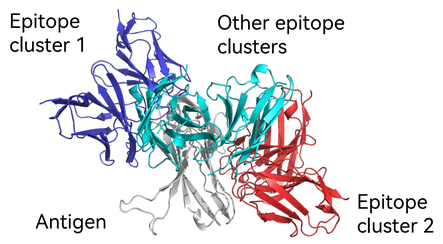


Figure 3. Prediction results of XtalFold™ and AlphaFold-Multimer for three antigen-antibody (Fv/Fab/VHH) complexes. DockQ: higher values indicate better predictions; DockQ ≥ 0.80 is considered high-quality, 0.8 > DockQ ≥ 0.49 is considered medium-quality, 0.49 > DockQ ≥ 0.23 is considered acceptable-quality, and 0.23 > DockQ ≥ 0 is considered incorrect prediction. Successful prediction (DockQ ≥ 0.23) includes acceptable, medium, and high quality.

Leveraging XtalFold™ in next-gen biologics discovery

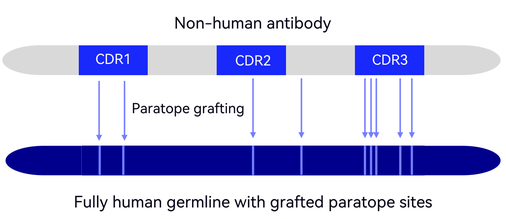
✓ High-throughput epitope mapping

The epitopes of 36 mAbs were mapped onto a tumor-associated antigen and grouped into several clusters, including two distinct non-overlapping clusters. Accurate epitope information can facilitate the design of biparatopic antibodies.



✓ Super-humanization

XtalFold™ identifies paratope residues that are subject to in silico evaluation of their relevance for binding affinity and structural integrity. Only the most important murine residues are grafted into a fully human germline to improve humanness and minimize immunogenicity risk.



✓ Improving developability

CDRs may contain high-risk hotspots that contribute to low developability, such as aggregation propensity and PTMs. Accurate identification of key paratope residues by XtalFold™ allow for precise engineering that avoids key binding residues

- ✓ **Affinity maturation**
- ✓ **pH selectivity engineering**
- ✓ **Cross-species engineering**
- ...

Reference

[1] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. Nature, 2021, 596(7873): 583-589.
[2] Evans R, O'Neill M, Pritzel A, et al. Protein complex prediction with AlphaFold-Multimer[J]. biorxiv, 2021: 2021.10. 04.463034.
[3] Ruffolo J A, Chu L S, Mahajan S P, et al. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies[J]. Nature communications, 2023, 14(1): 2389.
[4] Wu R, Ding F, Wang R, et al. High-resolution de novo structure prediction from primary sequence[J]. BioRxiv, 2022: 2022.07. 21.500999.
[5] Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model[J]. Science, 2023, 379(6637): 1123-1130.
[6] Yin R, Pierce B G. Evaluation of AlphaFold Antibody-Antigen Modeling with Implications for Improving Predictive Accuracy[J]. bioRxiv, 2023: 2023.07. 05.547832.

Scan the QR code to learn more about Ailux Biologics

Scan this QR code to view this poster online