OXFORD

Briefings in Bioinformatics, 2024, **25(5)**, bbae445

https://doi.org/10.1093/bib/bbae445

Problem Solving Protocol

AI-based IsAb2.0 for antibody design

Tianjian Liang[1,2,3,4,5,‡], Ze-Yu Sun[1,2,3,4,5,‡], Margaret G. Hines[6,‡], Kerri Jo Penrose[7], Yixuan Hao[1,2,3,4,5], Xiaojie Chu[6], John W. Mellors[6,7], Dimiter S. Dimitrov[6], Xiang-Qun Xie[1,2,3,4,5], Wei Li[6,*], Zhiwei Feng[1,2,3,4,5,*]

[1]Department of Pharmaceutical Sciences, Computational Chemical Genomics Screening Center, Pharmacometrics & System Pharmacology PharmacoAnalytics, School of Pharmacy, University of Pittsburgh, 335 Sutherland Drive, Pittsburgh, PA 15261, United States
[2]National Center of Excellence for Computational Drug Abuse Research, University of Pittsburgh, 3501 Terrace St, Pittsburgh, PA 15261, United States
[3]Drug Discovery Institute, University of Pittsburgh, 3501 Terrace St, Pittsburgh, PA 15261, United States
[4]Department of Computational Biology, School of Medicine, University of Pittsburgh, 3420 Forbes Avenue, Pittsburgh, PA 15261, United States
[5]Department of Structural Biology, School of Medicine, University of Pittsburgh, 3501 Fifth Ave, Pittsburgh, PA 15261, United States
[6]Division of Infectious Diseases, Department of Medicine, Center for Antibody Therapeutics, School of Medicine, University of Pittsburgh, 3550 Terrace Street, Pittsburgh, PA, United States
[7]Division of Infectious Diseases, Department of Medicine, Center for AIDS Research, School of Medicine, University of Pittsburgh, 3550 Terrace Street, Pittsburgh, PA, United States

*Corresponding authors. Wei Li, Department of Medicine, Center for Antibody Therapeutics, School of Medicine, University of Pittsburgh, 3550 Terrace Street, Pittsburgh, PA, United States. E-mail: liwei171@pitt.edu; Zhiwei Feng, Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, 3501 Terrace St, Pittsburgh, PA 15261, United States. E-mail: zhf11@pitt.edu

‡The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint first authors.

Abstract

Therapeutic antibody design has garnered widespread attention, highlighting its interdisciplinary importance. Advancements in technology emphasize the critical role of designing nanobodies and humanized antibodies in antibody engineering. However, current experimental methods are costly and time-consuming. Computational approaches, while progressing, faced limitations due to insufficient structural data and the absence of a standardized protocol. To tackle these challenges, our lab previously developed IsAb1.0, an *in silico* antibody design protocol. Yet, IsAb1.0 lacked accuracy, had a complex procedure, and required extensive antibody bioinformation. Moreover, it overlooked nanobody and humanized antibody design, hindering therapeutic antibody development. Building upon IsAb1.0, we enhanced our design protocol with artificial intelligence methods to create IsAb2.0. IsAb2.0 utilized AlphaFold-Multimer (2.3/3.0) for accurate modeling and complex construction without templates and employed the precise FlexddG method for *in silico* antibody optimization. Validated through optimization of a humanized nanobody J3 (HuJ3) targeting HIV-1 gp120, IsAb2.0 predicted five mutations that can improve HuJ3-gp120 binding affinity. These predictions were confirmed by commercial software and validated through binding and neutralization assays. IsAb2.0 streamlined antibody design, offering insights into future techniques to accelerate immunotherapy development.

Keywords: IsAb; antibody design; AlphaFold-Multimer; deep learning; humanized antibody; HIV-1 nanobody

Introduction

Antibodies are derived from plasma cells and play a key role in the immune system. Because of their high specificity as well as their high affinity [1] to a large variety of macromolecules, antibodies are effective in treating cancer [2], viral infections [3], addiction [4], autoimmune diseases [5], and are of particular interest when considering for the current coronavirus disease 2019 pandemic [6]. With the development of antibody engineering technology, nanobodies and humanized antibodies have been identified as potential tools for the treatment of these conditions as they exhibit significant potential in therapeutics. Nanobodies are the recombinant variable domains of heavy-chain only antibodies [7]. Their unique structure imbues them with features like small size, increased stability, excellent solubility, and the capacity for deep tissue penetration. All of these properties have led to the use of nanobodies in molecular imaging and therapy for diseases [8] such as, cancer [9], viral infections [10], and autoimmune diseases [11]. However, a large number of xenogenetically sourced nanobodies and antibodies with specificity for clinically relevant antigens cannot be used in the clinical applications due to their high immunogenicity in humans [12]. Therefore, humanization is an important method for reducing immunogenicity of xenogeneic antibodies and improving their activation in human immune system. As such, it is apparent that current antibody design techniques must be improved to evaluate these antibody derivatives.

Currently, there are many techniques available for empirical antibody rational design, but these tools come with significant limitations. For example, X-ray crystallography and electron microscopy-based methods [13] are used to obtain 3D antibody structures, and phage display libraries are available to optimize affinity. However, these methods are time-consuming, expensive, and labor-intensive. Computational tools or algorithms are complementary methods that can reduce the time and cost required for antibody design by decreasing the number of failures

Received: May 31, 2024. **Revised:** August 6, 2024. **Accepted:** August 27, 2024

and promoting the success rate of the experimental test [14]. Despite recent progress in these computational methods, several factors limiting their efficacy must be addressed. First, most antibody design tools are stand-alone and there is no free comprehensive antibody design protocol that provides users with detailed instruction from beginning to end. This requires users to spend considerable time creating a complete workflow of antibody design. Many current antibody design tools also require antibody–antigen complexes as an input for affinity maturation [15–17]. This information is not readily available for many complexes of therapeutic interest. Therefore, our lab previously established an *in silico* antibody design protocol IsAb1.0 [18]. However, the procedure of IsAb1.0 is not user-friendly enough. IsAb1.0 utilized homology modeling, requiring the input antibody to have homologous structures that are often unavailable for the novel antibodies. To ensure global docking accuracy, IsAb1.0 also requires users to input epitopes information, which is difficult to obtain in most cases. Finally, IsAb1.0 fails to consider the design of nanobodies and humanized antibodies, effectively excluding these potent molecules from consideration.

Recently, artificial intelligence (AI)-based and physics-based methods have made great progress in antibody engineering [19, 20]. Evans *et al.* [21] developed a new model called AlphaFold-Multimer by training AlphaFold, a deep learning-based tool for predicting single-chain protein structures, with multimeric inputs, incorporating support for multichain featurization and symmetry handling. AlphaFold-Multimer leverages Evoformer to represent pairwise relations between the different amino acids in the protein and multiple sequence alignment (MSA). The pair representation is used to predict the relative distances between the amino acids in the protein. The MSA information combines with pair representation to predict the final structure. AlphaFold-Multimer has been proved to produce high-accuracy predictions in protein complexes. Barlow *et al.* [22] developed FlexddG, a physics-based method to predict protein complex binding affinity changes upon mutation. FlexddG uses "backrub" to generate an ensemble of models, followed by torsion minimization, side chain repacking, and averaging these two processes to approximate interface $\Delta\Delta G$ values.

By creating an antibody-design protocol that can overcome these issues, the development of antibody therapies could be significantly improved and more importantly, accelerated. One such case in which this increased speed is essential is the treatment of human immunodeficiency virus type 1 (HIV-1). HIV-1 destroys CD4+ lymphocytes in their respective hosts, resulting in the development of acquired immunodeficiency syndrome (AIDS) [23,24]. The HIV-1 exterior envelope glycoprotein, gp120, binds to the T-cell surface receptor, CD4 [25,26], which triggers the fusion of the viral and cell membranes [27]. The receptor CD4 binding site is a conserved epitope that can be targeted by a wealth of broadly neutralizing antibodies that block the interaction of gp120 and CD4. Recently, Huang *et al.* [28] identified a llama VHH nanobody called J3 through neutralization screening of a phagemid VHH library. J3 exhibited a much broader and more potent neutralization ability than other similar nanobodies. J3 also displayed a range of neutralization abilities that can neutralize over 95% of circulation HIV-1 strains and its IC50 reaches 0.256 μg/ml, much higher than soluble CD4 and VRC01 [29]. While J3 could be a powerful tool against HIV-1, it requires humanization before use in therapeutic constructs in humans to avoid/decrease immunogenicity. However, antibody humanization typically results in reducing affinity, so affinity maturation after humanization is warranted.

Hence, our lab incorporated AlphaFold-Multimer2.3/3.0 and FlexddG to create a novel comprehensive *in silico* antibody design protocol called IsAb2.0. This new protocol is more accurate, and the procedure is more concise compared with the one previously developed by our lab. The protocol can also be applied to nanobody and humanized antibody design, meeting the demand for therapeutic antibodies in the treatment of complex diseases. IsAb2.0 only requires users to input sequences of the antibody and antigen. AlphaFold-Multimer first models the antibody–antigen complex; the output 3D structure serves as the possible binding pose. Then, the protocol applies SnugDock to refine the possible binding poses and outputs the final result. With the binding pose of the complex resolved, the protocol performs alanine scanning to predict the hotspots (or key residues) of the antibody that mediate antigen binding. Such information provides users with valuable insights for future antibody affinity engineering. Finally, FlexddG is conducted to perform single point mutation on antibody to improve its binding affinity and other properties.

In this study, we built IsAb2.0 based on the protocol previously established by our lab and validated it by optimizing HuJ3. To generate humanized J3 (HuJ3), we first humanized the nanobody J3 and found that the HuJ3 compromised HIV-1 Env binding and neutralization potency by three to five folds. Then, we utilized IsAb2.0 to model the 3D structure of HuJ3-gp120 complex and select point mutations that could improve HIV-1 neutralization. We also used BioLuminate [30–32] from Schrödinger to further assess the performance of the protocol. Finally, the predicted results from IsAb2.0 were validated by experimental methods, including wet enzyme-linked immunosorbent assay (ELISA) and HIV-1 neutralization assays. Using IsAb2.0, we successfully improved HuJ3 affinity by introducing a single point mutation, E44R, demonstrating the versatility of our antibody design protocol. Compared with IsAb1.0, we improved our design protocol mainly by (i) applying AlphaFold-Multimer2.3/3.0 to accurately construct the 3D structure of the antibody–antigen complex, and (ii) applying a more accurate method, FlexddG, to conduct *in silico* antibody single point mutation.

Method

Antibody–antigen structural complex generation

The COSMIC^2 server (https://cosmic-cryoem.org/) [33], a cloud-based computational platform developed by the University of Michigan for structural biology projects and openly accessible to academic researchers, facilitated the generation of an antibody–antigen structural complex. Initially, the protein sequence for gp120 Clade C was obtained from the Protein Data Bank (PDB: 7ri1) (https://www.rcsb.org/) and uploaded onto COSMIC^2, along with the sequence files for HuJ3 and gp120 provided in the Supplementary Material. Subsequently, an AlphaFold2 job was initiated on the server. Since the AlphaFold toolkit on COSMIC^2 does not provide configuration of the number of generated models, the "number of predictions per model" was set to 10 to enhance prediction diversity. Detailed submission parameters are documented in the Supplementary Material.

Structure refinement and local docking of antibody–antigen complex

The initial binding pose generated by AlphaFold-Multimer was refined using SnugDock ensemble docking [34–36]. Due to the inaccuracy of AlphaFold-Multimer in modeling protein secondary structures, crystal structure of gp120 from PDB:7ri1 and the 3D structure of HuJ3, modeled by SWISS-MODEL (template PDB: 7ri1),

were aligned to the complex generated by AlphaFold-Multimer and substituted for their corresponding components. This refined binding pose was then used as the input for SnugDock. The FastRelax function in Rosetta was employed to resolve potential clashes and identify energetically favorable conformations of the crystal gp120 and modeled HuJ3. Ten decoys were generated for both the crystal gp120 and modeled HuJ3, which were used as ensembles to enable flexibility of their backbone. These 10 relaxed gp120 and HuJ3 structures were prepacked using the Rosetta Prepack function to ensure low-energy starting side-chain conformations. Subsequently, the SnugDock function was used to dock the crystal gp120 ensemble with the modeled HuJ3 ensemble. SnugDock initiated the docking with a random perturbation of 3 Å translation and 8° rotation in each Cartesian direction. The Motif Dock Score was utilized during the low-resolution docking phase [37]. One thousand decoys were generated and ranked by their interface score(I_sc). The formation of the docking funnel involved the criteria of whether the local docking was successful or not. Once the local docking was successful, the lowest I_sc result was chosen as the final local docking result.

Computational alanine scanning and point mutation

Possible hotspots (or key residues) on the antibody were predicted by Rosetta alanine scanning program [38–40] (AlaScan.xml, https://github.com/Kortemme-Lab/ddg/, Fig. S1). The distance cutoff value of interface residues was set to 5 Å. Among the alanine scanning results, the residues whose ΔΔG higher than 1 kcal/mol were selected as hotspots on the antibody.

The "define_interface.py" was used to find out the interface residues on antibody (Fig. S1). FlexddG program [22] (https://github.com/Kortemme-Lab/flex_ddG_tutorial, Fig. S1) was applied to perform single point mutation on the antibody interface residues to increase the binding affinity of HuJ3-gp120 complex. The complex we wished to mutate was placed in the "inputs" folder. The antibody chain id was specified in the "chains_to_move.txt" file located in the "inputs" folder. FlexddG only allows users to mutate one residue at a time, so the "run_example_2.py" script was modified to specify the residue chain id and number we wanted to mutate each time. The parameters used to run the program were set as the recommended values. The "analyze_flex_ddG.py" file was used to analyze the output results from "run_example_2.py", which printed the wild type and mutant interface ΔG. ΔΔG score and the ΔΔG score reweighted with the fitted GAM model were calculated and listed in a .csv file through "analyze_flex_ddG.py". The mutant whose reweighted ΔΔG score was lower than 0 kcal/mol was accepted as the mutation which may increase the binding affinity of the complex.

BioLuminate point mutation validation

BioLuminate software was downloaded from the Schrödinger website. The target complex structure was first imported to BioLuminate. Then, "protein preparation and refinement" was used to fix the structure (optimize orientations of hydrogen-bonded groups, delete waters, and minimize the structure). The parameters used the default setting. The prepared structure was then imported to the "Residue Scanning" panel and the "Calculation type" selected the "Stability and Affinity" mode. In the residues table, the residues predicted by our protocol were selected and mutated to the potential mutations. The other parameters used the default settings. The mutations with affinity scores smaller than 0 kcal/mol were accepted.

Humanization of nanobody J3

Llama nanobody J3 was humanized by grafting its complementarity determining regions (CRDs) (CDR1–3) residues to the closest human Ig germline variable domain (VH) family VH3–23 scaffold. However, not all framework region (FR) residues are human. We deliberatively kept specific llama-derived FR residues, which include the vennerier region and packing residues that can modulate CDRs conformations and impact antigen binding. In addition, the "tetrad" in FR2, which is important to the nanobody stability and solubility, was left in its native state. Through this process, we obtained huJ3 with a 13 amino acid difference from the original nanobody J3. The "humanness" prediction results indicated the HuJ3 displayed a higher sequence similarity to naïve human antibody repertoire than the original J3, indicating a reduced risk of inducing immunogenicity in human applications.

Experimental validation of point mutation

All of the mutations of HuJ3 were selected for experimental validation. Their mutation sequence alignment was shown in supporting Fig. S2. These mutants were recombinantly expressed by the phagemid pComb3x [41] (Fig. S3), PComb3x was modified based on the original plasmid pComb3XSS, which is purchased from Addgene (Plasmid #63890). In the pComb3X, the lac promoter is used for transcription initiation followed by an OmpA signal peptide directing VH proteins secretion into periplasm. The six consecutive histidine tag and the following Flag tag were used for protein purification and detection. The pComb3x containing huJ3 was constructed by molecular subcloning to insert the huJ3 gene into the Sfi I linearized pComb3x backbone using the enzymes Sfi I and T4 ligase. The mutants plasmids were made through site-directed mutagenesis at BonOpus. Competent HB2151 *Escherichia coli* was transformed with these plasmids to express VH nanobodies. Antibodies were purified from *E. coli* periplasm after polumyxin B treatment followed by Ni-NTA chromatography. The molecular size and purity was verified by SDS-PAGE in Fig. S4. ELISA was performed to determine quantitative binding of each variant to the HIV-1 envelope glycoprotein. Transiently expressed gp140, a functional homolog of gp120 used to evaluate HIV-1 binding *in vitro*, served as the antigen for evaluation. Plates were coated with 50 ng gp140/well at 4°C overnight and subsequently blocked with 3% milk for 1 h at room temperature. Primary VH nanobodies, including the original J3, HuJ3, and each of the five variants, were serially diluted in 3% milk to achieve an 8-concentration gradient ranging from 1000 to 0.0128 nM, then incubated on the coated and blocked plate for 2 h at room temperature. After incubation, the plate was washed 4 times by 0.05% PBST. The detection antibody, anti-FLAG M2-peroxidase (A8592, Sig-ma-Aldrich, St. Louis, MO, USA), was added and incubated for 1 h at room temperature. The plate was washed another 4 times by 0.05% PBST. Binding activity was detected using 3,3′,5,5′-tetramethylbenzidine (TMB, Sigma-Aldrich, St. Louis, MO, USA). The reaction was stopped after 2 min by TMB stop buffer (ScyTek Laboratories, Logan, UT, USA) to prevent oversaturation and absorbance was read at 450 nm. The experiment was performed in duplicate. After analysis, another ELISA was performed in the same manner as above to validate and compare the gp140 affinity of variant E44R with J3 and HuJ3.

E44R was further validated in the laboratory-developed TZM-bl HIV-1 Phenotyping Assay to measure the susceptibility of HIV-1 (group M, subtype B, isolate BAL) to E44R compared with HuJ3. TZM-bl cells are an indicator cell line that allows quantitative analysis of HIV replication. TZM-bl cells were generated from

HeLa cells that stably express large amounts of CD4 and CCR5 and have separately integrated copies of the luciferase and ß-galactosidase genes under control of the HIV-1 promoter. These cells naturally express CXCR4 receptors. TZM-bl cells were plated at 10 000 cells per well overnight. The cells were treated with serial dilutions of HuJ3 antibodies, and infected with a dilution of infectious HIV-1 virus normalized to an output of 140 000 relative light units (RLU) as determined by endpoint dilution. After a 48-h incubation at 37°C, cells were lysed, and luminescence was measured in RLU using a commercially available luciferase detection system. The 50% *in vitro* concentration (IC_{50}) was calculated as the concentration of antibodies needed to inhibit 50% of HIV-1 replication in the assay. A batch control virus was run with each experimental setup.

HuJ3 designed by IsAb1.0

The crystal structure of gp120 CladeC (PDB: 7ri1) was obtained from the PDB (https://www.rcsb.org/). The sequence of HuJ3 was submitted to the SWISS-MODEL web server [42] (https://swissmodel.expasy.org/interactive). VHH nanobody J3 (PDB: 7ri1) was selected as the template to model the HuJ3.

ClusPro [43–46] was employed to generate the potential binding poses of the HuJ3-gp120 complex. To specify the antibody docking, "Antibody Mode" [47] was chosen. HuJ3 was designated as the receptor and gp120 as the ligand. Paratopes and epitopes were entered into the attraction sections of the receptor and ligand, respectively, to set the docking constraints. The cluster among the top 10 global docking results that closely resembled the binding pose of PDB: 7ri1 was selected as the potential binding pose.

The selected binding pose from global docking was input into the SnugDock function on the ROSIE web server, utilizing "thorough mode" to refine the structure. The success of the local docking is based on the formation of the docking funnel. Once the local docking was successful, the lowest I_sc result was chosen as the final local docking result.

Potential hotspots (or key residues) on the antibody were identified using the Rosetta alanine scanning program (AlaScan.xml). The distance cutoff for interface residues was set to 5 Å. Among the alanine scanning results, residues with a ΔΔG higher than 1 kcal/mol were identified as hotspots on the antibody.

Single State Design protocol was obtained from Dr. Jens Meiler's lab website. Initially, 'define_interface.py' was utilized to prepare a residue file (resfile), which defined the interface residues should be mutated. Subsequently, the Single State Design protocol (design.xml) was employed to modify the antibody.

Results

Workflow of IsAb 2.0

The general procedure of IsAb2.0 is outlined in Fig. 1. In the first step, users input sequences of the antibody and antigen into the protocol, with antibody sequences retrievable from the IMGT (https://www.imgt.org/) database. In Step 2, AlphaFold-Multimer2.3/3.0 generates 3D structures of the antibody, antigen, and their complex, effectively combining the functions of homology modeling and global docking from IsAb1.0. The per-residue confidence metric (pLDDT) is used to evaluate the quality of the model generated by AlphaFold-Multimer, assessing local accuracy. If the pLDDT scores are below 70, the antibody–antigen complex undergoes further refinement. Otherwise, the complex proceeds to Step 4 for local docking. In Step 3, several methods can be used for structural refinement. If the structure with lower pLDDT score has a crystal structure, this crystal structure will

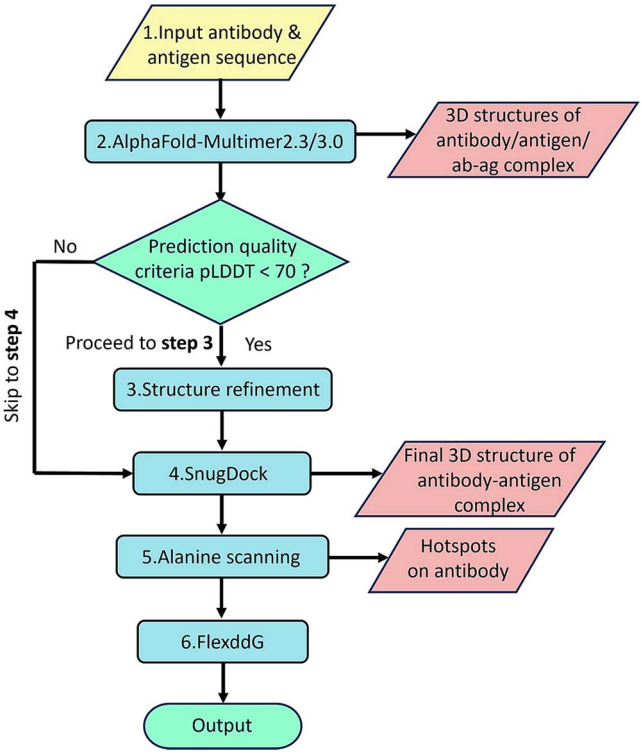

Figure 1. The workflow of the IsAb2.0; users are required to input the sequences of antibody and antigen into IsAb2.0 protocol; AlphaFold-Multimer 2.3/3.0 generates 3D structure of antibody–antigen complex based on their sequences; if the prediction quality (pLDDT) of complex is below 70, the structure undergoes refinement; SnugDock is then used to refine the antibody–antigen complex, followed by alanine scanning and FlexddG for hotspot identification and antibody optimization.

be optimized by Rosetta FastRelax to resolve potential clashes and identify energetically favorable conformations. The relaxed structures will then replace the model generated by AlphaFold-Multimer. If a crystal structure is unavailable, SWISS-MODEL will generate a new 3D homology model for low pLDDT structure and replace it. In Step 4, SnugDock performs local docking for the antibody and antigen, allowing flexibility of the CDR loops and interfacial side chains. SnugDock refines the potential binding poses provided by AlphaFold-Multimer and outputs the final antibody–antigen complex. In Step 5, after obtaining the 3D structure of antibody–antigen complex, alanine scanning is performed to predict possible hotspots on the antibody, facilitating future antibody design. Alanine scanning mutates the interface residues to alanine and calculates the change in energy to identify hotspots. In Step 6, point mutations are applied to mutate the antibody interface residues to the remaining 17 amino acids and identify mutations that can improve the binding affinity of the complex. The antibody–antigen complex is imported into FlexddG to perform point mutations, identifying mutations that enhance binding affinity.

Antibody–antigen structural complex modeling

As a result, we acquired 50 decoys along with their respective predicted local distance difference test (pLDDT) scores for subsequent analysis. By ranking the decoys based on their interface predicted template modeling and predicted template modeling (ipTM+pTM) scores, we identified the top 10 models with the highest scores. Remarkably, nine of these top models shared an identical binding position with J3-gp120. The model with the

Table 1. pLDDT and ipTM + pTM value for top 10 AlphaFold-Multimer predictions

Top 10 predictions	Average pLDDT	ipTM + pTM
Model_3_Pred_5	94.8754	0.9347
Model_3_Pred_2	94.6905	0.9325
Model_3_Pred_9	94.4099	0.9288
Model_3_Pred_3	94.3835	0.9277
Model_3_Pred_7	94.3328	0.9284
Model_3_Pred_4	94.2564	0.9269
Model_3_Pred_0	94.0165	0.9229
Model_3_Pred_6	93.9752	0.9252
Model_5_Pred_7	93.8252	0.9252
Model_1_Pred_2	91.6369	0.6868

Model_3_Pred_5 has the highest average pLDDT score and was selected as the result of AlphaFold-Multimer.

highest pLDDT and ipTM + pTM scores was selected as the likely binding pose. However, upon aligning the gp120 modeled by AlphaFold-Multimer with the crystal gp120, it was revealed that even the best model, Model_3_Pred_5, contained some structural errors (Fig. S5). These errors included a loss of alpha helix between residues LEU369-THR372 and two structural errors in the beta-sheet regions, specifically between residues LYS419-LYS421 and ASP464-GLU466. The pLDTT plot indicated that all these regions had lower pLDTT scores. For HuJ3, since no crystal structure is available, the model generated by AlphaFold-Multimer was evaluated based on the pLDDT scores. According to the pLDTT plot (Fig. S6), the pLDTT value in CDR1 between residues PHE29-GLN31 was lower than 60, indicating low confidence in this region. Structure analysis suggested that AlphaFold-Multimer may have failed to predict the alpha helix in this region. Similarly, the CDR3 residues SER105-GLY107 also had lower pLDTT scores. The structural analysis showed that this region may have failed in modeling an alpha helix. Since both CDR1 and CDR3 are important in the local docking step. Hence, structure refinement was conducted to correct the structure of these regions. Additionally, Table 1 summarizes the average pLDDT and ipTM+pTM scores for each model. Moreover, detailed pLDDT data for each residue was plotted in Fig. S6, enhancing the comprehensive overview of model performance metrics.

HuJ3-gp120 local docking

AlphaFold-Multimer encountered challenges in accurately modeling protein secondary structures. Consequently, the crystal structures of gp120 and modeled HuJ3 generated by SWISS-MODEL were used to replace the corresponding in the complex modeled by AlphaFold-Multimer. Then, using the complex binding pose to serve as starting structure to the local docking. The SnugDock ensemble docking was employed to perform local docking searches using backbone ensembles, which has been proved to improve docking accuracy. This ensemble docking approach relies on the conformational-selection mechanism for protein docking, utilizing a pre-generated ensemble of protein partners. The ensembles were generated using Rosetta Relax to sample various backbone conformations during docking. In the low-resolution stage, each docking run involves rigid-body translation and rotation around the protein partner, along with backbone swapping from the pre-generated ensemble. This allows for the sampling of diverse backbone conformations. In the high-resolution stage, an all-atom refinement is applied to the generated encounter complex, and the side-chains at the

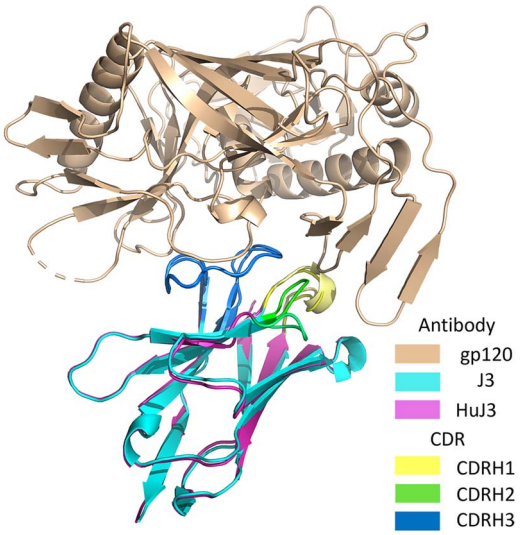

Figure 2. Comparison between HuJ3-gp120 and J3-gp120 3D structures. We overlapped the HuJ3-gp120 complex generated by IsAb2.0 with the J3-gp120 crystal complex. The general binding poses of HuJ3 and J3 with gp120 are highly similar, especially their CDR loops.

interface are packed for optimal binding. The interface root means square deviation (I_rmsd) of heavy atoms in the interface residues between the reference structure and resulting structure was used to evaluate the performance of local docking. According to the Critical Assessment of Protein Interactions [38], docking results with an I_rmsd value lower than 4 Å are considered "near-native" structure. According to the criteria from SnugDock [48], the robust evaluation of successful local docking is the presence of a "docking funnel", in which the "near-native" result (low I_rmsd) has lower energy (low I_sc) than the non-native result [49]. In this case, three of the five lowest interface score (intermolecular energy/I_sc) results [50] have an I_rmsd value lower than 4 Å ($N_5 >= 3$), and the performance will be identified as successful docking. From the I_sc versus I_rmsd plot (Fig. S7), we found that all the five lowest I_sc results had I_rmsd lower than 4 Å, which meant that this local docking was successful and the native HuJ3-gp120 binding pose has a higher chance of being similar to the binding pose we predicted. As for the final result, we chose the lowest I_sc result as our final binding pose of the HuJ3-gp120 complex. The predicted HuJ3-gp120 complex also overlapped with the J3-gp120 crystal structure and showed that the binding pose of HuJ3 to gp120 was highly similar to J3 (Fig. 2).

After the structural analysis of CD4-gp120 and J3-gp120, we found that residues D368, N425, M426, and V430 on gp120 interacted with both J3 and CD4. In this case, we proposed that these residues may be the epitopes on gp120. In addition, the binding site on gp120 of J3 and HuJ3 are the same as CD4. Therefore, we hypothesized that these epitopes on gp120 will form a connection with HuJ3. Structural analysis of the predicted HuJ3-gp120 complex (Fig. 3) revealed that gp120 formed 12 hydrogen bonds with HuJ3 (Table 2). The structure analysis of the HuJ3-gp120 complex revealed that D368 on gp120 formed hydrogen bonds with H56 and I101 (HuJ3). N425, M426, and V430 formed hydrogen bonds with Y100, T99, and N30 on HuJ3, respectively. Compared with Fig. 3B and D, we noticed that both these interactions could be found on the J3-gp120 and CD4-gp120 complexes, which matched with the hypothesis we made previously and increased the reliability of the HuJ3-gp120 complex we predicted.

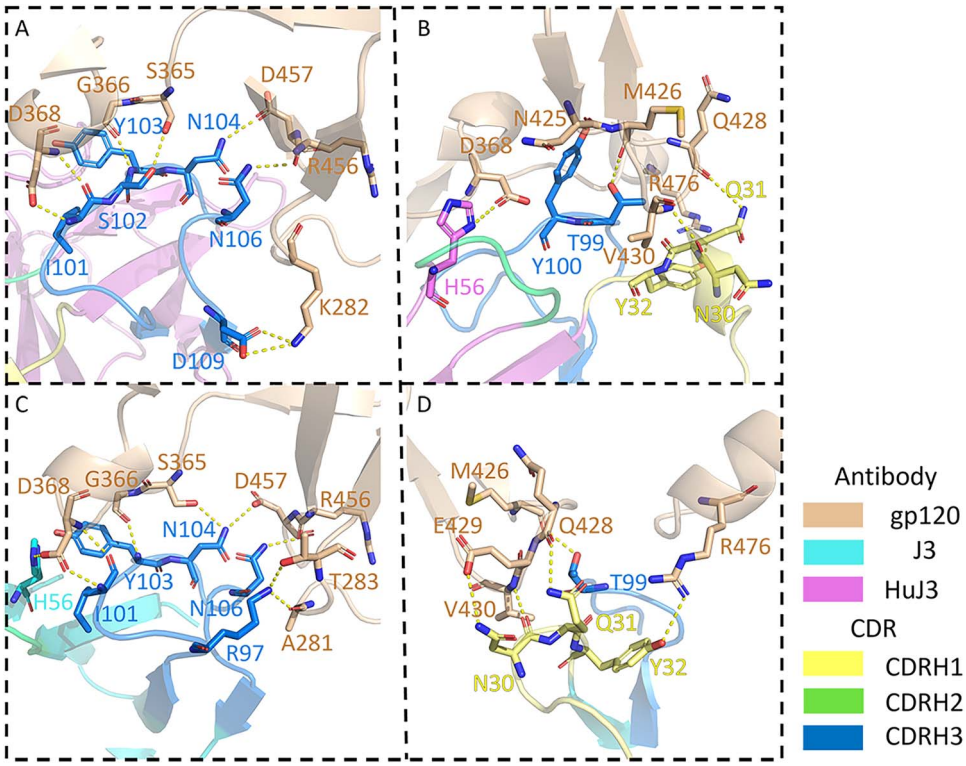

Figure 3. Comparison of hydrogen bonds between HuJ3 and J3 with gp120. Yellow dash line is hydrogen bond. (A) and (B) are the hydrogen bonds on HuJ3-gp120. (C) and (D) are the hydrogen bonds on J3-gp120. Most of the residue pairs forming hydrogen bonds are the same between HuJ3-gp120 and J3-gp120 complexes. Predicted epitopes of gp120 (D368, N425, M426, and V430) form hydrogen bonds with HuJ3.

Table 2. Residue pairs forming hydrogen bonds on the HuJ3-gp120 and J3-gp120 complexes and the number of hydrogen bonds they form

Residues on HuJ3	Residues on gp120	Number of hydrogen bonds
N30	V430	1
Y32	Q476	1
H56	D368	1
T99	M426	1
Y100	N425	1
I101	D368	2
Y103	G366	1
N104	D457	1
N106	R456	1
D109	K282	2
Residues on J3	Residues on gp120	
N30	E429/V430	2
Q31	Q428	1
Y32	R476	1
H56	D368	1
K97	A281/T283	2
T99	M426	1
I101	D368	2
Y103	G366	1
N104	S365/D457	2

HuJ3-gp120 alanine scanning and point mutation

Computational alanine scanning was applied to predict the hotspots on the HuJ3-gp120 interface. Based on the criteria that neutral residues and hotspots were defined as ΔΔG less or more than 1 kcal/mol during the mutation to alanine, six residues were defined as possible hotspots (Table 3).

Table 3. Hotspots predicted by IsAb2.0 and their ΔΔG

Hotspots	ΔΔG (kcal/mol)
Y32	1.00
H56	3.11
T99	2.50
Y100	6.51
Y103	1.79
D109	3.13

Residues whose ΔΔG larger than 1 kcal/Mol are considered as the hotspots.

To perform point mutation for the HuJ3-gp120 complex, the 'define_interface.py' was used to find the interface residues on HuJ3. The interface cutoff value was defined as 5 Å, and 18 residues on HuJ3 were identified as interface residues. Since previous experiments have tested mutations on CDR1 and Framework Region 1, our current focus is on mutating CDR2, CDR3 and interface residues. All these residues were input into FlexddG program and mutated to the remaining 17 amino acids. The mutants whose ΔΔG smaller than 0 kcal/mol were filtered from the FlexddG results, and E44, V54, S98, T99, and S102 were all identified as possible mutation sites. Structure rational analysis was used to reduce the number of possible mutations. ΔΔG of E44R was −0.394 kcal/mol, the second lowest mutation among the other E44 substitutions. Given that E44 is located within a negatively charged pocket, mutating it to a positively charged residue could enhance electrostatic interactions favorably (Fig. 4A). V54E had ΔΔG value −0.299 kcal/mol and E54 formed two hydrogen bonds with R432 on gp120 (Fig. 4B). Because S98 was in a hydrophobic pocket, it mutated into a hydrophobic residue which could imbue the complex with better electrostatic properties (Fig. 4C). T99R

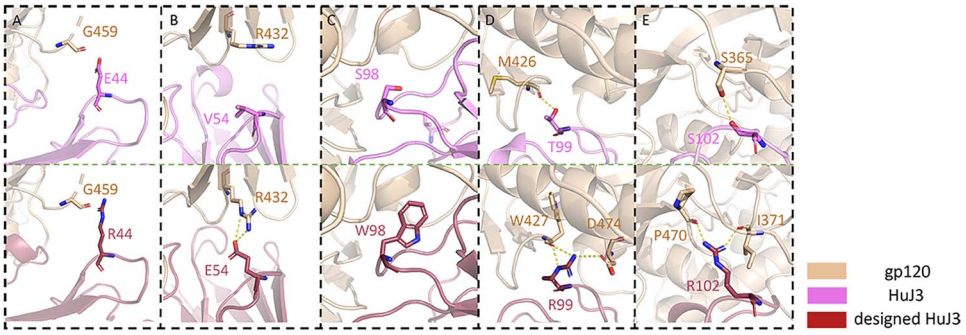

Figure 4. Comparison of hydrogen bonds between HuJ3 and designed HuJ3 with gp120. The yellow dashed line is the hydrogen bond. (A) E44R mutation, R44 enhances electrostatic interactions favorably. (B) V54E mutation, E54 forms two hydrogen bonds with R432 compared with V54. (C) S98W mutation, W98 increases the electrostatic property of the HuJ3-gp120 complex. (D) T99R mutation, T99 forms a hydrogen bond with M426, and R99 forms two hydrogen bonds with W427 and one with D474. (E) S102R mutation, R102 forms hydrogen bonds with P470 and I371 on gp120, respectively.

Table 4. Five potential mutations repeated by BioLuminate

Mutations	Δ Affinity (kcal/mol)
S102R	−23.19
T99R	−12.63
S98W	−4.09
V54E	2.52
E44R	−3.48

Δ affinity smaller than 0 kcal/Mol indicating that the mutation can increase binding affinity of the complex. Four of the five potential mutations predicted by IsAb2.0 have the same results as BioLuminate.

Table 5. Comparison of mutation energy predicted by FlexddG and binding affinity data (EC50) determined by ELISA

Mutations	ΔΔG (kcal/mol)	EC50 (nM)
E44R	−0.394	1.696
V54E	−0.299	2.305
S98W	−0.621	N/A
T99R	−0.979	N/A
S102R	−1.752	10.36

S98W showed no binding to the antigen, while T99R failed to express; therefore, there are no EC50 data for these two variants.

had the lowest ΔΔG value −0.979 kcal/mol among the other substitutions. Compared with T99 forming a hydrogen bond with gp120, R99 formed two hydrogen bonds with W427 and one with D474, which made the binding between HuJ3 and gp120 more stable (Fig. 4D). S102R connected with P470 and I371 on gp120 by forming a hydrogen bond with them respectively, its ΔΔG value −1.752 kcal/mol was the lowest among the other substitutions (Fig. 4E). Overall, E44R, V54E, S98W, T99R, and S102R were selected as the potential mutations.

To increase the reliability of our predictions, the powerful protein engineering software BioLuminate from Schrödinger was applied to test the five predicted mutants. Each potential mutation was performed by BioLuminate again. According to BioLuminate result analysis rules, the 'Δ Affinity' score lower than zero means the mutant binds more effectively than the original protein. Four of the five mutations predicted by our protocol with 'Δ Affinity' lower than 0 (Table 4), indicating that four of the five mutations predicted by the protocol and BioLuminate had the same results.

ELISA validation (Fig. 5B) showed that E44R had better binding affinity than HuJ3, which is consistent to the predicted result. V54E had the same binding affinity as HuJ3. The T99R mutant protein failed to express. S98W and S102R exhibit lower binding affinity than HuJ3. HIV-1 neutralization was consistent with ELISA results, with E44R showing an enhanced neutralization potency compared with HuJ3. However, its neutralization capacity was still lower than that of the original J3 nanobody (Table 5). A potential reason why V54E, S98W, and S102R failed to increase the binding affinity is that the side chain of valine and serine are uncharged, but glutamic acid and arginine have charged side chains. Also, tryptophan is significantly larger than serine. Such mutations may have caused a repulsive effect. Additionally, the loop regions are flexible, which may change the conformation of the loops, leading to decrease or no change to the binding affinity when mutations are introduced. As for T99R, it is likely that the mutation caused inaccurate folding and led to expression failure.

Comparisons between IsAb2.0 and IsAb1.0

To demonstrate the improvement of IsAb2.0, IsAb1.0 was applied to design HuJ3 and its predictions validated by the experimental methods (Fig. S8). The results of IsAb1.0 are provided in the Supplementary Materials. Figure 6 illustrated that the differences between the 3D structures of HuJ3-gp120 complexes generated by IsAb1.0 and IsAb2.0 were minimal. The modeling and docking results from both versions were closely aligned. However, IsAb2.0 does not require modeled structures to have homologous proteins as template, allowing it to generate 3D structures for novel antibodies and antigens. Additionally, IsAb2.0 does not require pre-existing binding information of the antibody and antigen, making it broadly applicable in various scenarios. The integration of AlphaFold-Multimer in IsAb2.0 enables the direct generation of antibody–antigen complexes from their sequences, achieving similar accuracy to complexes generated by IsAb1.0. This use of AlphaFold-Multimer effectively replaces the processes of homology modeling and global docking in IsAb1.0, significantly simplifying the protocol. Figure 5A showed that all four mutations predicted by IsAb1.0 exhibited either the same or lower binding affinity compared with HuJ3. In contrast, as discussed previously, one of the five predictions from IsAb2.0 was shown to increase the HuJ3-gp120 binding affinity, indicating that IsAb2.0 achieves higher prediction accuracy than IsAb1.0. Since protein–protein interaction can cause structure changes, Single State Design does not account for the conformational plasticity of proteins, leading to inaccurate predictions. Hence, FlexddG that uses the backrub method to generate ensemble models and allows structure flexibility can achieve higher accuracy in increasing antibody binding affinity.

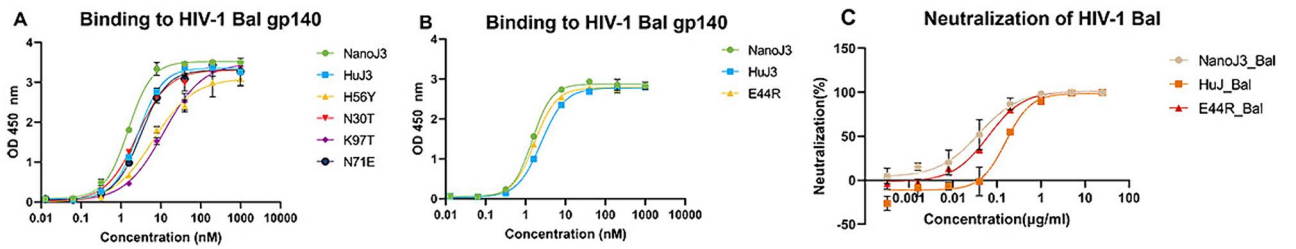

Figure 5. Experimental validation of designed HuJ3 variants. (A) Binding affinity of HuJ3 variants predicted by IsAb1.0 for HIV-1 Bal gp140. (B) Binding affinity of HuJ3 variants predicted by IsAb2.0 for HIV-1 Bal gp140. (C) Percent neutralization of HIV-1 Bal by HuJ3 variants predicted by IsAb2.0.

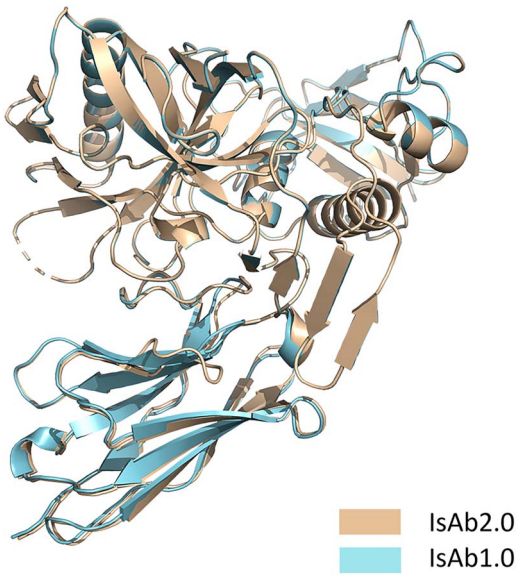

Figure 6. Alignment of the HuJ3-gp120 complexes generated by IsAb1.0 and IsAb2.0.

Discussion

In this study, we developed an advanced antibody design protocol, IsAb2.0, by integrating state-of-the-art AI-based and physical methods. IsAb2.0 can be utilized for the design of humanized antibodies and nanobodies, enabling the construction of more accurate models of antibody–antigen complexes without the need for template and additional binding information, using AlphaFold-Multimer. Furthermore, IsAb2.0 employs more accurate method FlexddG for predicting mutations. To validate our protocol and apply J3 to treat HIV-1 in human patients, we first humanized the llama nanobody J3 to HuJ3 to decrease its immunogenicity. We then applied the IsAb2.0 to design HuJ3 and increased its binding affinity with gp120. The IsAb2.0 modeled HuJ3-gp120 complex. Based on the HuJ3-gp120 complex, the IsAb2.0 predicted six hotspots on HuJ3 and found five potential mutations that could increase the binding affinity of the HuJ3-gp120 complex. All five mutations had been repeated by BioLuminate, and four of the five mutations gave the same results as our protocol, showing that IsAb2.0 yielded similar predictions as the well-known commercial protein design software. Among the predicted mutations, experimental validation by ELISA proved that E44R could increase the binding affinity of HuJ3-gp120. HIV-1 neutralization assays also showed the increased neutralization capacity of E44 compared with HuJ3. Although the prediction accuracy did not reach our expectations, these results support that IsAb2.0 has the potential to be applied to antibody design, including the design of nanobodies and humanized antibodies. In the future, we will further refine our protocol and resolve its limitations. For example, we will aim to improve the accuracy of point mutation which, at this point, is not sufficient. This may be due in part to the inaccuracies of the score function in FlexddG. The current score function may not evaluate the mutations accurately, leading to point mutation failure. Also, the current point mutation program does not consider the rationale of mutations, which contributes to the failure of prediction. The process of running FlexddG is complicated, leading to a prohibitively expensive computing time. Another limitation of our protocol is that it cannot run automatically. In some steps, to maintain prediction accuracy, it still requires users to choose the results manually. This problem means the protocol is not yet entirely user-friendly, especially for users who are not experienced in the area of antibody engineering.

To solve the limitations mentioned above, we will develop a novel score function suitable for FlexddG that enhances the precision of the result evaluation. We will also examine the rationale of amino acids substation during mutation. In addition, we will create a machine-learning-based method than can predict amino acids probability in each position based on the difference between the framework and CDR regions. This method will combine with FlexddG and build an advanced program. The advanced program will let FlexddG mutate the specified position to the amino acids with high occurrence probability which can improve the efficiency of FlexddG. Another solution is to leverage the advantages of graph neural networks in handling 3D structures to develop an AI-based antibody design model. This AI-based model can more effectively represent the interactions between antibodies and antigens, enabling it to learn the patterns of the interactions within complexes and design the antibodies with higher accuracy. To achieve the automation of the protocol and maintain prediction accuracy, we will create more advanced programs or modify existing methods to let the protocol select the correct results by itself.

Another limitation is that AlphaFold-Multimer often fails to accurately model secondary structures, which can affect local docking accuracy. Users can evaluate the quality of structures modeled by AlphaFold-Multimer based on the pLDDT scores of residues. According to the AlphaFold official document, regions with pLDDT scores lower than 70 have low confidence and should be treated cautiously. We suggest that if pLDDT scores of all regions, or the interaction critical regions, are higher than 70, users can directly input complex modeled by AlphaFold-Multer into SnugDock. In our case, some secondary structures of both the antibody and antigen were not accurately predicted. Hence, structure refinement is necessary to achieve higher structural accuracy for the antibody or antigen. The best way to perform structure refinement is to replace the structures of both the

antibody and antigen modeled by AlphaFold-Multimer with high-resolution crystal structures. If the crystal structures are unavailable, users can generate accurate 3D structures by some powerful modeling program, e.g. SWISS-MODEL.

Overall, the improvements made to the IsAb2.0 protocol has important implications for the future of antibody design. A standard, user-friendly antibody design tool could significantly improve and accelerate the development of antibody therapies across several fields, including the treatment of cancers, viral infections, and addiction. By developing IsAb2.0 protocol to include methods of nanobody and humanized antibody design while reducing its reliance on available structural data, we aim to expand treatment options for such conditions and reduce barriers to antibody engineering.

Key Points

- The novel antibody design protocol, IsAb 2.0, integrates AlphaFold-Multimer 2.3/3.0 to model the 3D structure of a humanized llama nanobody J3 and construct the binding complex of HuJ3-gp120. It can also predict the possible hotspots of HuJ3 and the potential mutations that can increase the binding affinity of HuJ3-gp120.
- Four of the five potential mutations predicted by our protocol are validated by powerful protein engineering commercial software. ELISA confirmed that one of the predicted mutations can increase the binding affinity of HuJ3 for gp120.
- The antibody design protocol provides comprehensive step-by-step instructions for users and addresses some challenges in antibody design. It has the potential to design antibodies with increased binding affinity, assisting researchers in meeting the high demand for therapeutic antibodies.

Acknowledgements

The authors thank the members of the University of Pittsburgh Center for AIDS Research Viral Pathogenesis and Persistence Laboratory for their efforts in assessing the HIV-1 neutralization capacity of the engineered binders. This research was supported in part by the University of Pittsburgh Center for Research Computing, RRID:SCR_022735, through the resources provided. Specifically, this work used the H2P cluster, which is supported by NSF award number OAC-2117681.

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Conflict of interest: The authors declare that they have no conflict of interest.

Author contributions

Z.F. and W.L. designed the project. T.L. and Z.-Y. S. tested the code. M.G.H. and X.C. conducted the experiments. T.L., M.G.H., and Z.-Y.S. prepared the figures and wrote the manuscript. All authors read and approved the final manuscript.

Funding

The authors would like to acknowledge the funding support to the Xie laboratory/Center from the National Institutes of Health National Institute on Drug Abuse (R01 DA052329 and P30 DA035778A1), and I4C 2.0: Immunotherapy for Cure (1UM1AI164556).

Data availability

Single State Design is accessible from http://www.meilerlab.org/index.php/rosetta-tutorials. FlexddG is accessible from https://github.com/Kortemme-Lab/flex_ddG_tutorial. All the data presented in this work can be found at https://github.com/zeysun/IsAb2.0.

References

1. Weitzner BD, Kuroda D, Marze N. *et al.* Blind prediction performance of RosettaAntibody 3.0: grafting, relaxation, kinematic loop modeling, and full CDR optimization. *Proteins* 2014;**82**:1611–23. https://doi.org/10.1002/prot.24534.
2. Paul S, Konig MF, Pardoll DM. *et al.* Cancer therapy with antibodies. *Nat Rev Cancer* 2024;**24**:399–426. https://doi.org/10.1038/s41568-024-00690-x.
3. Pantaleo G, Correia B, Fenwick C. *et al.* Antibodies to combat viral infections: development strategies and progress. *Nat Rev Drug Discov* 2022;**21**:676–96. https://doi.org/10.1038/s41573-022-00495-3.
4. Shen X, Kosten TR. Immunotherapy for drug abuse. *CNS Neurol Disord Drug Targets* 2011;**10**:876–9. https://doi.org/10.2174/187152711799219352.
5. Puthenpurail A, Rathi H, Nauli SM. *et al.* A brief synopsis of monoclonal antibody for the treatment of various groups of diseases. *World J Pharm Pharm Sci* 2021;**10**:14–22.
6. Chen J, Gao K, Wang R. *et al.* Review of COVID-19 antibody therapies. *Annu Rev Biophys* 2021;**50**:1–30. https://doi.org/10.1146/annurev-biophys-062920-063711.
7. Bao G, Tang M, Zhao J. *et al.* Nanobody: a promising toolkit for molecular imaging and disease therapy. *EJNMMI Res* 2021;**11**:6. https://doi.org/10.1186/s13550-021-00750-5.
8. Yang EY, Shah K. Nanobodies: next generation of cancer diagnostics and therapeutics. *Front Oncol* 2020;**10**:10. https://doi.org/10.3389/fonc.2020.01182.
9. Tolmachev V, Orlova A, Sörensen J. The emerging role of radionuclide molecular imaging of HER2 expression in breast cancer. *Semin Cancer Biol* 2021;**72**:185–97. https://doi.org/10.1016/j.semcancer.2020.10.005.
10. Levi-Schaffer F, de Marco A. Coronavirus disease 2019 and the revival of passive immunization: antibody therapy for inhibiting severe acute respiratory syndrome coronavirus 2 and preventing host cell infection: IUPHAR review: 31. *Br J Pharmacol* 2021;**178**:3359–72. https://doi.org/10.1111/bph.15359.
11. Wang J, Kang G, Yuan H. *et al.* Research progress and applications of multivalent, multispecific and modified nanobodies for disease treatment. *Front Immunol* 2022;**12**:12. https://doi.org/10.3389/fimmu.2021.838082.
12. Harris CT, Cohen S. Reducing immunogenicity by design: approaches to minimize immunogenicity of monoclonal antibodies. *BioDrugs* 2024;**38**:205–26. https://doi.org/10.1007/s40259-023-00641-2.
13. Marsh JA, Teichmann SA. Structure, dynamics, assembly, and evolution of protein complexes. *Annu Rev Biochem* 2015;**84**:551–75. https://doi.org/10.1146/annurev-biochem-060614-034142.

14. Song CM, Lim SJ, Tong JC. Recent advances in computer-aided drug design. *Brief Bioinform* 2009;**10**:579–91. https://doi.org/10.1093/bib/bbp023.
15. Schymkowitz J, Borg J, Stricher F. *et al.* The FoldX web server: an online force field. *Nucleic Acids Res* 2005;**33**:W382–8. https://doi.org/10.1093/nar/gki387.
16. Vangone A, Bonvin AMJJ. Contacts-based prediction of binding affinity in protein–protein complexes. *Elife* 2015;**4**:e07454. https://doi.org/10.7554/eLife.07454.
17. Xue LC, Rodrigues JP, Kastritis PL. *et al.* PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* 2016;**32**:3676–8. https://doi.org/10.1093/bioinformatics/btw514.
18. Liang T, Chen H, Yuan J. *et al.* IsAb: a computational protocol for antibody design. *Brief Bioinform* 2021;**22**:bbab143. https://doi.org/10.1093/bib/bbab143.
19. Joubbi S, Micheli A, Milazzo P. *et al.* Antibody design using deep learning: from sequence and structure design to affinity maturation. *Brief Bioinform* 2024;**25**:bbae307. https://doi.org/10.1093/bib/bbae307.
20. Hummer AM, Abanades B, Deane CM. Advances in computational structure-based antibody design. *Curr Opin Struct Biol* 2022;**74**:102379. https://doi.org/10.1016/j.sbi.2022.102379.
21. Richard E, Michael ON, Alexander P. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2022:2021.2010.2004.463034.
22. Barlow KA, Ó S, Thompson S. *et al.* Flex ddG: Rosetta Ensemble-based estimation of changes in protein–protein binding affinity upon mutation. *J Phys Chem B* 2018;**122**:5389–99.
23. Barré-Sinoussi F, Chermann JC, Rey F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983;**220**:868–71. https://doi.org/10.1126/science.6189183.
24. Gallo RC, Salahuddin SZ, Popovic M. *et al.* Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 1984;**224**:500–3. https://doi.org/10.1126/science.6200936.
25. Klatzmann D, Champagne E, Chamaret S. *et al.* T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 1984;**312**:767–8. https://doi.org/10.1038/312767a0.
26. Dalgleish AG, Beverley PCL, Clapham PR. *et al.* The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 1984;**312**:763–7. https://doi.org/10.1038/312763a0.
27. Chan DC, Fass D, Berger JM. *et al.* Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 1997;**89**:263–73. https://doi.org/10.1016/S0092-8674(00)80205-6.
28. Huang J, Ofek G, Laub L. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* 2012;**491**:406–12. https://doi.org/10.1038/nature11544.
29. Zhou T, Chen L, Gorman J. *et al.* Structural basis for llama nanobody recognition and neutralization of HIV-1 at the CD4-binding site. *Structure* 2022;**30**:862–875.e4. https://doi.org/10.1016/j.str.2022.03.012.
30. Zhu K, Day T, Warshaviak D. *et al.* Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins* 2014;**82**:1646–55. https://doi.org/10.1002/prot.24551.
31. Salam NK, Adzhigirey M, Sherman W. *et al.* Structure-based approach to the prediction of disulfide bonds in proteins. *Protein Eng Des Sel* 2014;**27**:365–74. https://doi.org/10.1093/protein/gzu017.
32. Beard H, Cholleti A, Pearlman D. *et al.* Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. *PloS One* 2013;**8**:e82849. https://doi.org/10.1371/journal.pone.0082849.
33. Cianfrocco MA, Wong-Barnum M, Youn C. *et al.* *COSMIC2: A Science Gateway for Cryo-Electron Microscopy Structure Determination. Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*. New Orleans, LA, USA: Association for Computing Machinery, 2017, Article 22.
34. Sircar A, Gray JJ. SnugDock: Paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PLoS Comput Biol* 2010;**6**:e1000644. https://doi.org/10.1371/journal.pcbi.1000644.
35. Lyskov S, Gray JJ. The RosettaDock server for local protein–protein docking. *Nucleic Acids Res* 2008;**36**:W233–8. https://doi.org/10.1093/nar/gkn216.
36. Lyskov S, Chou F-C, Conchúir SÓ. *et al.* Serverification of molecular Modeling applications: the Rosetta online server that includes everyone (ROSIE). *PloS One* 2013;**8**:e63906. https://doi.org/10.1371/journal.pone.0063906.
37. Marze NA, Roy Burman SS, Sheffler W. *et al.* Efficient flexible backbone protein-protein docking for challenging targets. *Bioinformatics* 2018;**34**:3461–9. https://doi.org/10.1093/bioinformatics/bty355.
38. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor Interface. *Science* 1995;**267**:383–6. https://doi.org/10.1126/science.7529940.
39. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein–protein complexes. *Proc Natl Acad Sci* 2002;**99**:14116–21. https://doi.org/10.1073/pnas.202485799.
40. Kortemme T, Kim DE, Baker D. Computational alanine scanning of protein-protein interfaces. *Sci STKE* 2004;**2004**:pl2–2. https://doi.org/10.1126/stke.2192004pl2.
41. Chen C, Sun Z, Liu X. *et al.* Protocol for constructing large size human antibody heavy chain variable domain (V(H)) library and selection of SARS-CoV-2 neutralizing antibody domains. *STAR Protoc* 2021;**2**:100617. https://doi.org/10.1016/j.xpro.2021.100617.
42. Waterhouse A, Bertoni M, Bienert S. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;**46**:W296–w303. https://doi.org/10.1093/nar/gky427.
43. Desta IT, Porter KA, Xia B. *et al.* Performance and its limits in rigid body protein-protein docking. *Structure* 2020;**28**:1071–1081.e3. https://doi.org/10.1016/j.str.2020.06.006.
44. Vajda S, Yueh C, Beglov D. *et al.* New additions to the ClusPro server motivated by CAPRI. *Proteins* 2017;**85**:435–44. https://doi.org/10.1002/prot.25219.
45. Kozakov D, Hall DR, Xia B. *et al.* The ClusPro web server for protein–protein docking. *Nat Protoc* 2017;**12**:255–78. https://doi.org/10.1038/nprot.2016.169.
46. Kozakov D, Beglov D, Bohnuud T. *et al.* How good is automated protein docking? *Proteins* 2013;**81**:2159–66. https://doi.org/10.1002/prot.24403.
47. Brenke R, Hall DR, Chuang GY. *et al.* Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics* 2012;**28**:2608–14. https://doi.org/10.1093/bioinformatics/bts493.
48. Chaudhury S, Berrondo M, Weitzner BD. *et al.* Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PloS One* 2011;**6**:e22477. https://doi.org/10.1371/journal.pone.0022477.
49. Tyka MD, Keedy DA, André I. *et al.* Alternate states of proteins revealed by detailed energy landscape mapping. *J Mol Biol* 2011;**405**:607–18. https://doi.org/10.1016/j.jmb.2010.11.008.
50. Wang C, Bradley P, Baker D. Protein-protein docking with backbone flexibility. *J Mol Biol* 2007;**373**:503–19. https://doi.org/10.1016/j.jmb.2007.07.050.